AREA PRESERVING SURFACE DIFFEOMORPHISMS WITH POLYNOMIAL DECAY OF CORRELATIONS ARE UBIQUITOUS

Y. PESIN, S. SENTI, AND F. SHAHIDI

Dedicated to the memory of Anatole Katok, our friend and mentor

ABSTRACT. We show that any smooth compact connected and oriented surface admits an area preserving $C^{1+\beta}$ diffeomorphism with non-zero Lyapunov exponents which is Bernoulli and has polynomial decay of correlations. We establish both upper and lower polynomial bounds on correlations. In addition, we show that this diffeomorphism satisfies the Central Limit Theorem and has the Large Deviation Property. Finally, we show that the diffeomorphism we constructed possesses a unique hyperbolic Bernoulli measure of maximal entropy with respect to which it has exponential decay of correlations.

1. INTRODUCTION

A classical problem in smooth dynamics known as the *smooth realization problem* asks whether there is a diffeomorphism f of a compact smooth manifold M which has a prescribed collection of ergodic properties with respect to a *natural* invariant measure μ such as the Riemannian volume (or a more general smooth measure, i.e., a measure that is equivalent to volume). Other interesting measures to consider include the measure of maximal entropy. A yet more interesting but substantially more difficult version of the smooth realization problem is to construct a volume preserving diffeomorphism f with prescribed ergodic properties on any given smooth manifold M. Starting with the basic ergodic property – ergodicity – Anosov and Katok [1] constructed an example of a volume preserving *ergodic* C^∞ map with some additional metric properties. Katok [12] gave an example of area preserving C^∞ diffeomorphism with non-zero Lyapunov exponents on any surface which is *Bernoulli* (see the definitions in the next section). Later Brin,

Date: July 8, 2023.

1991 *Mathematics Subject Classification.* 37D25, 37D35, 37A25, 37E30.

Y.P. is supported by the NSF grant DMS-2153053.

Feldman, and Katok [4] and then Brin [3] extended this result by constructing a volume preserving C^∞ diffeomorphism, which is Bernoulli, on any Riemannian manifold of dimension ≥ 5. In this example the map has all but one non-zero Lyapunov exponents. Finally, Dolgopyat and Pesin [6] constructed a volume preserving C^∞ Bernoulli diffeomorphism with *non-zero Lyapunov exponents* on any Riemannian manifold of dimension ≥ 2.

It is natural to ask if a compact smooth manifold admits a volume preserving Bernoulli diffeomorphism with non-zero Lyapunov exponents that enjoys other important statistical properties such as exponential or polynomial decay of correlations (that is rate of mixing), the Central Limit Theorem, and the Large Deviations property (all three with respect to a *natural* class of observables, e.g., functions which are Hölder continuous).

In one dimensional dynamics the famous Mauneville-Pomeau map [19] (with some modifications) provide some examples of a map with an indifferent fixed point preserving a measure which is absolutely continuous with respect to the (one-dimensional) Lebesgue measure. With respect to this measure the decay of correlations is polynomial, the Central Limit Theorem is satisfied, and the map has Large Deviation (with respect to the class of Hölder continuous observables; see [8, 10, 15]).

In the two dimensional case several examples have been constructed of area preserving maps on the 2-torus with polynomial decay of correlations (and in some cases with sharp polynomial lower and upper bounds), see [5, 7, 16, 15]. Also, starting from the work of Young, [24, 25], some general techniques for obtaining polynomial decay of correlations for maps admitting Young tower have been developed by Gouëzel [8] and Sarig, [22], see also [18, 23]. Also in [11] polynomial lower and upper bounds were obtained for almost Anosov diffeomorphisms. This class of diffeomorphisms was introduced by Hu in [10] and they preserve the Sinai-Ruelle-Bowen measure (which may not be volume) and the local dynamics of these maps is quite different than the one of the Katok map.

In the present paper we show that any surface admits an area preserving $C^{1+\beta}$ diffeomorphism with non-zero Lyapunov exponents which is Bernoulli and has polynomial decay of correlations – more precisely, it allows polynomial lower and upper bounds. It also satisfies the Central Limit Theorem and has Large Deviation.

Interestingly enough the map we construct also has the unique measure of maximal entropy with exponential decay of correlations. Thus we show that any surface allows a $C^{1+\beta}$ diffeomorphism with the unique measure of maximal entropy with respect to which it has non-zero Lyapunov exponents, is Bernoulli, has exponential decay of correlations, and satisfies the Central Limit Theorem.

While our proof follows the basic scheme of Katok construction in [12], we have to make many substantial changes which we outline here. Starting with a linear automorphism of the 2-torus T^2 which has 4 fixed points, we slow down trajectories in sufficiently small neighborhoods of these points and then *correct* the resulting map to obtain an area preserving diffeomorphism f_{T^2}. We stress that while in the Katok construction the slow down function can be chosen to be infinitely and arbitrarily flat at 0 (guaranteeing that f_{T^2} is C^∞), to ensure that f_{T^2} has polynomial decay of correlations, we have to choose the slow down function to be polynomial at 0.[1] This results in f_{T^2} to be of class $C^{2+2\kappa}$. It also has non-zero Lyapunov exponents and is Bernoulli.

To show that f_{T^2} has polynomial decay of correlations we represent this map as a Young diffeomorphism and study the symbolic map on the corresponding Young tower (see Section 5). This map preserves the measure that is the lift of the area (see [23, 24]). The decay of correlations of this symbolic map has been studied extensively (see for example, [8, 18, 22, 23, 24, 25]) and is tied to the decay of the tail of the return time (see Proposition 10.2). Thus to establish polynomial upper and lower bounds on the decay of correlations we need to obtain both upper and lower bounds on decay of the tail. This requires a deep understanding of the behavior of trajectories in the slow down domain and is done in Sections 6-8 which constitute the technically most difficult part of the work.

Our next step is to carry over the map of the torus to a map on a given surface. To achieve this we follow the approach in [12] and obtain a $C^{2+2\kappa}$ area preserving Bernoulli diffeomorphism f_{D^2} of the two dimensional disk with non-zero Lyapunov exponents which is identity on the boundary of the disk. In the original Katok's construction the map f_{D^2} is C^∞ and is infinitely and arbitrarily flat at the boundary of the disk[2], which allows one to apply some standard results to carry over f_{D^2} to a diffeomorphism f_M of a surface M. In our case f_{D^2} is only finitely flat at the boundary of the disk and we develop a specific construction of a diffeomorphism from the interior of the disk onto an open simply connected and dense subset of M, which extends to a

[1]We believe that the rate of decay of correlations of f_{T^2} is determined by the rate of decay of the slow down function at 0.

[2]This is due to the fact that the slow down function can be chosen arbitrarily flat at 0.

homeomorphism from the closed disk onto M and is area preserving. This diffeomorphism moves f_{D^2} to an area preserving Bernoulli diffeomorphism f_M of the surface with non-zero Lyapunov exponents, see Section 9.

Our next step is to use the conjugacy map and a representation of f_{T^2} as a Young diffeomorphism to obtain a similar representation for f_M, see Section 10. Now to obtain an upper polynomial bound for decay of correlation we use the results in [25] and [18] to choose an appropriate class of observables, which includes all Hölder continuous functions on the surface. To obtain a lower bound we use the result in [23] (which is a hyperbolic version of the result in [8]) and the class of Hölder continuous observables on the surface which vanish inside small neighborhoods of the fixed points, see Section 10.

We stress that the exponents in our polynomial lower and upper bounds are different. This is a result of our estimates on the behavior of trajectories of the Katok map in the slow down domain (see Section 6) and may be an artifact of our techniques. However, there is a particular class of observables for which our method gives the same exponent in the polynomial lower and upper bounds, see Statement 3(b)(ii) of the Main theorem 3.1.

Representing the map f_M as a Young tower also allows us to establish the Central Limit Theorem using results in [8, 14] as well as Polynomial Large Deviation using results in [17].

The paper is organized as follows. After we provide some definitions in the next section we state our Main Theorem in Section 3. In Section 4 we construct the map f_{T^2} on the 2-torus and state some of its properties including its class of smoothness. In Section 5 we recall the definition of Young diffeomorphisms and describe a representation of f_{T^2} as a Young diffeomorphism. The proof of the main result, Theorem 3.1, occupies Sections 6 through 10. In Section 6, we prove some technical results that establish new crucial properties of the slow down map. In Sections 7 and 8 we obtain polynomial respectively lower and upper bounds on the tail of the return time for the Katok map f_{T^2}. In Section 9 we show how to carry over the Katok map of the torus to a diffeomorphism of a given surface. Finally, in Section 10 we complete the proof of the main result.

Acknowledments. The first and the second authors would like to thank Bernoulli Center (CIB, Lausanne, Switzerland) where part of the work was done for their hospitality. S.S. was supported by the CNPQ grant. The authors also would like to thank the referee for valuable comments and suggestions.

2. Definitions and notations

Let X be a measurable space and $T : X \to X$ a measurable invertible transformation preserving a measure μ. For reader's convenience we recall the definitions of some properties of the map which are of interest to us in the paper.

2.1. **The Bernoulli property.** We say that (T, μ) has the *Bernoulli property* if it is metrically isomorphic to the Bernoulli shift (σ, κ) associated to some Lebesgue space (Y, ν), so that ν is metrically isomorphic to the Lebesgue measure on an interval together with at most countably many atoms and κ is given as the direct product of $\mathbb{Z}$ copies of ν on $Y^{\mathbb{Z}}$.

2.2. **Decay of correlations.** Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two classes of real-valued functions on X called *observables*. For $h_1 \in \mathcal{H}_1$ and $h_2 \in \mathcal{H}_2$ define the correlation function

$$\mathrm{Cor}_n(h_1, h_2) := \int h_1(T^n(x))h_2(x)\, d\mu - \int h_1(x)\, d\mu \int h_2(x)\, d\mu.$$

We say that T has *polynomial decay of correlations* (more precisely, *polynomial upper bound on correlations*) with respect to classes $\mathcal{H}_1$ and $\mathcal{H}_2$ if there exists $\gamma_1 > 0$ such that for any $h_1 \in \mathcal{H}_1$, $h_2 \in \mathcal{H}_2$, and any $n > 0$,

$$|\mathrm{Cor}_n(h_1, h_2)| \le C n^{-\gamma_1},$$

where $C = C(h_1, h_2) > 0$ is a constant.

We say that T admits a *polynomial lower bound on correlations* with respect to classes $\mathcal{H}_1$ and $\mathcal{H}_2$ of observables if there exists $\gamma_2 > 0$ such that for any $h_1 \in \mathcal{H}_1$, $h_2 \in \mathcal{H}_2$, and any $n > 0$,

$$|\mathrm{Cor}_n(h_1, h_2)| \ge C' n^{-\gamma_2},$$

where $C' = C'(h_1, h_2) > 0$ is a constant.

We say that T has *exponential decay of correlations* with respect to classes $\mathcal{H}_1$ and $\mathcal{H}_2$ if there exists $\gamma_3 > 0$ such that for any $h_1 \in \mathcal{H}_1$, $h_2 \in \mathcal{H}_2$, and any $n > 0$,

$$|\mathrm{Cor}_n(h_1, h_2)| \le C'' e^{-\gamma_3 n},$$

where $C'' = C''(h_1, h_2) > 0$ is a constant.

2.3. **The Central Limit Theorem.** We say that T satisfies the *Central Limit Theorem (CLT)* with respect to a class $\mathcal{H}$ of observables on

X if there exists $\sigma > 0$ such that for any $h \in \mathcal{H}$ with $\int h \, d\mu = 0$ the sum

$$\frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} h(f^i(x))$$

converges in law to a normal distribution $N(0, \sigma)$.

2.4. **Large Deviation.** We say that T has *Polynomial Large Deviation* with respect to a class $\mathcal{H}$ of observables on X if there is $\beta > 0$ such that for any $h \in \mathcal{H}$, any $\varepsilon > 0$, and any sufficiently large $n > 0$

$$\mu\Big(\Big|\frac{1}{n}\sum_{i=0}^{n-1} h(T^i(x)) - \int h\Big| > \varepsilon\Big) < Kn^{-\beta},$$

where $K = K(\varepsilon, \beta, h) > 1$ is a constant.

2.5. **Lyapunov exponents.** Let $f \colon M \to M$ be a diffeomorphism of a compact smooth Riemannian manifold M. Given a point $x \in M$ and a vector $v \in T_xM$, the number

$$\chi(x, v) := \limsup_{n\to\infty} \frac{1}{n} \log \|df_x^n v\|$$

is called the *Lyapunov exponents* of v at x. One can show that for every $x \in M$ the function $\chi(x, \cdot)$ takes on finitely many values which we denote by $\chi_1(x) \le \cdots \le \chi_p(x)$, where $p = \dim M$. The functions $\chi_i(x)$, $i = 1, \dots, p$ are Borel measurable and f-invariant.

If μ is an f-invariant measure, then for μ-almost every $x \in M$ and any $v \in T_xM$,

$$\chi(x, v) = \lim_{n\to\infty} \frac{1}{n} \log \|df_x^n v\|.$$

We say that f has *nonzero Lyapunov exponents* with respect to μ or that μ is *hyperbolic* if for μ-almost every x we have that $\chi_i(x) \neq 0$, $i = 1, \dots, p$ and that $\chi_1(x) < 0$ while $\chi_p(x) > 0$.

Note that if μ is ergodic, then $\chi_i(x)$ is a constant for all $i = 1, \dots, p$ and μ-almost every $x \in M$, which we denote by $\chi_i(\mu)$.

3. Main Results

Let M be a smooth compact connected oriented surface with area m. Without loss of generality we assume that $m(M) = 1$. Given $\rho > 0$, let $C^\rho := C^\rho(M)$ be the class of all Hölder continuous functions on M with exponent ρ.

Consider a nested sequence of subsets $\{M_j\}$ that exhausts M that is $M_1 \subset M_2 \subset \cdots \subset M$ and $\bigcup_{j\ge1} M_j = M$. Given such a sequence, let $\mathcal{G} = \mathcal{G}(\{M_j\})$ be the class of observables $h \subset C^\rho$ for which there is $k = k(h)$ such that $\operatorname{supp}(h) \subset M_k$.

Theorem 3.1. *Let M be a compact smooth connected and oriented surface. There are numbers $\beta > 0$, $\rho > 0$, $\gamma_2 > \gamma_1 > 0$, and a $C^{1+\beta}$ diffeomorphism f of M preserving area m and satisfying:*

(1) *f has the Bernoulli property with respect to m;*
(2) *f has non-zero Lyapunov exponents almost everywhere with respect to m;*
(3) *f admits polynomial upper and lower bounds on correlations with respect to m; more precisely:*
   (a) *for any $h_i \in C^\rho$, $i = 1, 2$*

$$|Cor_n(h_1, h_2)| \le C_1 n^{-\gamma_1},$$

   where $C_1 = C_1(\|h_1\|_{C^\rho}, \|h_2\|_{C^\rho}) > 0$;
   (b) *if $\int h_1 dm \int h_2 dm > 0$, then there is a nested sequence of subsets $\{M_j\}$ which exhausts M such that for any $h_i \in \mathcal{G}(\{M_j\})$, $i = 1, 2$,*

$$|Cor_n(h_1, h_2)| \ge C_2 n^{-\gamma_2},$$

   where $C_2 = C_2(\|h_1\|_{C^\rho}, \|h_2\|_{C^\rho}) > 0$;
(4) *the map f satisfies the CLT for the class of observables $h \in C^\rho$, $\int h\, dm = 0$ with $\sigma = \sigma(h)$ given by*

$$\sigma^2 = -\int h^2 dm + 2\sum_{n=0}^{\infty} \int h \cdot h \circ f^n dm,$$

where $\sigma > 0$ if and only if h is not cohomologous to zero, i.e., $h \circ f \neq g \circ f - g$ for any measurable function g;
(5) *the map f has Polynomial Large Deviation with respect to the class C^ρ of observables with the constant K of the form $K = K(\|h\|_{C^\rho})\varepsilon^{-2\beta}$. In addition, for an open and dense subset of observables in C^ρ and sufficiently small $\varepsilon > 0$*

$$n^{-\beta} < m\Big(\Big|\frac{1}{n}\sum_{i=0}^{n-1} h(f^i(x)) - \int h\Big| > \varepsilon\Big)$$

for infinitely many n;
(6) *f has a unique measure of maximal entropy (MME) with respect to which it has the Bernoulli property, non-zero Lyapunov exponents almost everywhere, exponential decay of correlations and satisfies the CLT with respect to the class C^ρ of observables.*

4. A slow down map of the 2-torus

4.1. **The definition of a slow down map.** Consider the automorphism of the two-dimensional torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ given by the matrix $A := \left(\begin{smallmatrix} 5 & 8 \\ 8 & 13 \end{smallmatrix}\right)$. It has four fixed points $x_1 = (0,0)$, $x_2 = (\frac{1}{2}, 0)$, $x_3 = (0, \frac{1}{2})$, and $x_4 = (\frac{1}{2}, \frac{1}{2})$. For $i = 1, 2, 3, 4$ consider the disk $D^i_r = \{(s_1, s_2) : {s_1}^2 + {s_2}^2 \le r^2\}$ of radius r centered at x_i and set $D_r = \bigcup_{i=1}^4 D^i_r$. Here (s_1, s_2) is the coordinate system obtained from the eigendirections of A and originated at x_i. Let $\lambda > 1$ be the largest eigenvalue of A. There are $r_2 > r_1 > r_0$ such that

$$D^i_{r_0} \subset A(D^i_{r_1}), \quad A(D^i_{r_1}) \cup A^{-1}(D^i_{r_1}) \subset D^i_{r_2} \tag{1}$$

and the disks $D^i_{r_2}$ are pairwise disjoint. Fix i and consider the system of differential equations in $D^i_{r_1}$

$$\frac{ds_1}{dt} = s_1 \log \lambda, \quad \frac{ds_2}{dt} = -s_2 \log \lambda. \tag{2}$$

Observe that $A|_{D^i_{r_1}}$ is the time-1 map of the local flow generated by this system.

We choose a number $0 < \alpha < 1$, $0 < r_0 < 1$, and a function $\psi : [0,1] \to [0,1]$ satisfying:

- (K1) ψ is of class C^∞ everywhere on $(0,1]$ but at the origin;
- (K2) $\psi(u) = 1$ for $r_0 \le u \le 1$;
- (K3) $\psi'(u) > 0$ for $0 < u < r_0$;
- (K4) $\psi(u) = (u/r_0)^\alpha$ for $0 \le u \le \frac{r_0}{2}$.

Using the function ψ, we slow down trajectories of the flow by perturbing the system (2) in $D^i_{r_1}$ as follows

$$\begin{aligned} \frac{ds_1}{dt} &= \quad s_1 \psi({s_1}^2 + {s_2}^2) \log \lambda \\ \frac{ds_2}{dt} &= -\, s_2 \psi({s_1}^2 + {s_2}^2) \log \lambda. \end{aligned} \tag{3}$$

This system of differential equations generates a local flow g^i_t and we denote by g^i the time-1 map of this flow. The choices of ψ, r_0 and r_1 (see (1)) guarantee that the domain of g^i contains $D^i_{r_1}$. Furthermore, g^i is of class C^∞ in $D^i_{r_1} \setminus \{x_i\}$ and it coincides with A in some neighborhood of the boundary $\partial D^i_{r_1}$. Therefore, the map

$$G(x) = \begin{cases} A(x) & \text{if } x \in \mathbb{T}^2 \setminus D_{r_1}, \\ g^i(x) & \text{if } x \in D^i_{r_1} \end{cases} \tag{4}$$

defines a homeomorphism of the torus $\mathbb{T}^2$, which is a C^∞ diffeomorphism everywhere except at the fixed points x_i. Since $0 < \alpha < 1$, we

have that
$$\int_0^1 \frac{du}{\psi(u)} < \infty.$$
This implies that the map G preserves the probability measure
$$d\nu = q_0^{-1} q \, dm, \tag{5}$$
where m is the area and the density q is a positive C^∞ function that is infinite at x_i and is defined by
$$q(s_1, s_2) := \begin{cases} (\psi({s_1}^2 + {s_2}^2))^{-1} & \text{if } (s_1, s_2) \in D^i_{r_1}, \\ 1 & \text{in } \mathbb{T}^2 \setminus D_{r_1} \end{cases}$$
and
$$q_0 := \int_{\mathbb{T}^2} q \, dm.$$
We further perturb the map G by a coordinate change ϕ in $\mathbb{T}^2$ to obtain an area preserving map. To achieve this, define a map ϕ in $D^i_{r_1}$ by the formula
$$\phi(s_1, s_2) := \frac{1}{\sqrt{q_0({s_1}^2 + {s_2}^2)}} \left(\int_0^{{s_1}^2 + {s_2}^2} \frac{du}{\psi(u)} \right)^{1/2} (s_1, s_2) \tag{6}$$
and set $\phi = \mathrm{Id}$ in $\mathbb{T}^2 \setminus D_{r_1}$. Clearly, ϕ is a homeomorphism and is a C^∞ diffeomorphism outside the points x_1, x_2, x_3, x_4. One can show that ϕ transfers the measure ν into the area and that the map $f_{\mathbb{T}^2} = \phi \circ G \circ \phi^{-1}$ is a homeomorphism and is a C^∞ diffeomorphism outside the points x_1, x_2, x_3, x_4. It is called a *slow down map* (see [12] and also [2]). The following proposition describes some basic properties of this map.

Proposition 4.1 ([12],[2]). *The map $f_{\mathbb{T}^2}$ has the following properties:*

(1) *It is topologically conjugated to A via a homeomorphism H.*
(2) *It admits two transverse invariant continuous stable and unstable distributions $E^s(x)$ and $E^u(x)$ and for almost every point x with respect to area m it has two non-zero Lyapunov exponents, positive in the direction of $E^u(x)$ and negative in the direction of $E^s(x)$. Moreover, the only invariant measure with zero Lyapunov exponents is the atomic measure supported on the fixed points x_i.*
(3) *It admits two continuous, uniformly transverse, invariant foliations with smooth leaves which are the images under the conjugacy map of the stable and unstable foliations for A respectively.*

(4) *For every* $\varepsilon > 0$ *one can choose* $r_0 > 0$ *such that*

$$|\int_{D^i_{r_0}} \log |D f_{\mathbb{T}^2}|E^u| \, dm - \log \lambda| < \varepsilon.$$

(5) *It is ergodic with respect to the area* m.

The following proposition establishes regularity of the map $f_{\mathbb{T}^2}$.

Proposition 4.2. *The map* $f_{\mathbb{T}^2}$ *is of class of smoothness* $C^{2+2\kappa}$, *where* $\kappa = \frac{\alpha}{1-\alpha}$.

Proof. Fix $i \in \{1, 2, 3, 4\}$ and consider the vector field in $D^i_{r_1}$ given by the right-hand side of (2). It is Hamiltonian with respect to the area and the Hamiltonian function $H_1(s_1, s_2) = s_1 s_2 \log \lambda$. The vector field given by (3) is obtained from (2) by a time change and hence, is also Hamiltonian with respect to the measure ν (see (5)) and the same Hamiltonian function. The map $f_{\mathbb{T}^2}$ is conjugate via φ (see (6)) to the time-1 map of the flow generated by (3). Since $\phi_* \nu = m$, $f_{\mathbb{T}^2}$ is the time-1 map of the flow which is Hamiltonian with respect to the area and the Hamiltonian function $H_2 = H_1 \circ \phi^{-1}$. Using (6), we find that H_2 can be given in $D^i_{r_1}$ as follows (see [12]):

$$H_2(s_1, s_2) = \frac{s_1 s_2 h(\sqrt{s_1^2 + s_2^2})}{s_1^2 + s_2^2} \log \lambda,$$

where by $(K4)$, $h(u) = u^{\frac{2}{1-\alpha}}$ and $u = s_1^2 + s_2^2$. To prove that $f_{\mathbb{T}^2}$ is of the desired class of smoothness we will show that the Hamiltonian H_2 has Hölder continuous partial derivatives of second order with Hölder exponent 2κ. To this end we consider the function $g(x, y) = xy(x^2 + y^2)^\kappa$ with $\kappa = \frac{\alpha}{1-\alpha}$ and show that g has Hölder continuous partial derivatives of second order with Hölder exponent 2κ. Note that g is of class C^∞ except for $(x, y) = (0, 0)$, so we only need to show Hölder continuity of partial derivatives at the origin. Note also that the function g is symmetric, so we only show that $\frac{\partial^2 g}{\partial x^2}$ and $\frac{\partial^2 g}{\partial x \partial y}$ are Hölder continuous. Since $\frac{\partial g}{\partial x}(0, 0) = \lim\limits_{\Delta x \to 0} \frac{g(\Delta x, 0) - g(0,0)}{\Delta x} = 0$, we have that

$$\frac{\partial g}{\partial x} = \begin{cases} y(x^2 + y^2)^\kappa + 2\kappa x^2 y (x^2 + y^2)^{\kappa - 1}, & (x, y) \neq (0, 0); \\ 0 & (x, y) = (0, 0). \end{cases}$$

Note that

$$\frac{\partial^2 g}{\partial x \partial y}(0, 0) = \lim_{\Delta y \to 0} \frac{\frac{\partial g}{\partial x}(0, \Delta y) - \frac{\partial g}{\partial x}(0, 0)}{\Delta y} = \lim_{\Delta y \to 0} \frac{\Delta y (\Delta y)^{2\kappa} - 0}{\Delta y} = 0$$

and hence,

$$\frac{\partial^2 g}{\partial x \partial y} = \begin{cases} (1+2\kappa)(x^2+y^2)^{\kappa} + \\ \quad +4(\kappa-1)x^2y^2(x^2+y^2)^{\kappa-2} & (x,y) \neq (0,0); \\ 0 & (x,y) = (0,0). \end{cases}$$

Since the function $\frac{\partial^2 g}{\partial x\partial y}$ is differentiable for all $(x,y) \neq (0,0)$, it is Hölder continuous for all pairs of nonzero points (x,y). It remains to show Hölder continuity for pairs of points one of which is zero. It is easy to see that

$$\Big|\frac{\partial^2 g}{\partial x\partial y}(x,y) - \frac{\partial^2 g}{\partial x\partial y}(0,0)\Big| \le K(x^2+y^2)^{\kappa} = Kd((x,y),(0,0))^{2\kappa},$$

where $K > 0$ and d denotes the usual distance. Thus $\frac{\partial^2 g}{\partial x\partial y}$ is Hölder continuous with Hölder exponent 2κ.

Now we consider $\frac{\partial^2 g}{\partial x^2}$. Observe that $\frac{\partial^2 g}{\partial x^2}(0,0) = 0$ and

$$\frac{\partial^2 g}{\partial x^2} = \begin{cases} 6\kappa xy(x^2+y^2)^{\kappa-1} + 4(\kappa-1)x^3y(x^2+y^2)^{\kappa-2} & (x,y) \neq (0,0); \\ 0, & (x,y) = (0,0). \end{cases}$$

It is easy to see that

$$\Big|\frac{\partial^2 g}{\partial x^2}(x,y) - \frac{\partial^2 g}{\partial x^2}(0,0)\Big| \le K(x^2+y^2)^{\kappa} = Kd((x,y),(0,0))^{2\kappa}.$$

Hence, $\frac{\partial^2 g}{\partial x^2}$ is Hölder continuous with Hölder exponent 2κ. □

5. Proof of Theorem 3.1: representing $f_{\mathbb{T}^2}$ as a Young diffeomorphism

5.1. **Young diffeomorphisms.** Let $f : M \to M$ be a $C^{1+\epsilon}$ diffeomorphism of a compact smooth Riemannian manifold M. Following [24] we describe a collection of conditions on the map f. [3]

An embedded C^1-disk $\gamma \subset M$ is called an *unstable disk* (respectively, a *stable disk*) if for all $x, y \in \gamma$ we have that $d(f^{-n}(x), f^{-n}(y)) \to 0$ (respectively, $d(f^n(x), f^n(y)) \to 0$) as $n \to +\infty$. A collection of embedded C^1 disks $\Gamma^u = \{\gamma^u\}$ is called a *continuous family of unstable disks* if there exists a homeomorphism $\Phi : K^s \times D^u \to \cup\gamma^u$ satisfying:

- $K^s \subset M$ is a Borel subset and $D^u \subset \mathbb{R}^d$ is the closed unit disk for some $d < \dim M$;

[3] Our requirements are slightly different than those in [24] since we do not assume the inducing domain Λ to be compact.

- $x \to \Phi|\{x\} \times D^u$ is a continuous map from K^s to the space of C^1 embeddings of D^u into M which can be extended to a continuous map of the closure $\overline{K^s}$;
- $\gamma^u = \Phi(\{x\} \times D^u)$ is an unstable disk.

A *continuous family of stable disks* is defined similarly.

We allow the sets K^s to be non-compact in order to deal with overlaps which appear in most known examples including the Katok map.

A set $\Lambda \subset M$ has *hyperbolic product structure* if there exists a continuous family $\Gamma^u = \{\gamma^u\}$ of unstable disks γ^u and a continuous family $\Gamma^s = \{\gamma^s\}$ of stable disks γ^s such that

- $\dim \gamma^s + \dim \gamma^u = \dim M$;
- the γ^u-disks are transversal to γ^s-disks with an angle uniformly bounded away from 0;
- each γ^u-disks intersects each γ^s-disk at exactly one point;
- $\Lambda = (\cup \gamma^u) \cap (\cup \gamma^s)$.

A subset $\Lambda_0 \subset \Lambda$ is called an *s-subset* if it has hyperbolic product structure and is defined by the same family Γ^u of unstable disks as Λ and a continuous subfamily $\Gamma_0^s \subset \Gamma^s$ of stable disks. A *u-subset* is defined analogously.

We define the *s-closure* $scl(\Lambda_0)$ of an *s-subset* $\Lambda_0 \subset \Lambda$ by

$$scl(\Lambda_0) := \bigcup_{x \in \overline{\Lambda_0 \cap \gamma^u}} \gamma^s(x) \cap \Lambda$$

and the *u-closure* $ucl(\Lambda_1)$ of a given *u-subset* $\Lambda_1 \subset \Lambda$ similarly:

$$ucl(\Lambda_1) := \bigcup_{x \in \overline{\Lambda_1 \cap \gamma^s}} \gamma^u(x) \cap \Lambda.$$

Assume the map f satisfies the following conditions:

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

$$\Lambda_i^s := \bigcup_{\gamma \in \Gamma_i^s} \big(\gamma \cap \Lambda\big) \subset \Lambda \tag{7}$$

are pairwise disjoint and satisfy:

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

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

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

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

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

(Y2) The sets Λ_i^u are pairwise disjoint.

For any $x \in \Lambda_i^s$ define the *inducing time* by $\tau(x) := \tau_i$ and the *induced map* $\tilde{f} : \bigcup_{i\in\mathbb{N}} \Lambda_i^s \to \Lambda$ by

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

(Y3) There exists $0 < a < 1$ such that for any $i \in \mathbb{N}$ we have:

(a) For $x \in \Lambda_i^s$ and $y \in \gamma^s(x)$,
$$d(\tilde{f}(x), \tilde{f}(y)) \le a\, d(x, y);$$

(b) For $x \in \Lambda_i^s$ and $y \in \gamma^u(x) \cap \Lambda_i^s$,
$$d(x, y) \le a\, d(\tilde{f}(x), \tilde{f}(y)).$$

For $x \in \Lambda$ let $Jacf(x) = \det|Df|_{E^u(x)}|$ and $Jac\tilde{f}(x) = \det|D\tilde{f}|_{E^u(x)}|$ denote the Jacobian of $Df|_{E^u(x)}$ and $D\tilde{f}|_{E^u(x)}$ respectively.

(Y4) There exist $c > 0$ and $0 < b < 1$ such that:

(a) For all $n \ge 0$, $x \in \tilde{f}^{-n}(\bigcup_{i\in\mathbb{N}} \Lambda_i^s)$ and $y \in \gamma^s(x)$ we have
$$\left|\log \frac{\mathrm{Jac}\tilde{f}(\tilde{f}^n(x))}{\mathrm{Jac}\tilde{f}(\tilde{f}^n(y))}\right| \le cb^n;$$

(b) For any $i_0, \ldots, i_n \in \mathbb{N}$, $\tilde{f}^k(x), \tilde{f}^k(y) \in \Lambda_{i_k}^s$ for $0 \le k \le n$ and $y \in \gamma^u(x)$ we have
$$\left|\log \frac{\mathrm{Jac}\tilde{f}(\tilde{f}^{n-k}(x))}{\mathrm{Jac}\tilde{f}(\tilde{f}^{n-k}(y))}\right| \le cb^k.$$

(Y5) For every $\gamma^u \in \Gamma^u$ one has
$$\mu_{\gamma^u}(\gamma^u \cap \Lambda) > 0, \quad \mu_{\gamma^u}\left(\overline{(\Lambda \setminus \cup\Lambda_i^s) \cap \gamma^u}\right) = 0,$$
where μ_{γ^u} is the leaf volume on γ^u.

(Y6) There exists $\gamma^u \in \Gamma^u$ such that
$$\sum_{i=1}^{\infty} \tau_i \mu_{\gamma^u}(\Lambda_i^s \cap \gamma^u) < \infty.$$

It is shown in [21] that the Katok map is a Young diffeomorphism. We will briefly outline the argument.

5.2. **A tower representation of the automorphism A.** Consider a finite Markov partition $\tilde{\mathcal{P}}$ for the automorphism A. Recall that by definition of Markov partitions, $\tilde{P} = \overline{\text{int}\tilde{P}}$ for any $\tilde{P} \in \tilde{\mathcal{P}}$. Let $\tilde{P} \in \tilde{\mathcal{P}}$ be a partition element which does not intersect any of the disks $D^i_{r_0}$, $i = 1, 2, 3, 4$. Given $\delta > 0$, we can always choose the Markov partition $\tilde{\mathcal{P}}$ in such a way that diam $(\tilde{P}) < \delta$. For a point $x \in \tilde{P}$ denote by $\tilde{\gamma}^s(x)$ (respectively, $\tilde{\gamma}^u(x)$) the connected component of the intersection of $\tilde{P}$ with the stable (respectively, unstable) leaf of x, which contains x. We say that $\tilde{\gamma}^s(x)$ and $\tilde{\gamma}^u(x)$ are *full length* stable and unstable curves through x.

Given $x \in \tilde{P}$, let $\tilde{\tau}(x)$ be the first return time of x to $\text{int}\tilde{P}$. For all x with $\tilde{\tau}(x) < \infty$ denote by

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

where $\tilde{U}^u(x) \subseteq \tilde{\gamma}^u(x)$ is an interval containing x and open in the induced topology of $\tilde{\gamma}^u(x)$, and $\tilde{A}^u(x) \subset \tilde{U}^u(x)$ is the set of points which either lie on the boundary of the Markov partition or never return to the set $\tilde{P}$. Note that $\tilde{A}^u(x)$ has zero one-dimensional Lebesgue measure in $\tilde{\gamma}^u(x)$. One can choose $\tilde{U}^u(x)$ such that

(1) for any $y \in \tilde{\Lambda}^s(x)$ we have $\tilde{\tau}(y) = \tilde{\tau}(x)$;
(2) for any $y \in \tilde{P}$ such that $\tilde{\tau}(y) = \tilde{\tau}(x)$ we have $y \in \tilde{\Lambda}^s(x)$.

Moreover, the image under $A^{\tilde{\tau}(x)}$ of $\tilde{\Lambda}^s(x)$ is a u-subset containing $A^{\tilde{\tau}(x)}(x)$. It is easy to see that for any $x, y \in \tilde{P}$ with finite first return time the sets $\tilde{\Lambda}^s(x)$ and $\tilde{\Lambda}^s(y)$ are either coincide or disjoint. Thus we have a countable collection of disjoint sets $\tilde{\Lambda}^s_i$ and numbers $\tilde{\tau}_i$ which give a representation of the automorphism A as a Young diffeomorphism for which the set

$$\tilde{\Lambda} = \bigcup_{i \geq 1} \tilde{\Lambda}^s_i$$

is the base of the tower, the sets $\tilde{\Lambda}^s_i$ are the s-sets and the numbers $\tilde{\tau}_i$ are the inducing times, see [21] for details.

5.3. **A tower representation for the slow down map $f_{\mathbb{T}^2}$.** Applying the conjugacy map H, one obtains the element $P = H(\tilde{P})$ of the Markov partition $\mathcal{P} = H(\tilde{\mathcal{P}})$. Since the map H is continuous, given ε, there is $\delta > 0$ such that diam $(P) < \varepsilon$ for any $P \in \mathcal{P}$ provided diam $(\tilde{P}) < \delta$. Further we obtain the set $\Lambda = H(\tilde{\Lambda})$, which has direct product structure given by the full length stable $\gamma^s(x) = H(\tilde{\gamma}^s(x))$ and unstable $\gamma^u(x) = H(\tilde{\gamma}^u(x))$ curves. We thus obtain a representation of the slow down map as a Young diffeomorphism for which

(1) $\Lambda_i^s = H(\tilde{\Lambda}_i^s)$ are s-sets;
(2) $\Lambda_i^u = H(\tilde{\Lambda}_i^u) = f_{\mathbb{T}^2}^{\tau_i}(\Lambda_i^s)$ are u-sets;
(3) $\tau_i = \tilde{\tau}_i$ – the inducing times – are the first return time of points in Λ_i^s to Λ;
(4) $\tilde{f}(x) = f_{\mathbb{T}^2}^{\tau(x)}(x)$ is the induced map.

Note that for all x with $\tau(x) < \infty$

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

where $U^u(x) = H(\tilde{U}^u(x)) \subseteq \gamma^u(x)$ is an interval containing x and open in the induced topology of $\gamma^u(x)$, and $A^u(x) = H(\tilde{A}^u(x)) \subset U^u(x)$ is the set of points which either lie on the boundary of the Markov partition or never return to the set P. Note that $A^u(x)$ has zero one-dimensional Lebesgue measure in $\gamma^u(x)$.

In what follows we will always assume that a Markov partition and the slow down domain are chosen such that the following statement holds.

Proposition 5.1. *Given $Q > 0$, one can choose a Markov partition $\mathcal{P}$ and the number r_0 in the construction of the map $f_{\mathbb{T}^2}$ such that*

(1) *there is a partition element P for which $f_{\mathbb{T}^2}^j(x) \notin D_{r_0}$ for any $0 \le j \le Q$ and for any point x for which either $x \in \Lambda$ or $x \notin f_{\mathbb{T}^2}(D_{r_0})$ while $f_{\mathbb{T}^2}^{-1}(x) \in D_{r_0}^i$ for some $i = 1, 2, 3, 4$.*
(2) *if P_i is the element of the Markov partition containing x_i, $i = 1, 2, 3, 4$, then $x_i \in D_{r_0}^i \subset Int\ P_i$.*

To prove this proposition observe that it holds for the automorphism A and hence, it remains to apply the conjugacy homeomorphism H.

Proposition 5.2 ([21], Proposition 6.2)**.** *There exists $Q > 0$ such that the collection of s-subsets Λ_i^s satisfies Conditions (Y1)-(Y6).*

5.4. **Lifting the slow down map to the tower.** We define *Young tower* with the base Λ by setting

$$\hat{Y} = \{(x, k) \in \Lambda \times \mathbb{N} : 0 \le k < \tau(x)\}.$$

and the *tower map* $\hat{f}_{\mathbb{T}^2} : \hat{Y} \to \hat{Y}$ by $\hat{f}_{\mathbb{T}^2}(x, k) = (x, k+1)$ if $k < \tau(x) - 1$ and $\hat{f}_{\mathbb{T}^2}(x, k) = (\tilde{f}(x), 0)$ if $k = \tau(x) - 1$ where $\tilde{f} : \Lambda \to \Lambda$ is the induced map. The map $\hat{f}_{\mathbb{T}^2}$ is the lift of the slow down map to the tower and it preserves the *lift measure* $\hat{m} = m \times \text{counting}/(\int_\Lambda \tau)$. We have that

$\hat{f}_{\mathbb{T}^2}$ is a measurable bijection from $\Lambda_i^s \times \{k-1\}$ to $\Lambda_i^s \times \{k\}$ for all $1 \le k < \tau_i - 1$ and from $\Lambda_i^s \times \{\tau_i - 1\}$ to $\Lambda \times \{0\}$.

6. Proof of Theorem 3.1: Technical Lemmas

We establish here several technical results on the solutions of the nonlinear systems of differential equations (3). Throughout this section we fix a number $0 < \alpha < 1$ and $i \in \{1,2,3,4\}$ and for simplicity we drop the index i in the notation of the disk D_r^i. We also set $f := f_{\mathbb{T}^2}$.

Lemma 6.1 ([21], Lemma 5.1)**.** *For $s = (s_1, s_2) \in D_{\frac{r_0}{2}}$ let*

$$d_{i,j} = d_{i,j}(s_1, s_2) := \frac{\partial^2}{\partial s_i \partial s_j} s_2 \psi(s_1^2 + s_2^2).$$

Then

$$\max_{i,j=1,2} |d_{i,j}| \le \frac{6\alpha}{r_0^\alpha} (s_1^2 + s_2^2)^{\alpha - \frac{1}{2}}.$$

Consider a solution $s(t) = (s_1(t), s_2(t))$ of Equation (3) with an initial condition $s(0) = (s_1(0), s_2(0))$. Assume it is defined on the *maximal* time interval $[0, T]$ for which $f^{-1}(s(0)) \notin D_{\frac{r_0}{2}}$ and $f(s(T)) \notin D_{\frac{r_0}{2}}$ but $s(t) \in D_{\frac{r_0}{2}}$ for all $0 \le t \le T$. In particular, $s_1(t) \ne 0$ and $s_2(t) \ne 0$. Setting $T_1 = \frac{T}{2}$ we have that $s_1(t) \le s_2(t)$ for all $0 \le t \le T_1$ and $s_1(t) \ge s_2(t)$ for all $T_1 \le t \le T$. The following statement provides effective lower and upper bounds on the functions $s_1(t)$ and $s_2(t)$. For the proof see Lemma 5.2 in the erratum to the paper [21].

Lemma 6.2. *The following statements hold:*

$$|s_2(t)| \ge |s_2(a)| \left(1 + 2^\alpha C_1 s_2^{2\alpha}(a)(t-a)\right)^{-\frac{1}{2\alpha}}, \quad 0 \le a \le t \le T_1;$$
$$|s_2(t)| \le |s_2(a)| \left(1 + C_1 s_2^{2\alpha}(a)(t-a)\right)^{-\frac{1}{2\alpha}}, \quad 0 \le a \le t \le T;$$
$$|s_1(t)| \ge |s_1(b)| \left(1 + 2^\alpha C_1 s_1^{2\alpha}(b)(b-t)\right)^{-\frac{1}{2\alpha}}, \quad T_1 \le t \le b \le T;$$
$$|s_1(t)| \le |s_1(b)| \left(1 + C_1 s_1^{2\alpha}(b)(b-t)\right)^{-\frac{1}{2\alpha}}, \quad 0 \le t \le b \le T.$$

where $C_1 = \frac{2\alpha \log \lambda}{r_0^\alpha}$ is a constant.

Consider another solution $\tilde{s}(t) = (\tilde{s}_1(t), \tilde{s}_2(t))$ of Equation (3) satisfying an initial condition $\tilde{s}(0) = (\tilde{s}_1(0), \tilde{s}_2(0))$. For $i = 1, 2$, we set

$$\Delta s_i(t) = \tilde{s}_i(t) - s_i(t).$$

For the proof of the next result see Lemma 5.3 in the erratum to the paper [21].

Lemma 6.3. *Fix $0 < \mu < 1$ and assume that $s_1(t) \neq 0$, $s_2(t) \neq 0$ for all $0 \le t \le T$, and that*

(1) *$\Delta s_2(t) > 0$ and $|\Delta s_1(t)| \le \mu \Delta s_2(t)$ for $t \in [0, T]$;*
(2) $\left| \frac{\Delta s_2}{s_2}(0) \right| < \frac{1-\mu}{72}$.

Then

$$\Delta s_2(t) \le \frac{\Delta s_2(0)}{s_2(0)} s_2(t) \left(1 + 2^\alpha C_1 s_2^{2\alpha}(0)\, t\right)^{-\beta'}, \qquad 0 \le t \le T_1;$$

$$\Delta s_2(t) \le \frac{\Delta s_2(T_1)}{s_1(T_1)} s_1(t) \left(\frac{1 + 2^\alpha C_1 s_1^{2\alpha}(b)(b-t)}{1 + 2^\alpha C_1 s_1^{2\alpha}(b)(b-T_1)} \right)^{\beta'}, \quad T_1 \le t \le b \le T,$$

where $\beta' = \frac{1-\mu}{2^{\alpha+2}}$ and C_1 is the constant in Lemma 6.2. In addition,

$$\|\Delta s(T)\| \le \sqrt{1+\mu^2}\, \frac{s_1(T)}{s_2(0)}\, \|\Delta s(0)\|.$$

Given $0 < \alpha < 1$ and $0 < \mu < 1$, denote by

$$\gamma = \frac{1}{2\alpha} + 2^{\alpha-1}(1+\mu) + \frac{1-\mu}{6}, \quad \gamma' = \frac{1}{2\alpha} + \frac{1-\mu}{2^{\alpha+2}}. \tag{8}$$

It is easy to see that $\gamma > \gamma' > 2$ for all $0 < \alpha < \frac{1}{4}$ and $0 < \mu < \frac{1}{2}$.

In what follows till the end of the next section we denote by C_2 through C_{10} positive constants that are independent of the time t and the choice of the solution $(s_1(t), s_2(t))$.

Lemma 6.4. *Under the assumptions of Lemma 6.3 for any $0 \le t \le T_1$ we have*

$$\Delta s_2(t) \le C_2 \Delta s_2(0) t^{-\gamma'}.$$

Proof. By Lemma 6.2 (the second estimate), one has

$$s_2(t) \le s_2(0)(1 + C_1 s_2^{2\alpha}(0) t)^{-\frac{1}{2\alpha}}.$$

Therefore, Lemma 6.3 implies that

$$\begin{aligned}\Delta s_2(t) &\le \frac{\Delta s_2(0)}{s_2(0)} s_2(t) \left(1 + 2^\alpha C_1 s_2^{2\alpha}(0)\, t\right)^{-\beta'} \\ &\le \Delta s_2(0) \left(1 + C_1 s_2^{2\alpha}(0)\, t\right)^{-\gamma'}.\end{aligned}$$

Since $(1 + C_1 s_2^{2\alpha}(0))t > C_1 s_2^{2\alpha}(0) t$, we have

$$\Delta s_2(t) \le \Delta s_2(0) (C_1 s_2^{2\alpha}(0))^{-\gamma'} t^{-\gamma'}$$

and the desired estimate follows, since $s_2(0)$ is of order r_0. □

Lemma 6.5. *Under the assumptions of Lemma 6.3 we have*

$$\Delta s_2(t) \geq \frac{\Delta s_2(0)}{s_2(0)} s_2(t) \left(1 + C_1\, s_2^{2\alpha}(0)\, t\,\right)^{-\beta}, \qquad 0 \leq t \leq T_1;$$

$$\Delta s_2(t) \geq \frac{\Delta s_2(T_1)}{s_1(T_1)} s_1(t) \left(1 + C_1\, s_1^{2\alpha}(T_1)\,(t - T_1)\right)^{-\beta_1}, \quad T_1 \leq t \leq T,$$

where $\beta = (1+\mu)2^{\alpha-1} + \frac{1-\mu}{6}$ *and* $\beta_1 = \beta + \frac{2\alpha}{\alpha}$.

Proof. Let $s_1 = s_1(t)$, $s_2 = s_2(t)$, $u := s_1^2 + s_2^2$, and $\tilde{u} = \tilde{s}_1^2 + \tilde{s}_2^2$. Assume $s_1(t)$ and $s_2(t)$ are strictly positive (the proof in the case when $s_1(t)$ and $s_2(t)$ are strictly negative follows by symmetry). By Equation (3), we have

$$\begin{aligned} \frac{d}{dt}\Delta s_2(t) &= \frac{d}{dt}\tilde{s}_2(t) - \frac{d}{dt}s_2(t) = -(\log\lambda)\left(\tilde{s}_2\psi(\tilde{u}) - s_2\psi(u)\right) \\ &= -\log\lambda \left(\frac{\partial}{\partial s_1}\Big(s_2\psi(u)\Big)\Delta s_1 + \frac{\partial}{\partial s_2}\Big(s_2\psi(u)\Big)\Delta s_2\right) \\ &\quad - \frac{\log\lambda}{2}\sum_{i,j=1,2} d_{i,j}(\xi_1,\xi_2)(\Delta s_i)(\Delta s_j) \end{aligned} \tag{9}$$

for some $\xi = (\xi_1, \xi_2)$ for which ξ_i lies between $s_i(t)$ and $\tilde{s}_i(t)$ for $i = 1, 2$ (see Lemma 6.1 for the definition of $d_{i,j}(\xi_1, \xi_2)$). Note that

$$\frac{\partial}{\partial s_1}\Big(s_2\psi(u)\Big) = 2s_1 s_2\psi', \quad \frac{\partial}{\partial s_2}\Big(s_2\psi(u)\Big) = 2s_2^2\psi' + \psi$$

and hence,

$$\begin{aligned} \frac{d}{dt}\left(\frac{\Delta s_2}{s_2}\right) &= \frac{1}{s_2}\left(\frac{d}{dt}\Delta s_2\right) - \frac{\Delta s_2}{s_2^2}\left(\frac{ds_2}{dt}\right) \\ &= -\log\lambda\left(2\psi'(s_1\Delta s_1 + s_2\Delta s_2) + \frac{\Delta s_2}{s_2}\psi\right) \\ &\quad + (\log\lambda)\frac{\Delta s_2}{s_2}\psi - \frac{\log\lambda}{2}\sum_{i,j=1,2} d_{i,j}(\xi_1,\xi_2)\frac{\Delta s_i \Delta s_j}{s_2} \\ &= -\frac{2\alpha\log\lambda}{r_0^\alpha}\,(s_1^2 + s_2^2)^{\alpha-1}(s_1\Delta s_1 + s_2\Delta s_2) \\ &\quad - \frac{\log\lambda}{2}\sum_{i,j=1,2} d_{i,j}(\xi_1,\xi_2)\frac{\Delta s_i\Delta s_j}{s_2}. \end{aligned}$$

Note that for $0 \leq t \leq T_1$ we have $0 < s_1(t) \leq s_2(t)$. Since $|\Delta s_1| \leq \mu\Delta s_2 < \Delta s_2$, we have

$$s_1\Delta s_1 + s_2\Delta s_2 \leq (s_1\mu + s_2)\Delta s_2 \leq (1+\mu)s_2\Delta s_2$$

and Lemma 6.1 yields

$$\sum_{i,j=1,2} d_{i,j}(\xi_1,\xi_2)\Delta s_i \Delta s_j \le \frac{24\,\alpha}{r_0^\alpha}\ (\xi_1^2+\xi_2^2)^{\alpha-\frac{1}{2}}(\Delta s_2)^2. \tag{10}$$

Therefore, we have

$$\begin{aligned}\frac{d}{dt}(\frac{\Delta s_2}{s_2}) &\ge -(1+\mu)\frac{2\alpha\log\lambda}{r_0^\alpha}(s_1^2+s_2^2)^{\alpha-1}s_2^2\frac{\Delta s_2}{s_2}\\ &- \frac{12\alpha\log\lambda}{r_0^\alpha}s_2^{2\alpha}(\frac{\xi_1^2+\xi_2^2}{s_2^2})^{\alpha-\frac{1}{2}}\Big(\frac{\Delta s_2}{s_2}\Big)^2.\end{aligned}$$

Using again the fact that $0 < s_1(t) \le s_2(t),\ 0 < t < T_1$ we obtain

$$\begin{aligned}\frac{d}{dt}\Big(\frac{\Delta s_2}{s_2}\Big) &\ge -(1+\mu)\frac{\alpha\log\lambda}{r_0^\alpha}2^\alpha(s_2)^{2\alpha}\frac{\Delta s_2}{s_2}\\ &- \frac{12\alpha\log\lambda}{r_0^\alpha}s_2^{2\alpha}(\frac{\xi_1^2+\xi_2^2}{s_2^2})^{\alpha-\frac{1}{2}}\Big(\frac{\Delta s_2}{s_2}\Big)^2.\end{aligned}$$

Let $\chi = \chi(t) = \frac{\Delta s_2}{s_2}(t)$. Then the above inequality can be written as

$$\frac{d\chi}{dt} \ge -\frac{\alpha\log\lambda}{r_0^\alpha}s_2^{2\alpha}\chi\Big((1+\mu)2^\alpha + 12\left(\frac{\xi_1^2+\xi_2^2}{s_2^2}\right)^{\alpha-\frac{1}{2}}\chi\Big).$$

Following arguments in [21] (see page 17) one can derive from here that $\left(\frac{\xi_1^2+\xi_2^2}{s_2^2}\right)^{\alpha-\frac{1}{2}} \le 2$ and that

$$\frac{d\chi}{dt} \ge -A\frac{\alpha\log\lambda}{r_0^\alpha}s_2^{2\alpha}(t)\chi(t),$$

where $A = (1+\mu)2^\alpha + \frac{1-\mu}{3}$. By Gronwall's inequality (applied to $-\chi(t)$) and the second inequality in Lemma 6.2, we obtain that

$$\begin{aligned}\chi(t) &\ge \chi(0)\exp\left(-A\frac{\alpha\log\lambda}{r_0^\alpha}\int_0^t s_2^{2\alpha}(\tau)\,d\tau\right)\\ &\ge \chi(0)\exp\left(-A\frac{\alpha\log\lambda}{r_0^\alpha}\int_0^t s_2^{2\alpha}(0)(1+C_1s_2^{2\alpha}(0)\,\tau)^{-1}\,d\tau\right)\\ &= \chi(0)\exp\left(-A\frac{\alpha\log\lambda}{r_0^\alpha}\frac{1}{C_1}\log(1+C_1s_2^{2\alpha}(0)\,t\ )\right)\\ &= \chi(0)\exp\left(-\frac{A}{2}\log(1+C_1s_2^{2\alpha}(0)\,t\ )\right)\\ &= \chi(0)\left(1+C_1s_2^{2\alpha}(0)\,t\ \right)^{-\beta},\end{aligned}$$

where $\beta = \frac{A}{2} = (1+\mu)2^{\alpha-1} + \frac{1-\mu}{6}$.

In the case $s_2 < 0$ one can show using argument similar to the above that the same estimate for $\chi(t)$ holds but with exponent $(1+\mu)2^{\alpha-1} - \frac{1-\mu}{6} < \beta$. This completes the proof of the first estimate.

To prove the second estimate, using (9), we obtain that

$$\frac{d}{dt}\left(\frac{\Delta s_2}{s_1}\right) = -\log\lambda\left(2s_2\psi'\Delta s_1 + (2s_2^2\psi' + \psi)\frac{\Delta s_2}{s_1}\right) - \log\lambda\psi\frac{\Delta s_2}{s_1}$$
$$-\frac{\log\lambda}{2}\sum_{i,j=1,2} d_{i,j}(\xi_1,\xi_2)\frac{\Delta s_i\Delta s_j}{s_1}.$$

By the assumption $|\Delta s_1| \le \mu\Delta s_2$ and positivity of s_1, s_2, ψ', and Δs_2, we obtain

$$\frac{d}{dt}\left(\frac{\Delta s_2}{s_1}\right) \ge -\log\lambda(2\mu s_1 s_2\psi' + 2s_2^2\psi' + 2\psi)\frac{\Delta s_2}{s_1}$$
$$-\frac{\log\lambda}{2}\sum_{i,j=1,2} d_{i,j}(\xi_1,\xi_2)\frac{\Delta s_i\Delta s_j}{s_1}.$$

Since $s_2(t) \le s_1(t)$ on $[T_1, T]$, we have

$$2\mu s_1 s_2\psi' + 2s_2^2\psi' + 2\psi = 2\psi'(\mu s_1^2 + s_1^2 + \frac{1}{\alpha}2s_1^2) \le 2^\alpha\frac{\alpha\mu+\alpha+2}{r_0^\alpha}s_1^{2\alpha}.$$

Using this fact along with the estimate (10), we find that

$$\frac{d}{dt}\left(\frac{\Delta s_2}{s_1}\right) \ge -\log\lambda 2^\alpha\frac{\alpha\mu+\alpha+2}{r_0^\alpha}s_1^{2\alpha}\frac{\Delta s_2}{s_1}$$
$$-\frac{12\alpha\log\lambda}{r_0^\alpha}s_1^{2\alpha}(\frac{\xi_1^2+\xi_2^2}{s_1^2})^{\alpha-\frac{1}{2}}\left(\frac{\Delta s_2}{s_1}\right)^2.$$

Let $\tilde\chi = \tilde\chi(t) = \frac{\Delta s_2}{s_1}(t)$. Then the above inequality can be written as

$$\frac{d\tilde\chi}{dt} \ge -\frac{\log\lambda}{r_0^\alpha}s_1^{2\alpha}\tilde\chi\left(2^\alpha(\alpha\mu+\alpha+2) + 12\alpha\left(\frac{\xi_1^2+\xi_2^2}{s_1^2}\right)^{\alpha-\frac{1}{2}}\tilde\chi\right). \tag{11}$$

It is shown in [21] (see page 19) that

$$\left(\frac{\xi_1^2+\xi_2^2}{s_1^2}\right)^{\alpha-\frac{1}{2}} \le \begin{cases}(1-\tilde\chi)^{2\alpha-1} & 0<\alpha\le\frac{1}{2},\\ 2^{\alpha-\frac{1}{2}}(1+\tilde\chi)^{2\alpha-1} & \frac{1}{2}\le\alpha<1\end{cases} \tag{12}$$

and that $\tilde\chi = \frac{\Delta s_2}{s_2}$ is positive and decreasing (or negative and increasing). Observing that $s_1(T_1) = s_2(T_1)$ and using Assumption (2) of the lemma, we obtain

$$0 \le \tilde\chi(T_1) = \frac{\Delta s_2(T_1)}{s_1(T_1)} = \frac{\Delta s_2(T_1)}{s_2(T_1)} \le \frac{\Delta s_2(0)}{s_2(0)} < \frac{1-\mu}{72}.$$

$$(13)\quad \tilde\chi(t)=\frac{\Delta s_2(t)}{s_1(t)}\le\frac{\Delta s_2(T_1)}{s_1(T_1)}(1+2^\alpha C_1 s_1^{2\alpha}(t)(t-T_1))^{-\beta}\le\frac{\Delta s_2(T_1)}{s_1(T_1)}.$$

Setting $B=2^\alpha(\alpha\mu+\alpha+2)+\frac{1-\mu}{3}\alpha$ and combining (11), (12), and (13), we obtain that

$$(14)\qquad \frac{d\tilde\chi}{dt}\Big|_{t=T_1}\ge-\frac{B\log\lambda}{r_0^\alpha}s_1^{2\alpha}(T_1)\tilde\chi(T_1).$$

Therefore, Gronwall's inequality and the fourth inequality in Lemma 6.2 now yield

$$\begin{aligned}\tilde\chi(t)&\ge\tilde\chi(T_1)\exp\left(-B\frac{\log\lambda}{r_0^\alpha}\int_{T_1}^t s_1^{2\alpha}(\tau)\,d\tau\right)\\&\ge\tilde\chi(T_1)\exp\left(-B\frac{\log\lambda}{r_0^\alpha}\int_{T_1}^t s_1^{2\alpha}(t)(1+C_1s_1^{2\alpha}(0)\,(t-\tau))^{-1}\,d\tau\right)\\&=\tilde\chi(T_1)\exp\left(-B\frac{\log\lambda}{r_0^\alpha}\frac{1}{C_1}\log(1+C_1s_1^{2\alpha}(t)\,(t-T_1)\;)\right)\\&=\tilde\chi(T_1)\exp\left(-B\log(1+C_1s_1^{2\alpha}(t)\,(t-T_1)\;)\right)\\&=\tilde\chi(T_1)\left(1+C_1s_1^{2\alpha}(t)\,(t-T_1)\;\right)^{-\beta_1},\end{aligned}$$

where $\beta_1=\frac{B}{2\alpha}$. It follows that

$$\Delta s_2(t)\ge\frac{\Delta s_2(T_1)}{s_1(T_1)}s_1(t)\left(1+C_1\,s_1^{2\alpha}(T_1)\,(t-T_1)\right)^{-\beta_1},\ T_1\le t\le T.$$

In the case $s_1<0$ one can show using argument similar to the above that the same estimate for $\tilde\chi(t)$ holds but with exponent $2^{\alpha-1}(1+\mu)+\frac{2^\alpha}{\alpha}-\frac{1-\mu}{6}<\beta_1$. This completes the proof of the second estimate. □

Lemma 6.6. *Under the assumptions of Lemma 6.3 for all* $0\le t\le T_1$ *we have*

$$\Delta s_2(t)\ge C_3\Delta s_2(0)t^{-\gamma}.$$

Proof. By assumption, Lemma 6.5 holds and the first estimate together with the first inequality in Lemma 6.2 yield

$$\begin{aligned}\Delta s_2(t)&\ge\frac{\Delta s_2(0)}{s_2(0)}s_2(t)(1+C_1s_2^{2\alpha}(0)t)^{-\beta}\\&\ge\Delta s_2(0)(1+C_1s_2^{2\alpha}(0)t)^{-(\frac{1}{2\alpha}+\beta)}.\end{aligned}$$

Since $1+C_1s_2^{2\alpha}(0)t\le(1+C_1s_2^{2\alpha}(0))t$ for $t\ge1$, the above implies

$$\Delta s_2(t)>\Delta s_2(0)(1+C_1s_2^{2\alpha}(0))^{-(\frac{1}{2\alpha}+\beta)}t^{-(\frac{1}{2\alpha}+\beta)}.$$

It remains to observe that $s_2(0)$ is of order r_0 and so $(1+C_1s_2^{2\alpha}(0))^{-(\frac{1}{2\alpha}+\beta)}$ is a constant and $\gamma=\frac{1}{2\alpha}+\beta$.

For $0 < t \leq 1$ the orbit stays bounded away from the region of perturbation and so the inequality holds for some constant. □

Lemma 6.7. *Under the assumptions of Lemma 6.3*

$$C_4 \Delta s_2(T_1) \geq \Delta s_2(T) \geq C_5 \Delta s_2(T_1).$$

Proof. First, we prove the lower bound. The forth inequality in Lemma 6.2 with $t = T_1$ and $b = T$ yields

$$s_1(T) \geq s_1(T_1)(1 + C_1 s_1^{2\alpha}(T)(T - T_1))^{\frac{1}{2\alpha}}.$$

This and the second estimate in Lemma 6.5 implies that

$$\begin{aligned}\Delta s_2(T) &\geq \frac{\Delta s_2(T_1)}{s_1(T_1)} s_1(T_1) \left(1 + C_1\, s_1^{2\alpha}(T)(T - T_1)\right)^{\frac{1}{2\alpha}} \\ &\times \left(1 + C_1\, s_1^{2\alpha}(T_1)(T - T_1)\right)^{-\beta_1} \\ &\geq \Delta s_2(T_1) \left(1 + C_1\, s_1^{2\alpha}(T_1)(T - T_1)\right)^{\frac{1}{2\alpha} - \beta_1},\end{aligned}$$

where we recall that $s_1(T) \geq s_1(T_1)$ and $\beta_1 = 2^{\alpha-1}(1 + \mu) + \frac{1-\mu}{6} + \frac{2^\alpha}{\alpha}$, $0 < \alpha < \frac{1}{4}$, and $0 < \mu < \frac{1}{2}$. It is easy to see that $\beta_1 > \frac{1}{2\alpha}$ and hence, to complete the proof of the lower bound it suffices to show that $s_1^{2\alpha}(T_1)(T - T_1)$ is bounded above by a constant that is independent of T_1, T, and the choice of the solution $(s_1(t), s_2(t))$.

Since $s_1(T_1) = s_2(T_1)$, applying the second inequality in Lemma 6.2 with $a = 0$ and $t = T_1$, we obtain that

$$s_1^{2\alpha}(T_1) \leq s_2^{2\alpha}(0) \left(1 + C_1 s_2^{2\alpha}(0) T_1\right)^{-1}.$$

This implies that

$$\begin{aligned}s_1^{2\alpha}(T_1)(T - T_1) &\leq \frac{s_2^{2\alpha}(0)(T - T_1)}{1 + C_1 s_2^{2\alpha}(0) T_1} \\ &\leq \frac{s_2^{2\alpha}(0)(T - T_1)}{C_1 s_2^{2\alpha}(0) T_1} = \frac{T - T_1}{C_1 T_1}.\end{aligned}$$

Since $\frac{T-T_1}{T_1}$ is of order 1 regardless of the choice of the solution $(s_1(t), s_2(t))$, this completes the proof of the lower bound.

To prove the upper bound we use the existence of invariant stable and unstable cones at every point x in the disk D_r^i, $i = 1, 2, 3, 4$. More precisely, let $K^-(x)$ (respectively, $K^+(x)$) be the cone at x around vertical (respectively, horizontal) line of angle $\frac{\pi}{4}$. One can show (see [12], Proposition 4.1) that these families of cones are invariant under the flow g_t^i (given by Equation (3)) that is

$$dg_\tau^i(K^+(x)) \subset K^+(g_\tau^i(x)), \quad d(g_\tau^i)^{-1}(K^-(x)) \subset K^-((g_\tau^i)^{-1}(x)).$$

Moreover, let us fix positive numbers a and b and consider the region

$$\{(s_1, s_2) \colon |s_1 s_2| \le a,\ |s_1| \le b,\ |s_2| \le b\} \subset D_r^i$$

and a segment of hyperbola

$$\{s_1 s_2 = \varepsilon,\ 0 \le s_1 \le b,\ 0 \le s_2 \le b\} \text{ for } 0 < \varepsilon \le a.$$

One can show (see [12], Proposition 4.1) that for every (s_1, s_2) on this hyperbola every angle inside the cone $K^-(x)$ is contracted under $(dg_t^i)^{-1}$ by the factor $d = C\frac{a^2}{b^4}$, where $C > 0$ is a constant independent of x, a, and b (note that given $b > 0$, one can choose a so small to ensure that $d < 1$).

Let $s(t) = (s_1(t), s_2(t))$ and $\tilde{s}(t) = (\tilde{s}_1(t), \tilde{s}_2(t))$ be two solutions of Equation (3) satisfying the assumptions of Lemma 6.3. It follows from what was said above that for any $T_1 < t < T$ and any $\tau \ge 0$ such that $T_1 \le t - \tau$ we have that $\Delta s(t) \subset K^-(s(t))$ and $\|\Delta s(t-\tau)\| \le \|\Delta s(t)\|$. Since $\Delta s(t) = \Delta s_1(t) + \Delta s_2(t)$ and, by assumption, $|\Delta s_1(t)| \le \mu \Delta s_2(t)$, choosing $t = T$ and $\tau = T - T_1$, we obtain the desired upper bound. This completes the proof of the lemma. □

Consider a set Λ_i^s and note that it consists of full length s-curves. Let us fix one of these curves, say σ.

Lemma 6.8. *Assume that σ enters the slow down disk $D_{\frac{r_0}{2}}$ at time n, so that the intersection $f^n(\sigma) \cap D_{\frac{r_0}{2}}$ is not empty. Assume that σ then exits $D_{\frac{r_0}{2}}$ at time m, $m > n > 1$. Then*

$$C_6(m-n)^{-\gamma} \le \frac{L(f^m(\sigma))}{L(f^n(\sigma))} \le C_7(m-n)^{-\gamma'},$$

where γ, γ' are as in (8), and L denotes the length of the curve.

Proof. Let x and y be the endpoints of the curve σ. For $k \ge 0$ set $x_k = f^k(x)$ and $y_k = f^k(y)$. It is easy to see that there is $K_0 > 0$ such that for all $k \ge 1$,

$$K_0^{-1} d(x_k, y_k) \le L(f^k(\sigma)) \le K_0 d(x_k, y_k), \tag{15}$$

where d denotes the usual distance.

Let $s, \tilde{s} : [0, N] \to R^2$ be the solutions of Equation (2) with initial conditions $s(0) = x_n$ and $\tilde{s}(0) = y_n$ respectively. Also, define $\Delta s_i(t) = \tilde{s}_i(t) - s_i(t)$, $i = 1, 2$ and $\Delta s = (\Delta s_1, \Delta s_2)$. Note that there is $K_1 > 0$ such that for all $n, m > 0$ and $n \le j \le m$

$$K_1^{-1} ||\Delta s(j)|| \le d(x_j, y_j) \le K_1 ||\Delta s(j)||. \tag{16}$$

In what follows we will use Lemma 6.3 and we need to check the assumptions of this Lemma. Assumption (1) is satisfied, since y is contained in the stable cone at x. Assumption (2) requires $d(x_i, y_i), i = n_0, m_0$ to be sufficiently small. In view of Proposition 6.1 and (16), this can be ensured, if we choose the number r_0 in the construction of the slow down map sufficiently small to guarantee that Q is sufficiently large.

We have that $f^j(\sigma) \subset D_{\frac{r_0}{2}} \cap \{(s_1, s_2) : s_2 \geq s_1\}$ for $n \leq j \leq \frac{n+m}{2}$ and $f^j(\sigma) \subset D_{\frac{r_0}{2}} \cap \{(s_1, s_2) : s_2 < s_1\}$ for $\frac{n+m}{2} < j < m$. Applying Lemmas 6.3, 6.6 with $0 < t \leq \frac{n+m}{2}$ and $\frac{n+m}{2} < t \leq m$ and using lower bound in Lemma 6.7 as well as (15) and (16), we obtain that

$$\begin{aligned}
L(f^m(\sigma)) &\geq K_0^{-1} d(x_m, y_m) \geq K_0^{-1} K_1^{-1} ||\Delta s(m-n)|| \\
&\geq K_0^{-1} K_1^{-1} \Delta s_2(m-n) > K_0^{-1} K_1^{-1} C_5 \Delta s_2 \big(\frac{m-n}{2}\big) \\
&\geq K_0^{-1} K_1^{-1} C_5 C_3 \Delta s_2(0) \big(\frac{m-n}{2}\big)^{-\gamma} \\
&\geq K_0^{-1} K_1^{-1} C_5 C_3 2^{\gamma} (m-n)^{-\gamma} \frac{1}{\sqrt{1+\mu^2}} ||\Delta s(0)|| \\
&\geq K_0^{-2} K_1^{-2} C_5 C_3 2^{\gamma} (m-n)^{-\gamma} \frac{1}{\sqrt{1+\mu^2}} L(f^n(\sigma)).
\end{aligned}$$

Therefore, for some $C_6 > 0$,

$$\frac{L(f^m(\sigma))}{L(f^n(\sigma))} \geq C_6 (m-n)^{-\gamma}.$$

To prove the upper bound we use Lemmas 6.3, 6.4 with $0 < t \leq \frac{n+m}{2}$ and $\frac{n+m}{2} < t \leq m$, the upper bound in Lemma 6.7 as well as inequalities (15) and (16). Then the arguments similar to the above yield that for some $C_7 > 0$,

$$\frac{L(f^m(\sigma))}{L(f^n(\sigma))} \leq C_7 (m-n)^{-\gamma'}.$$

This completes the proof of the lemma. □

7. Proof of Theorem 3.1: A lower bound for the tail of the return time

In this section we establish a polynomial lower bound on the decay of the tail of the return time that is $m(\{x \in \Lambda : \tau(x) > n\})$. Consider the Markov partitions $\tilde{\mathcal{P}}$ and $\mathcal{P}$ for the automorphism A and the map $f_{\mathbb{T}^2}$ respectively and let $\tilde{P} \in \tilde{\mathcal{P}}$ and $P \in \mathcal{P}$ be the elements of the partitions as in Section 5.3. Fix the number Q as in Proposition 5.2.

We assume that the partition $\mathcal{P}$ and the number r_0 are chosen such that Proposition 5.1 holds and we set again $f := f_{\mathbb{T}^2}$. Finally, we denote by

$$\mathcal{N} = \{n \in \mathbb{N} : \text{ there is } x \in P \text{ such that } n = \tau(x)\}.$$

Lemma 7.1. *There exists an integer $Q_1 > 0$ such that for any $N > 0$ one can find $n > N$ with $n \in \mathcal{N}$, an s-subset Λ^s_ℓ with $\tau(\Lambda^s_\ell) = n$ and numbers $0 < m_1 < m_2$ satisfying $m_1 < Q_1$, $n - m_2 < Q_1$ such that $f^k(\Lambda^s_\ell) \cap D^1_{r_0} = \emptyset$ for all $0 \le k < m_1$ or $m_2 < k \le n$ and $f^k(\Lambda^s_\ell) \cap D^1_{r_0} \ne \emptyset$ for all $m_1 \le k \le m_2$.*

Proof. It suffices to show that there is $Q_1 > 0$ such that for any $N > 0$ there is an admissible word of length $n > N$ with $n \in \mathcal{N}$ of the form

$$P\bar{W}_1\bar{P}_i\bar{W}_2P, \tag{17}$$

where the words $\bar{W}_1$ and $\bar{W}_2$ are of length $l(\bar{W}_j) < Q_1$ for $j = 1, 2$ and do not contain any of the symbols P or P_k (the element of the Markov partition containing x_k for $k = 1, 2, 3, 4$), and the word $\bar{P}_i$ consists of the symbol P_i which is repeated $n - 2 - l(\bar{W}_1) - l(\bar{W}_2)$ times. Since the map f is topologically conjugate to A, it is enough to find an admissible word of the form (17) which consists of the corresponding elements of the partition $\tilde{\mathcal{P}}$.

Note that $A = B^3$ where B is an automorphism of the torus given by the matrix $B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. Therefore the result would follow if we find an admissible word of the type of (17) for the automorphism B. To this end consider the stable and unstable separatrices through the origin and denote the "first" connected component of their intersection with P by γ^s and γ^u respectively. It takes finitely many iterates of G and G^{-1} for each of these curves to completely enter the disk $D^1_{r_0}$. Now for each sufficiently large $n > 0$ with $n \in \mathcal{N}$ there is an s-set Λ^s_ℓ with $\tau(\Lambda^s_\ell) = n$ which completely enters $D^1_{r_0}$ (under iterates of G and G^{-1}) at the same time as γ^s and γ^u respectively. This completes the proof of the lemma. □

Lemma 7.2. *There exists a constant $C_8 > 0$ such that*

$$m(\{x \in \Lambda : \tau(x) > n\}) > C_8 n^{-(\gamma-1)},$$

where γ is defined in (8).

Proof. Using the conjugacy (6), it suffices to prove the lemma for the map G. Write

$$\begin{aligned} m(\{x\in\Lambda:\tau(x)>n\}) &= \sum_{N=n+1}^{\infty} m(\{x\in\Lambda:\tau(x)=N\}) \\ &= \sum_{N=n+1}^{\infty}\ \sum_{\Lambda_p^s:\tau(\Lambda_p^s)=N} m(\Lambda_p^s) > \sum_{N=n+1}^{\infty} m(\Lambda_\ell^s), \end{aligned}$$

where Λ_ℓ^s is the set constructed in Lemma 7.1. We wish to obtain a polynomial bound for the measure of the set Λ_ℓ^s.

Given $x\in\Lambda_\ell^s$, denote by $\gamma_\ell^s(x) := \gamma^s(x)\cap\Lambda_\ell^s$ (recall that $\gamma^s(x)$ is the full length stable curve through x in the element P of the Markov partition). There is $K_1>0$ such that

$$m(\Lambda_\ell^s) = m(G^N(\Lambda_\ell^s)) = K_1 L(G^N(\gamma_\ell^s(x))), \tag{18}$$

where L stands for the length of the curve.

Let $x_j = G^j(x)$ for $j=0,\dots,n$. Assume that x enters the region $D_{r_0}^1$ at time k_1 and exits at time k_2, i.e.,

(1) $G^j(x)\notin D_{r_0}^1$ if $0\le j<k_1$, or $k_2<j\le N$;
(2) $G^j(x)\in D_{r_0}^1$ if $k_1\le j\le k_2$.

Note that for $0\le j<k_1$ and $k_2<j\le N$ the curve $G^j(\gamma_\ell^s(x))$ lies in the stable cone for the automorphism A at x_j and indeed, is an admissible manifold for A (i.e., for any $y\in\gamma_\ell^s(x)$ the line $T_y\gamma_\ell^s(x)$ lies in the stable cone at y). So the length of the curve $\gamma_\ell^u(x)$ expands exponentially outside of the region D_{r_0}. Since by Lemma 7.1, $k_1<Q_1$ and $N-k_2<Q_1$, we have that

$$L(\gamma_\ell^s(x)) = \lambda^{k_1} L(G^{k_1}(\gamma_\ell^s(x))) \le \lambda^{Q_1} L(G^{k_1}(\gamma_\ell^s(x))) \tag{19}$$

and

$$L(G^N(\gamma_\ell^s(x))) = \lambda^{-(N-k_2)} L(G^{k_2}(\gamma_\ell^s(x))) \ge \lambda^{-Q_1} L(G^{k_2}(\gamma_\ell^s(x))), \tag{20}$$

where λ is the largest eigenvalue of the matrix A.

By Lemma 5.6 in [21], the time the trajectory spends in $D_{r_0}^1\setminus D_{\frac{r_0}{2}}^1$ is uniformly bounded. Thus, by Lemma 6.8,

$$L(G^{k_2}(\gamma_\ell^s(x))) > C_6(k_2-k_1)^{-\gamma} L(G^{k_1}(\gamma_\ell^s(x))). \tag{21}$$

Since $k_2-k_1<N$, combining Equations (18)-(21) yields

$$\begin{aligned} m(\Lambda_\ell^s) &\ge K_2 L(G^N(\gamma_\ell^s(x))) = K_2\lambda^{-Q_1} L(G^{k_2}(\gamma_\ell^s(x))) \\ &\ge K_2 C_6 \lambda^{-Q_1}(k_2-k_1)^{-\gamma} L(G^{k_1}(\gamma_\ell^s(x))) \\ &\ge K_2 C_6 \lambda^{-2Q_1}(k_2-k_1)^{-\gamma} L(\gamma_\ell^s(x)) \ge K_3 N^{-\gamma}, \end{aligned}$$

where $K_2 > 0$ is a constant and $K_3 = K_2 C_6 \lambda^{2Q_1} L(\gamma_\ell^u(x))$.

Note that $\gamma_\ell^s(x)$ is a full length stable curve in P and hence, has length which is independent of N. It follows that

$$m(\{x \in \Lambda : \tau(x) > n\}) > \sum_{N=n+1}^{\infty} m(\Lambda_\ell^s) > C_8 \frac{1}{n^{\gamma-1}},$$

where $C_8 > 0$ is a constant. The desired lower bound follows. □

8. Proof of Theorem 3.1: An upper bound for the tail of the return time

In this section we obtain an upper polynomial bound for the decay of the tail of the return time. As before we assume that the Markov partition and the number r_0 are chosen such that Proposition 5.1 holds. Recall that D_r is the union of the disks D_r^i around the points x_i and P_i is the element of the partition containing x_i, $i = 1, 2, 3, 4$. We have that $D_{r_0}^i \subset P_i$. Using the conjugacy (6), it suffices to establish that upper bound for the map G.

Given an s-set $\Lambda_i^s \subset P$ with $\tau(\Lambda_i^s) = n$, choose any numbers $k = k(\Lambda_i^s)$, $p = p(\Lambda_i^s)$, and two finite collections of numbers $\{k_m \geq 0\}_{m=1,\dots,p}$ and $\{l_m \geq 0\}_{l=0,\dots,p}$ such that

(1) $k_1 + k_2 + \cdots + k_p = k$ and $l_1 + l_2 + \cdots + l_{p+1} = n - k$;
(2) the trajectory of the set Λ_i^s under G^j, $0 \leq j \leq n$, consecutively spends l_m-times outside D_{r_0} and k_m-times inside D_{r_0}.

Given $0 < p < k < n$, consider the collections

$$\mathcal{S}_{k,n,p} = \{\Lambda_i^s \subset P : \tau(\Lambda_i^s) = n,\, k = k(\Lambda_i^s),\, p = p(\Lambda_i^s)\}.$$

Lemma 8.1. *There are $0 < h < h_{top}(f)$, $\varepsilon_0 > 0$, and $C_9 > 0$ such that $\varepsilon_0 < h_{top}(f) - h$ and*

$$\mathit{Card}\ \mathcal{S}_{k,n,p} \leq C_9 \frac{1}{p^2} e^{(h+\varepsilon_0)(n-k)}.$$

Proof. Note that the cardinality of $\mathcal{S}_{k,n,p}$ does not exceed the number of symbolic words of length n that start and end at P and contain exactly k symbols P_j, $j = 1, 2, 3, 4$. Since A is topologically mixing, the latter is exactly the number of words of length $n - k$ that start and end at P. By Corollary 1.9.12 and Proposition 3.2.5 in [13], the number of such words grows exponentially with an exponent that does not exceed $(n - k)h$ where $0 < h < h_{\text{top}}(A)$.

The number of different ways the iterates of Λ_i^s can enter D_{r_0} exactly p times and stay in this set exactly k times does not exceed the number of ways in which the number k can be written as a sum of p positive integers (where order matters) which is equal to $\binom{k-1}{p-1}$. The number of different ways the iterates of Λ_i^s can spend outside D_{r_0} exactly $p+1$ times is equal to the number of ways in which the number $n-k$ can be written as a sum of $p+1$ positive integers which is $\binom{n-k-1}{p}$. Since iterates of Λ_i^s may enter any of the disks $D_{r_0}^i$, $i = 1, 2, 3, 4$, we obtain

$$\text{Card } \mathcal{S}_{k,n,p} \le C_9 4^p \binom{k-1}{p-1} \binom{n-k-1}{p} e^{h(n-k)}.$$

Write

$$\text{Card } \mathcal{S}_{k,n,p} \le \frac{C_9}{p^2} p^2 4^p \binom{k-1}{p-1} \binom{n-k-1}{p} e^{h(n-k)}.$$

To prove the lemma we wish to estimate $p^2 4^p \binom{k-1}{p-1}\binom{n-k-1}{p}$ and we claim that there is $\varepsilon_0 > 0$ such that $\binom{k-1}{p-1} < e^{\varepsilon_0(n-k)}$.

To this end note that by Propositions 5.1 and 5.2, it takes Λ_i^s at least Q iterates before it enters D_{r_0} again. This implies that $n = k + l_1 + \cdots + l_{p+1} > k + (p+1)Q$ that is $p + 1 < \frac{n-k}{Q}$.

For a fixed k note that $\binom{k-1}{p-1}$ achieves its maximum when $p-1 = [\frac{k-1}{2}]$ or $p - 1 = [\frac{k-1}{2}] + 1$. We may assume $p - 1 = \frac{k-1}{2}$. Then using the asymptotic formula $\binom{m}{l} \sim (\frac{me}{l})^l$, we obtain that

$$\binom{k-1}{p-1} < \binom{2p-2}{p-1} < 4^{p-1} = e^{(p-1)\ln 4} < e^{\frac{n-k}{Q}\ln 4}.$$

To estimate $\binom{n-k-1}{p}$ observe that p does not exceed $\frac{n-k}{Q}$. Hence, using the above asymptotic formula, we find that

$$\binom{n-k-1}{p} < \binom{n-k}{\frac{n-k}{Q}} < \left(\frac{(n-k)e}{\frac{n-k}{Q}}\right)^{\frac{n-k}{Q}}$$
$$< e^{\frac{n-k}{Q}\ln\frac{(n-k)e}{\frac{n-k}{Q}}} < e^{\frac{n-k}{Q}\ln(Qe)}.$$

Finally, note that

$$p^2 4^p < e^{2\ln p + p\ln 4} < e^{2p + p\ln 4} < e^{\frac{n-k}{Q}(\ln 4 + 2)}.$$

Now, given any sufficiently small $\varepsilon_0 > 0$, one can choose Q large enough so that $\frac{\ln 4 + 2}{Q} + \frac{\ln 4}{Q} + \frac{\ln(Qe)}{Q} < \varepsilon_0$. Combining the above estimates we obtain

$$4^p \binom{k-1}{p-1}\binom{n-k-1}{p} e^{h(n-k)} < e^{(n-k)\varepsilon_0} e^{h(n-k)} = e^{(n-k)\varepsilon_0 + h}$$

and hence,

$$\text{Card } \mathcal{S}_{k,n,p} \leq C_9 \frac{1}{p^2} e^{(h+\varepsilon_0)(n-k)}.$$

This completes the proof of the lemma. □

Lemma 8.2. *There exists $\varepsilon_0 > 0$ such that for any $\Lambda_i^s \in \mathcal{S}_{k,n,p}$,*

$$m(\Lambda_i^s) \leq C_{10} k^{-\gamma'} e^{(-\log\lambda+\varepsilon_0)(n-k)},$$

where $C_{10} > 0$ is a constant and γ' is given by (8)*.*

Proof. Note that by (18)), $m(\Lambda_i^s) = m(G^n(\Lambda_i^s)) = K_1 L(G^n(\gamma_i^s(x)))$ and that the length of the backward iterates of γ_i^s lying outside the region D_{r_0} are stretched by the largest eigenvalue λ of the matrix A. Note also that every time the iterates of γ_i^s enter the region we have an upper estimate for its length according to Lemma 6.8 (note that we can apply this lemma in the region D_{r_0} since by Lemma 5.6 in [21] the time spent in $D_{r_0} \setminus D_{\frac{r_0}{2}}$ is uniformly bounded). Thus,

$$\begin{aligned} m(\Lambda_i^s) &= m(G^n(\Lambda_i^s)) = K_1 L(G^n(\gamma_i^s(x))) \\ &= K_1 \lambda^{-l_{p+1}} L(G^{n-l_{p+1}}(\gamma_i^s(x))) \\ &\leq K_1 C_7 \lambda^{-l_{p+1}} k_p^{-\gamma'} L(G^{n-(l_1+k_1)}(\gamma_i^s(x))) \leq \cdots \\ &\leq K_1 C_7^p \lambda^{-(l_{p+1}+\cdots+l_1)} k_p^{-\gamma'} k_{p-1}^{-\gamma'} \cdots k_1^{-\gamma'} L(\gamma_i^s(x)). \end{aligned}$$

Note that $\gamma_i^s(x)$ is a full length stable curve in P and hence, has length independent of n. One can also assume that $k_i \geq 2$ by making r_0 smaller if necessary. This implies

$$k_1 k_2 \cdots k_p \geq k_{\max} 2^{p-1} \geq k_{\max} p \geq \sum_{i=1}^{p} k_i = k,$$

where $k_{\max}$ denotes the largest of $k_i's$.

In addition, $C_6^p = e^{p \ln C_6} < e^{\frac{n-k}{Q} \ln C_6} < e^{\varepsilon_0 (n-k)}$ for sufficiently small $\varepsilon_0 > 0$ if one chooses Q large. Therefore,

$$m(\Lambda_i^s) < K_1 e^{\varepsilon_0(n-k)} \lambda^{-(n-k)} k^{-\gamma'} < C_{10} k^{-\gamma'} e^{(-\log\lambda+\varepsilon_0)(n-k)}.$$

This completes the proof of the lemma. □

Lemma 8.3. *There exists $C_{11} > 0$ such that*

$$m(\{x \in \Lambda : \tau(x) > n\}) < C_{11} n^{-(\gamma'-1)},$$

see (8) *for the definition of γ'.*

Note that

$$m(\{x \in \Lambda : \tau(x) = n\}) \le \sum_{k=1}^{n} \sum_{p=1}^{k} \max_{\Lambda_i^s \in \mathcal{S}_{k,n,p}} \{m(\Lambda_i^s)\} \text{Card } \mathcal{S}_{k,n,p}.$$

Therefore, by Lemmas 8.1 and 8.2, we have

$$\begin{aligned} m(\{x \in \Lambda : \tau(x) = n\}) & \\ & \le \sum_{k=1}^{n} \sum_{p=1}^{k} \frac{1}{p^2} C_9 e^{(h+\varepsilon_0)(n-k)} C_{10} e^{(\varepsilon_0 - \log \lambda)(n-k)} k^{-\gamma'} \\ & < C_9 C_{10} \frac{\pi^2}{6} e^{-\delta n} \sum_{k=1}^{n} e^{\delta k} k^{-\gamma'}, \end{aligned} \tag{22}$$

where $\delta = 2\varepsilon_0 + \log \lambda - h > 0$ if ε_0 is sufficiently small.

To estimate $\sum_{k=1}^{n} e^{\delta k} k^{-\gamma'}$ set $u_k = e^{\delta k} k^{-\gamma'}$ and note that $u_{k+1} - u_k \sim e^{\delta k} k^{-\gamma'} = u_k$. Since $\sum_{k=1}^{n} u_k$ is positive and diverges, by Stolz-Cesaro theorem,

$$\sum_{k=1}^{n} u_k \sim \sum_{k=1}^{n} u_{k+1} - u_k = u_{n+1} - u_1 \sim e^{\delta n} n^{-\gamma'}.$$

Therefore,

$$m(\{x \in \Lambda : \tau(x) = n\}) \le C_9 C_{10} e^{-\delta n} \sum_{k=1}^{n} e^{\delta k} k^{-\gamma'} < C_9 C_{10} n^{-\gamma'}.$$

Thus, we have the following estimate of the tail

$$m(\{x \in \Lambda : \tau(x) > n\}) = \sum_{k>n} m(\{x \in \Lambda : \tau(x) = k\}) < C_{11} n^{-(\gamma'-1)}$$

for some $C_{11} > 0$. This concludes the proof of the Lemma and the upper bound.

9. Proof of Theorem 3.1: Carrying the slow-down map to a surface

In this section we show how to carry over the slow-down map of the torus to a measure preserving diffeomorphism of any surface. Following [12], we will construct the maps $\varphi_1, \varphi_2, \varphi_3$ such that the following diagram is commutative:

$$\begin{array}{ccccccc}
\mathbb{T}^2 & \xrightarrow{\varphi_1} & S^2 & \xrightarrow{\varphi_2} & D^2 & \xrightarrow{\varphi_3} & M \\
{\scriptstyle f_{\mathbb{T}^2}}\downarrow & & \downarrow{\scriptstyle f_{S^2}} & & \downarrow{\scriptstyle f_{D^2}} & & \downarrow{\scriptstyle f_M} \\
\mathbb{T}^2 & \xrightarrow{\varphi_1} & S^2 & \xrightarrow{\varphi_2} & D^2 & \xrightarrow{\varphi_3} & M
\end{array}$$

We stress that while our construction of the maps φ_1 and φ_2 follows [12], our construction of the map φ_3 is quite different, since we have to deal with finite regularity of the slow-down map.

First, using the slow down map we construct a diffeomorphism of the sphere S^2.

Proposition 9.1 (see [12])**.** *There exists a map $\varphi_1 \colon \mathbb{T}^2 \to S^2$ satisfying:*

(1) *φ_1 is a double branched covering, is one-to-one on each branch, and C^∞ everywhere except at the points x_i, $i = 1, 2, 3, 4$ where it branches;*
(2) *$\varphi_1 \circ I = \varphi_1$ where $I \colon \mathbb{T}^2 \to \mathbb{T}^2$ is the involution map given by $I(t_1, t_2) = (1 - t_1, 1 - t_2)$;*
(3) *φ_1 preserves area, i.e., $(\varphi_1)_* m = m_{S^2}$ where m_{S^2} is the area in S^2;*
(4) *there exists a coordinate system in each disk $D^i_{r_0}$ such that*
$$\varphi_1(s_1, s_2) = \left(\frac{{s_1}^2 - {s_2}^2}{\sqrt{{s_1}^2 + {s_2}^2}}, \frac{2 s_1 s_2}{\sqrt{{s_1}^2 + {s_2}^2}} \right);$$
(5) *The map $f_{S^2} := \varphi_1 \circ f_{\mathbb{T}^2} \circ \varphi_1^{-1}$ preserves the area.*

The sphere can be unfolded onto the unit disk D^2 and the map f_{S^2} can be carried over to an area preserving map f_{D^2} of the disk which is identity on the boundary of the disk. To see this set $p_i = \varphi_1(x_i)$, $i = 1, 2, 3, 4$. In a small neighborhood of the point p_4 we define a map φ_2 by

$$\varphi_2(\tau_1, \tau_2) = \left(\frac{\tau_1 \sqrt{1 - \tau_1^2 - \tau_2^2}}{\sqrt{\tau_1^2 + \tau_2^2}}, \frac{\tau_2 \sqrt{1 - \tau_1^2 - \tau_2^2}}{\sqrt{\tau_1^2 + \tau_2^2}} \right).$$

One can extend φ_2 to an area preserving C^∞ diffeomorphism (still denoted by φ_2) between $S^2 \setminus \{p_4\}$ and the interior of the unit disk D^2. The map

$$f_{D^2} := \begin{cases} \varphi_2 \circ f_{S^2} \circ \varphi_2^{-1} & \text{on } \operatorname{int} D^2 \\ Id & \text{on } \partial D^2 \end{cases} \tag{23}$$

is a diffeomorphism of D^2 that preserves area m_{D^2}.

Proposition 9.2. *The maps f_{S^2} and f_{D^2} are of class of smoothness $C^{2+2\kappa}$ where $\kappa = \frac{\alpha}{1-\alpha}$.*

Proof. Using the explicit local expressions for f_{S^2} and f_{D^2} and following arguments in Proposition 4.2, we find that the maps f_{S^2} and f_{D^2} are Hamiltonian with respect to the area and the Hamiltonian functions are given as

$$H_3(\tau_1, \tau_2) = \frac{\tau_2 h(\sqrt{\tau_1^2 + \tau_2^2})}{\sqrt{\tau_1^2 + \tau_2^2}} \log \lambda$$

and

$$H_4(x_1, x_2) = \frac{x_2 h(\sqrt{1 - x_1^2 - x_2^2})}{\sqrt{x_1^2 + x_2^2}} \log \lambda$$

respectively. Here, as before, $h(u) = u^{\frac{2}{1-\alpha}}$.

To show that the maps f_{S^2} and f_{D^2} are of the desired class of smoothness, we will show that the Hamiltonian functions H_3 and H_4 have Hölder continuous second order partial derivatives with Hölder exponent 2κ. Since H_3 and H_4 are of the same regularity we consider only one of them and we set $g(x, y) = y(x^2 + y^2)^\delta$ where $\delta = \frac{1}{1-\alpha} - \frac{1}{2}$.

Obviously, $\frac{\partial g}{\partial x}(0, 0) = 0$ and

$$\frac{\partial g}{\partial x}(x, y) = \begin{cases} 2\delta xy(x^2 + y^2)^{\delta-1}, & (x, y) \neq (0, 0); \\ 0, & (x, y) = (0, 0). \end{cases}.$$

The function $\frac{\partial g}{\partial x}$ is symmetric, so we will only study Hölder continuity of $\frac{\partial^2 g}{\partial x^2}$ as Hölder continuity of $\frac{\partial^2 g}{\partial x \partial y}$ is immediate. Note that

$$\frac{\partial^2 g}{\partial x^2}(x, y) = \begin{cases} 2\delta \frac{y(x^2+y^2)^{1-\delta} - 2(1-\delta)x^2 y(x^2+y^2)^{-\delta}}{(x^2+y^2)^{2-2\delta}} & (x, y) \neq (0, 0); \\ 0 & (x, y) = (0, 0). \end{cases}$$

Since the function $\frac{\partial^2 g}{\partial x^2}(x, y)$ is differentiable for all $(x, y) \neq (0, 0)$, it is Hölder continuous for all pairs of nonzero points (x, y). It remains to show Hölder continuity for pairs of points one of which is zero. We can write

$$\frac{\partial^2 g}{\partial x^2} = K_1 y(x^2 + y^2)^{\delta-1} - K_2 x^2 y(x^2 + y^2)^{\delta-2},$$

where $K_1 > 0$ and $K_2 > 0$ are some constants. Choose $(x, y) \neq (0, 0)$ and note that

$$|y(x^2+y^2)^{\delta-1}| \leq (x^2+y^2)^{\frac{1}{2}}(x^2+y^2)^{\delta-1} = (x^2+y^2)^{\delta-\frac{1}{2}} = d((x, y), (0, 0))^{2\delta-1}.$$

Similarly,

$$|x^2 y(x^2 + y^2)^{\delta-2}| \leq d((x, y), (0, 0))^{2\delta-1}.$$

Hence, $\frac{\partial^2 g}{\partial x^2}$ is Hölder continuous with Hölder exponent

$$2\delta - 1 = 2\Big(\frac{1}{1-\alpha} - \frac{1}{2}\Big) - 1 = 2\kappa.$$

Further, each of the functions $\frac{\partial^2 g}{\partial y^2}$ and $\frac{\partial^2 g}{\partial y \partial x}$ can be written as a linear combinations of functions

$$x(x^2+y^2)^{\delta-1}, \quad y(x^2+y^2)^{\delta-1}, \quad x^2y(x^2+y^2)^{\delta-2}, \quad y^2x(x^2+y^2)^{\delta-2} \tag{24}$$

and are 0 at the point $(x, y) = (0, 0)$. Arguing as above one can show that each function in (24) is Hölder continuous with Hölder exponent 2κ. This completes the proof of the proposition. □

Consider a smooth compact connected oriented surface M. It can be cut along closed geodesics in such a way that the resulting surface with boundary is homeomorphic to a regular polygon via a homeomorphism which we denote by T; this is a well-known topological construction.

Let now f be a C^∞ diffeomorphism of the disk, which is identity on the boundary ∂D^2 and is *infinitely flat*, i.e., given a sequence $\rho_n \to 0$ and a sequence of open domains $V_n \subset D^2$ satisfying

$$V_n \subset V_{n+1} \text{ and } \bigcup_{n\ge 1} V_n = D^2, \tag{25}$$

we have that for every $n \ge 1$,

$$\|f - \mathrm{Id}\|_{C^n(V_n)} \le \rho_n.$$

For such an f it is shown in [12] (see also [2]) that there is a homeomorphism $\varphi : \overline{D^2} \to M$ such that

(1) φ is of class C^∞ in the interior of the disk;
(2) φ is area preserving, i.e., $h_* m_{D^2} = m_M$;
(3) the map $\varphi \circ f \circ \varphi^{-1}$ is a C^∞ area preserving diffeomorphism of the surface.

In our case however, the map $f = f_{D^2}$ is only of class $C^{2+2\kappa}$ and hence, is only *finitely flat* at the boundary, i.e., there is a sequence of open domains $V_n \subset D^2$, satisfying (25), such that for every $0 < \beta < 2 + 2\kappa$,

$$\|f_{D^2} - \mathrm{Id}\|_{C^{1+\beta}(V_n)} \le (r_{n-1})^{2+2\kappa-\beta}, \tag{26}$$

where $r_n = \mathrm{dist}(V_n, \partial D^2)$. This requires us to develop a specific construction of the homeomorphism φ which guarantees that the map f_M is an area preserving diffeomorphism of class $C^{1+\beta}$ for some $\beta > 0$. More precisely, the following statement holds.

Theorem 9.3. *Given a smooth compact connected oriented surface* M *and numbers* $\frac{1}{9} < \alpha < \frac{1}{4}$ *and* $0 < \mu < \frac{1}{2}$, *there exist* $\beta = \beta(\alpha, \mu) > 0$ *and a continuous map* $\varphi_3 \colon \overline{D^2} \to M$ *such that*

(1) *the restriction* $\varphi_3 | \text{int } D^2$ *is a diffeomorphic embedding;*
(2) $\varphi_3(\overline{D^2}) = M$;
(3) φ_3 *preserves area; more precisely,* $(\varphi_3)_* m_{D^2} = m_M$ *where* m_M *is the area in* M*; moreover,* $m_M(M \setminus \varphi_3(\text{int } D^2)) = 0$;
(4) *the map* $f_M := \varphi_3 \circ f_{D^2} \circ \varphi_3^{-1}$ *is a* $C^{1+\beta}$ *area preserving diffeomorphism of the surface.*

Proof. One can represent a compact smooth oriented surface M as a regular p-polygon P (the number p is even) whose angles are $\alpha = \frac{\pi(p-2)}{p}$. Let $A_1, A_2, \dots, A_p$ be vertices of the polygon and O its center. For each $i = 1, \dots, p$ denote by B_i the points on the segment A_iO for which $\frac{|A_iB_i|}{|A_iO|} = \frac{1}{3}$. In what follows we assume that $A_{p+1} = A_1$ and $B_{p+1} = B_1$. Denote by

$$P^* := \Big(\bigcup_{i=1}^{p} A_iA_{i+1}\Big) \cup \Big(\bigcup_{i=1}^{p} A_iB_i\Big). \tag{27}$$

Note that the complement to P^* is an open simply connected set.

We now construct a homeomorphism from the unit disk D^2 onto P.

Proposition 9.4. *There exist*

(1) *a nested sequence of open simply connected sets* $U_0 \subset U_1 \subset \cdots \subset U_n \subset \cdots$ *satisfying* $\bigcup_n U_n = P \setminus P^*$;
(2) *a sequence of* C^∞ *diffeomorphisms* $h_n : U_n \to U_{n+1}$ *for* $n \ge 0$;
(3) *a number* $\beta > 0$

such that setting $h(x) = \lim_{n\to\infty} h_{n-1} \circ \cdots \circ h_1 \circ h_0$, *we have that the map* $f_P : P \to P$ *given by*

$$f_P = \begin{cases} (h \circ f_{D^2} \circ h^{-1})(x), & x \in P \setminus P^*, \\ Id & \textit{otherwise} \end{cases}$$

is a $C^{1+\beta}$ *diffeomorphism.*

Proof of the proposition. We split the proof into three steps.

Step 1. We first construct a sequence of open sets U_n (see Figure 1)

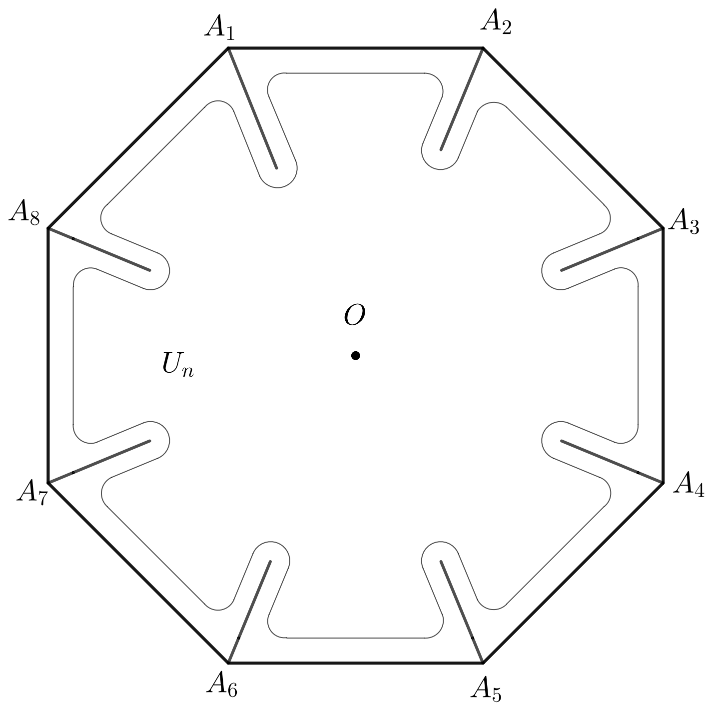

Figure 1: The shape of the set U_n

Fix $n > 0$, $t \in [n, n+1]$ and let $r(t) > 0$ be a strictly monotonically decreasing continuous function on $[1, \infty)$ which will be determined later in Step 4. For $i = 1, \dots, p$ consider the following collection of points associated with the point A_i (see Figure 2):

- K_{ti}, the point that is determined uniquely by the requirements that the angle $\angle(K_{ti}A_iA_{i+1}) = \frac{1}{8}\alpha$ and $\mathrm{dist}(K_{ti}, A_iA_{i+1}) = r(t)$;
- L_{ti}, the point that is determined uniquely by the requirements that the angle $\angle(OA_iL_{ti}) = \frac{1}{8}\alpha$ and $\mathrm{dist}(L_{ti}, A_iO) = r(t)$;
- M_{ti}, the image of L_{ti} under the reflection about the line OA_i;
- N_{ti}, the image of K_{ti} under the reflection about the line OA_i;
- E_{ti}, the point on the line through B_i which is perpendicular to the line A_iB_i and such that $\mathrm{dist}(E_{ti}, B_i) = r(t)$;
- F_{ti}, the image of E_{ti} under the reflection about the line A_iB_i.

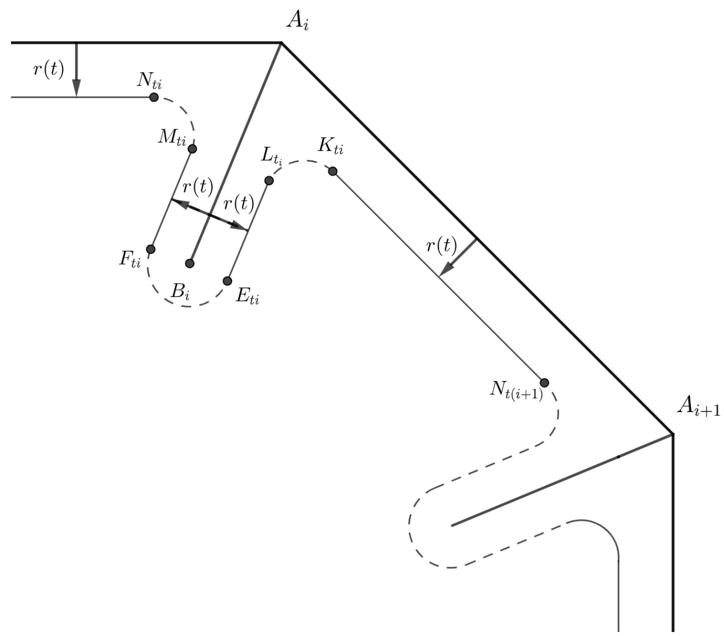

Figure 2: The collection of marked points

We introduce the following curves: for $i = 1, \dots, p$ let

- $\gamma_{ti}^{(1)}$ be a line segment connecting the points K_{ti} and $N_{t(i+1)}$ where we assume that $N_{t(p+1)} = N_{t1}$;
- $\gamma_{ti}^{(2)}$ be a curve connecting the points K_{ti} and L_{ti} to be determined later in Step 2;
- $\gamma_{ti}^{(3)}$ be a curve connecting the points M_{ti} and N_{ti} to be determined later in Step 2.
- $\gamma_{ti}^{(4)}$ be the line segment connecting the points L_{ti} and E_{ti};
- $\gamma_{ti}^{(5)}$ be the line segment connecting the points M_{ti} and F_{ti};
- $\gamma_{ti}^{(6)}$ be a curve connecting the points E_{ti} and F_{ti} to be determined later in Step 2.

Let τ_t be the curve given by

$$\tau_t = \bigcup_{i=1}^{p} \bigcup_{j=1}^{6} \gamma_{ti}^{(j)}.$$

By construction, τ_t, $t \in [n, n+1]$, is a closed connected continuous curve which bounds an open simply connected domain in the polygon

P. We denote this domain by U_t. In particular, U_n is the desired open set.

Step 2. We show how to choose the curves $\gamma_{ti}^{(2)}$ and $\gamma_{ti}^{(6)}$. The curve $\gamma_{ti}^{(3)}$ can be chosen in a similar way.

Let $\varphi, \psi : [-1, 1] \to \mathbb{R}$ be two continuous functions satisfying:

(1) $\varphi \in C^\infty$ on $[0, 1]$ and $\psi \in C^\infty$ on $(0, 1)$;
(2) $\varphi(-x) = \varphi(x)$ and $\psi(-x) = \psi(x)$;
(3) $\varphi(0) = a$ where $1 - \tan\frac{\alpha}{4} < a < \cot\frac{\alpha}{4}$ and $\varphi(-1) = \varphi(1) = \cot\frac{\alpha}{8}$;
(4) $\psi(0) = 1$ and $\psi(-1) = \psi(1) = 0$;
(5) $0 < \varphi'(x) \le \cot\frac{\alpha}{4}$ for $0 < x \le 1$ and $-\cot\frac{\alpha}{4} \le \varphi'(x) < 0$ for $-1 \le x < 0$;
(6) $\varphi'(-1) = -\cot\frac{\alpha}{4}$ and $\varphi'(1) = \cot\frac{\alpha}{4}$;
(7) ψ is infinitely vertically flat at -1 and 1.

For $i = 1, \dots, p$ consider the orthogonal coordinate system with origin at A_i whose vertical axis is the bisector of the angle $\angle(B_i A_i A_{i+1})$ and for every $t \in [n, n+1]$ we let $\gamma_{ti}^{(2)}$ be the graph of the function $\varphi_t(x) = r(t)\varphi(\frac{x}{r(t)})$ where $-r(t) \le x \le r(t)$. It is easy to see that $\gamma_{ti}^{(2)}$ is a C^∞ curve that connects the points K_{ti} and L_{ti} and is infinitely tangent to the lines $K_{ti}N_{ti}$ and $L_{ti}E_{ti}$.

Now consider the orthogonal coordinate system with origin at B_i whose vertical axis is the line A_iB_i. We let $\gamma_{ti}^{(6)}$ be the graph of the function $\psi_t(x) = r(t)\psi(\frac{x}{r(t)})$ where $-r(t) \le x \le r(t)$. It is easy to see that $\gamma_{ti}^{(6)}$ is a C^∞ curve that connects the points E_{ti} and F_{ti} and is infinitely tangent to the lines $L_{ti}E_{ti}$ and $M_{ti}F_{ti}$. Hence, with the above choice of curves γ_{ti}^j, $j = 1, \dots, 6$, the curve τ_t is of class C^∞.

We show that the curves τ_t corresponding to different values of t are disjoint. To this end fix $n \le t_1 < t_2 \le n+1$. It suffices to show that the curves $\gamma_{ti}^{(2)}$, $\gamma_{ti}^{(6)}$, and $\gamma_{ti}^{(3)}$ with $t = t_1$ and $t = t_2$ are disjoint. We will prove this for the curve $\gamma_{ti}^{(2)}$ only as the proof for other curves is similar. By Property (3),

$$\varphi_{t_1}(0) = ar(t_1), \quad \varphi_{t_1}(r(t_1)) = r(t_1)\cot\frac{\alpha}{8}, \quad \varphi'_{t_1}(r(t_1)) = \cot\frac{\alpha}{4}.$$

Similarly,

$$\varphi_{t_2}(0) = ar(t_2), \quad \varphi_{t_2}(r(t_2)) = r(t_2)\cot\frac{\alpha}{8}, \quad \varphi'_{t_2}(r(t_2)) = \cot\frac{\alpha}{4}.$$

In view of Property (5) the desired result would follow if we show that

$$\varphi_{t_1}(r(t_2)) = r(t_1)\varphi\Big(\frac{r(t_2)}{r(t_1)}\Big) \ge \varphi_{t_2}(r(t_2)) = r(t_2)\cot\frac{\alpha}{8}.$$

Setting $x = \frac{r(t_2)}{r(t_1)}$, the above inequality amounts to $\varphi(x) \geq x \cot \frac{\alpha}{8}$ and immediately follows from Properties (3) and (4) of the function φ.

Step 3. We now construct maps h_n. By the Riemann Mapping theorem, there is a C^∞ diffeomorphism $h_0 : D^2 \to U_1$. For each $n = 1, 2, 3, \dots$ we will construct maps $h_n : U_n \to U_{n+1}$ such that $h_n \big| U_{n-1} = \mathrm{Id}$.

Given two numbers $n-1 \leq s \leq n$ and $n-1 \leq t \leq n+1$ such that $s < t$ we construct a C^∞ diffeomorphism $\hat{h}_{st} : \tau_s \to \tau_t$ in the following way.

- $\hat{h}_{st} : \gamma_{si}^{(1)} \to \gamma_{ti}^{(1)}$ is a linear map, given by $\hat{h}_{st}(z) = \frac{r(t)}{r(s)}(z)$, $z \in \gamma_{si}^{(1)}$ and $i = 1, \dots, p$;
- $\hat{h}_{st} : \gamma_{si}^{(j)} \to \gamma_{ti}^{(j)}$ is a map, given by $\hat{h}_{st}(z) = v$, where $z = (y, \varphi_s(y))$ and $v = (\frac{r(t)}{r(s)} y, \varphi_t(y))$ for $-r(s) \leq y \leq r(s)$, $i = 2, \dots, p$, $j = 2, \dots, 6$.

Now given $n-1 \leq s \leq n$, define the map $\hat{h}_s := \hat{h}_{st}$ with $t = 2(s-n+1)+n-1$. The desired map $h_n : U_n \to U_{n+1}$ is now given as follows: for $A \in U_n$ choose a unique s such that $A \in \tau_s$ with $n-1 \leq s \leq n$ and then set $h_n(A) = B$ where $B = \hat{h}_s(A) \in \tau_{2(s-n+1)+n-1} \subset U_{n+1}$. It is easy to see that h_n is a C^∞ diffeomorphism.

It follows that the map $h = \lim_{n\to\infty} h_{n-1} \circ \cdots \circ h_1 \circ h_0$ is a well defined C^∞ diffeomorphism from $\mathrm{int} D^2$ onto $P \setminus P^*$ where P^* is given by (27). It also follows from the construction of the map h that there is $C > 0$ such that

$$\|h\|_{C^1} \leq C, \quad \|h^{-1}\|_{C^1} \leq C. \tag{28}$$

Step 4. It remains to show that the map $f_P = h \circ f_{D^2} \circ h^{-1}$ is a $C^{1+\beta}$ diffeomorphism for some $\beta > 0$.

Observe that $f_P(P \setminus P^*) = P \setminus P^*$, $f_P(P^*) = P^*$, and $f_P|P^* = \mathrm{Id}$. In particular, $f_P|P \setminus P^*$ is a C^∞ diffeomorphism. It remains to show that f_P is of class $C^{1+\beta}$ on P^*. To do so we will show the following:

$$\|f_P - \mathrm{Id}\|_{C^{1+\beta}(U_{n+1} \setminus U_n)} \to 0 \tag{29}$$

as $n \to \infty$.

First, we will prove the following lemma.

Lemma 9.5. *Let $r(n)$ be a decreasing sequence such that $0 < r(1) < 1$ and*

$$r(n+1) = r^2(n), \quad 0 < r(1) < 1. \tag{30}$$

Then the sequence of open sets $V_n = h^{-1}(U_n)$ *satisfies* (25) *and*

$$V_{n-1} \subset f_{D^2}(V_n) \subset V_{n+1}.$$

Proof of the Lemma. Note that there are $C_2 \geq C_1 > 0$ such that

$$C_1 r(n) \leq \operatorname{dist}(U_n, P^*) \leq C_2 r(n).$$

Furthermore, in view of (28) there are $C_4 \geq C_3 > 0$ such that

$$C_3 r(n) \leq \operatorname{dist}(V_n, \partial D^2) \leq C_4 r(n).$$

Since the map f_{D^2} is identity on ∂D^2 and is of class of smoothness $2+2\kappa$, we obtain that for all sufficiently small r_n, any x in the neighborhood $U_{r_n}(\partial D^2)$ and any $\beta > 0$

$$\operatorname{dist}(x, f_{D^2}(x)) < r(n)^{2+2\kappa-\beta}.$$

Therefore, for some $0 < a < 1$,

$$C_3 r(n) - r(n)^{2+a} \leq \operatorname{dist}(f_{D^2}(x), \partial D^2) \leq C_4 r(n) + r(n)^{2+a}.$$

To prove the desired inclusion we will show that

$$C_4 r(n+2) < C_3 r(n) - r(n)^{2+a} \leq C_4 r(n) + r(n)^{2+a} < C_3 r(n-2).$$

We prove the leftmost inequality. Since $r(n+2) = r^4(n)$, we have

$$\begin{gathered} C_4 r(n+2) < C_3 r(n) - r(n)^{2+a} \Leftrightarrow \\ C_4 \frac{r(n+2)}{r(n)} < C_3 - r(n)^{1+a} \Leftrightarrow \\ C_4 r(n)^3 < C_3 - r(n)^{1+a}. \end{gathered}$$

For large values of n both $C_4 r(n)^3$ and $r(n)^{1+a}$ are small. Hence, the last inequality holds.

We now prove the rightmost inequality (the inequality in the middle is obvious).

$$\begin{gathered} C_4 r(n) + r(n)^{2+a} < C_3 r(n-2) \Leftrightarrow \\ C_4 + r(n)^{1+a} < C_3 \frac{r(n-2)}{r(n)} \Leftrightarrow \\ C_4 + r(n)^{1+a} < C_3 r(n)^{-\frac{3}{4}}. \end{gathered}$$

For large values of n the left hand side of the last inequality is close to C_4 while the right hand side gets large. This completes the proof of the lemma.

The above lemma allows us to write

$$\begin{gathered} \|h \circ (f_{D^2} - \mathrm{Id}) \circ h^{-1}\|_{C^{1+\beta}(U_{n+1} \setminus U_n)} \leq \\ \|h\|_{C^{1+\beta}(V_{n+2} \setminus V_{n-1})} \|f_{D^2} - \mathrm{Id}\|_{C^{1+\beta}(V_{n+1} \setminus V_n)} \|h^{-1}\|_{C^{1+\beta}(U_{n+1} \setminus U_n)}. \end{gathered} \tag{31}$$

Observe that $\|f_{D^2} - \mathrm{Id}\|_{C^{1+\beta}}$ admits Estimate (26) and it remains to estimate the norms $\|h\|_{C^{1+\beta}(V_{n+2}\setminus V_{n-1})}$ and $\|h^{-1}\|_{C^{1+\beta}(U_{n+1}\setminus U_n)}$.

We write $h|_{V_{n+2}} = h_{n+1}|_{U_{n+1}}$, so in order to estimate the norm of h we will estimate the norm of h_{n+1}. In order to estimate the norm of h^{-1} we will need to estimate the norms of h_n^{-1} for each $n = 0, 1, 2, \dots$.

Further, it suffices to estimate the norm of h restricted to the boundary of the sets U_n. Recall that τ_n, the boundary of U_n, is a union of the curves $\gamma_{ni}^{(j)}$, $i = 1, \dots, p$, $j = 1, \dots, 6$. Note that the map h acts linearly on the curves $\gamma_{ni}^{(1)}$ and hence, the norm of h restricted to these parts of the curve τ_n is bounded.

The curves $\gamma_{ni}^{(2)}$ and $\gamma_{ni}^{(3)}$ are the graphs of the function φ_n and the curves $\gamma_{ni}^{(6)}$ are the graphs of the function ψ_n. We shall only give an estimate of the norm of h restricted to the curves $\gamma_{ni}^{(2)}$, since the estimates of the norm of h restricted to other curves are similar. Also, we can assume that i and $j > 1$ are fixed.

Now for a fixed curve $\gamma_{ni}^{(2)}$ we define the orthogonal coordinate system centered at the vertex A_i with the vertical axis A_iO (recall that O is the center of the polygon P). In this coordinate system the map h_n is given by

$$h_n : (x, \varphi_n(x)) \to \Big(\frac{r(n+1)}{r(n)}x, \varphi_{n+1}(x)\Big), \ -r(n) \le x \le r(n).$$

Since φ_n is symmetric, we can further assume that $x > 0$. We have that

$$\|h_n\|_{C^{1+\beta}} = \max\Big(\|h_n\|_{C^0}, \|dh_n\|_{C^0}, \|dh_n\|_{C^\beta}\Big),$$

where

$$\|dh_n\|_{C^\beta} = \max\Big(\sup \frac{\|\partial_x h_n(x) - \partial_y h_n(y)\|}{\|x-y\|^\beta}, \sup \frac{\|\partial_x h_n(x) - \partial_y h_n(y)\|}{\|x-y\|^\beta}\Big)$$

and $\partial_x h_n$ and $\partial_y h_n$ are the partial derivatives of h with respect to the first and the second variables respectively.

Let us write $h_n = (h_n^{(1)}, h_n^{(2)})$, $y = \varphi_n(x)$ where

$$h_n^{(1)}(x,y) = \frac{r(n+1)}{r(n)}x \ \text{ and } \ h_n^{(2)}(x,y) = \varphi_{n+1}(x).$$

Since

$$h_n^{(1)}(x,y) = \frac{r(n+1)}{r(n)}x \le r(n+1) < 1 \ \text{ and } \ h_n^{(2)}(x,y) = \varphi_{n+1}(x) < K_1,$$

we obtain that $\sup \|h_n\| < K_1$.

Further, we have that $\partial_x h_n^{(1)}(x,y) = \frac{r(n+1)}{r(n)} < 1$ and

$$\begin{aligned}\partial_y h_n^{(2)}(x,y) &= \frac{\partial}{\partial y}\Big(\frac{r(n+1)}{r(n)}x\Big) = \frac{\partial}{\partial y}\Big(\frac{r(n+1)}{r(n)}\varphi_n^{-1}(y)\Big)\\ &= r(n+1)(\varphi^{-1})'\Big(\frac{x}{r(n)}\Big)\frac{1}{r(n)} < 1.\end{aligned}$$

Also, $\partial h_n^{(2)}(x,y) = \varphi'_{n+1}(x) < \cot\frac{\alpha}{4}$ and

$$\begin{aligned}\partial h_n^{(2)}(x,y) &= \frac{\partial}{\partial y}\Big(r(n+1)\varphi\Big(\frac{r(n)\varphi^{-1}(\frac{y}{r(n)})}{r(n+1)}\Big)\Big)\\ &= r(n+1)\varphi'\frac{r(n)}{r(n+1)}\varphi^{-1}\Big(\frac{y}{r(n)}\Big)(\varphi^{-1})'\frac{1}{r(n)} < K_2.\end{aligned}$$

Thus, the partial derivatives of the functions h_1 and h_2 are bounded. However, a similar calculation shows that $\|dh_n\|_{C^\beta}$ tends to infinity as $r(n)^{-\beta}$. Therefore, we conclude that

$$\|h\|_{C^{1+\beta}(V_{n+2}\setminus V_{n-1})} \le r(n+1)^{-\beta} = r(n)^{-2\beta}. \tag{32}$$

Similar computations holds for $h_n^{-1} : U_{n+1} \to U_n$, with the only difference that the partial derivatives estimated by $\frac{r(n)}{r(n+1)}$ which is unbounded. Therefore, the Hölder norm of the first derivatives of h_n^{-1} are bounded by

$$\frac{r(n)}{r(n+1)}\frac{1}{r(n)^{-\beta}} = \frac{r(n)^{1-\beta}}{r(n+1)}.$$

It follows that

$$\begin{aligned}\|h^{-1}\|_{C^{1+\beta}(U_{n+1}\setminus U_n)} &= \|h_n^{-1}\circ h_{n-1}^{-1}\circ\cdots\circ h_0^{-1}\|_{C^{1+\beta}(U_{n+1}\setminus U_n)}\\ &\le \prod_{i=0}^{n}\frac{r(n)^{1-\beta}}{r(n+1)} = \prod_{i=0}^{n} r(n)^{-1-\beta}.\end{aligned} \tag{33}$$

Since $r(n-1) = r(n)^{\frac{1}{2}}$, we obtain that

$$\prod_{i=0}^{n} r(n)^{-1-\beta} = r(n)^{-(1+\beta)(1+\frac{1}{2}+\frac{1}{4}+\cdots+\frac{1}{2^{n+1}})} = r(n)^{-2(1+\beta)(1-\frac{1}{2^{n+2}})}. \tag{34}$$

Finally, using (31), we find that

$$\begin{aligned}&\|h\circ(f_{D^2}-\mathrm{Id})\circ h^{-1}\|_{C^{1+\beta}(U_n\setminus U_{n-1})}\\ &\qquad\le r(n)^{-2\beta}r(n)^{2+2\kappa-\beta}r(n)^{-2(1+\beta)(1-\frac{1}{2^{n+2}})}\\ &\qquad= r(n)^{2\kappa-5\beta+\frac{1}{2^{n+1}}-\frac{\beta}{2^{n+1}}}.\end{aligned}$$

One can choose β such that $2\kappa - 5\beta + \frac{1}{2^{n+1}} - \frac{\beta}{2^{n+1}} > 0$ and conclude that

$$\|h \circ (f_{D^2} - \mathrm{Id}) \circ h^{-1}\|_{C^{1+\beta}(U_n \setminus U_{n-1})} \to 0$$

as $n \to \infty$. Therefore, f_P is tangent to Id near ∂P and hence, the map f_P is of class $C^{1+\beta}$.

Proof of Theorem 9.3. By construction, the map f_P generates via a homeomorphism T a $C^{1+\beta}$ diffeomorphism f_M of the surface M. We construct a C^∞ diffeomorphism $\psi : D^2 \to D^2$ such that $\varphi_3 := T \circ h \circ \psi$ is the desired area preserving diffeomorphism (that is $(\varphi_3)_* m_{D^2} = m_M$) which can be continuously extended to the closure of D^2.

Denote $\mu = (h^{-1} \circ T^{-1})_* m_M$. Since both m_{D^2} and m_M are normalized Lebesgue measures, we have

$$\int_{D^2} d\, m_{D^2} = 1 = \int_M d\, m_M = \int_{D^2} d\mu.$$

To obtain the desired result it suffices to show that there is a C^∞ diffeomorphism $\psi : D^2 \to D^2$ that can be continuously extended to ∂D^2 such that $\psi_* \mu = m_{D^2}$.

Set $\mu_1 = m_{D^2}$ and for $n > 1$ define a sequence of measures μ_n such that

(i) $\mu_n \in C^\infty(D^2)$ that is the measure μ_n is absolutely continuous with respect to m_{D^2} with density function of class C^∞;
(ii) $\mu_n = \mu$ on $h^{-1}(U_{n-1})$;
(iii) $\int_{h^{-1}(U_n)} d\mu_n = \int_{h^{-1}(U_n)} d\mu$.

It is clear that for any $n \geq 1$, $\int_{D^2} d\mu_n = \int_{D^2} d\mu = 1$.

We need the following version of Moser's theorem (see [9], Lemma 1).

Lemma 9.6. *Let ω and μ be two volume forms on an oriented manifold M and let K be a connected compact set such that the support of $\omega - \mu$ is contained in the interior of K and $\int_K d\omega = \int_K d\mu$. Then there is a C^∞ diffeomorphism $\hat{\psi} : M \to M$ such that $\hat{\psi}|(M \setminus K) = Id|(M \setminus K)$ and $\hat{\psi}_* \omega = \mu$.*

Applying Lemma 9.6 to each compact sets $K_n = h^{-1}(\overline{U}_n \setminus U_n)$ and volume forms $\mu_{n+1}|(\overline{U}_n \setminus U_n)$ and $\mu_n|(\overline{U}_n \setminus U_n)$, we obtain a C^∞ diffeomorphism $\hat{\psi}_n : D^2 \to D^2$ such that $(\hat{\psi}_n)_* \mu_{n+1} = \mu_n$ and $\hat{\psi}_n | h^{-1}(U_{n-1}) =$ Id. Then we let

$$\psi_n = \hat{\psi}_n \circ \cdots \circ \hat{\psi}_1 \quad \text{and} \quad \psi = \lim_{n \to \infty} \psi_n.$$

The construction gives $\hat{\psi}_n(h^{-1}(U_n \setminus U_n)) = h^{-1}(U_n \setminus U_n)$. Recalling that $r(n)$ satisfies (30) and using (32), (33), and (34), we find that

$\operatorname{diam} \hat{\psi}_n^{-1}(U_n) \le C d_n$ where $C > 0$ is a constant and d_n is a decreasing sequence of numbers such that $\sum_{n=1}^{\infty} d_n < \infty$. This implies that $d(x, \hat{\psi}_n(x)) \le C d_n$ for any $x \in D^2$. It follows that for any $x \in D^2$ and $n > j > 0$,

$$\begin{aligned} d(\psi_j(x), \psi_n(x)) &\le \sum_{i=j}^{n-1} d(\psi_i(x), \psi_{i+1}(x)) \\ &\le \sum_{i=j}^{n-1} d\big(\psi_i(x), \hat{\psi}_i(\psi_i(x))\big) \le C \sum_{i=j}^{n-1} d_i. \end{aligned}$$

This implies that the sequence ψ_n is uniformly Cauchy and hence, ψ is well defined and continuous on D^2. We can also get that $\psi : D^2 \to D^2$ is a C^∞ diffeomorphism.

By construction, we know that $(\psi_n)_* \mu_{n+1} = \mu_1 = m_{D^2}$. Note that $D^2 = \cup_{n \ge 1} h^{-1}(U_n)$. Hence, for any $x \in D^2$ there is $n > 0$ and a neighborhood of x on which $\mu_{n+i} = \mu_n$ for any $i > 0$. It follows that $\psi_* \mu = (\psi_n)_* \mu_n = m_{D^2}$ on the neighborhood and hence, $\psi_* \mu = m_{D^2}$ on D^2. ☐

10. Completion of the proof of Theorem 3.1

10.1. **Representing the map f_M as a Young diffeomorphism.** Consider a smooth compact connected oriented surface M of genius $g \ge 0$ and the diffeomorphism $f_M : M \to M$ given by Statement 4 of Theorem 9.3. In this section we represent the map f_M as a Young diffeomorphism.

Proposition 10.1. *The map f_M is a Young diffeomorphism. More precisely, one can choose the number r_2 in* (1) *so small that the collection of s-subsets satisfies Conditions (Y1)-(Y6).*

Proof. First note that we already know that the map f_{T^2} is a Young diffeomorphism, so we can assume that the genius $g \ge 1$. Consider the collection of s-subsets Λ_i^s and the return time $\tau : \Lambda \to \mathbb{N}$ for the map $f_{\mathbb{T}^2}$ defined in Section 5.3. Define $\Delta_i^s := \varphi_3(\varphi_2(\varphi_1(\Lambda_i^s)))$ with the return time on M (again denoted by τ) given by $\tau(\varphi_3(\varphi_2(\varphi_1(x))) = \tau(x)$, $x \in \mathbb{T}^2$. Let $\Delta = \bigcup_i \Delta_i^s$. We claim that f_M is a Young diffeomorphism with respect to the collection of s-subsets Δ_i^s.

To prove this we need to check Conditions (Y1)–(Y6). Since the maps φ_i, $i = 1, 2, 3$ are homeomorphisms and (Y1) and (Y2) are satisfied for the map $f_{\mathbb{T}^2}$, then these conditions are also satisfied for f_M. In addition, (Y5) and (Y6) hold true for f_M since the maps φ_i, $i = 1, 2, 3$ preserve the area.

To show (Y3) and (Y4) observe that the element of the Markov partition P in the Young tower representation for the map $f_{\mathbb{T}^2}$ is away from the critical points $x_i,\ i = 1, 2, 3, 4$. This implies that the map $\varphi_3 \circ \varphi_2 \circ \varphi_1$ is a smooth diffeomorphism from P onto its image. Since the return time function $\tilde{f}_M = f_M^\tau$ is defined on $\Delta \subset P$, (Y4) follows as it hold true for the map $f_{\mathbb{T}^2}$ by Proposition 5.2. Note that (Y3) holds for the map $f_{\mathbb{T}^2}$ for some constant $0 < a < 1$. If the number r_2 in (1) is chosen sufficiently small, then the fact that (Y3) holds for the map f_M can be shown by applying the argument in the proof of Proposition 6.2. in [21].

Finally, arguing similarly it is easy to show that the diffeomorphism f_{S^2} of the sphere S^2 is a Young diffeomorphism. This completes the proof of the proposition. □

10.2. **Lower and upper polynomial bounds on the decay of correlations.** Let f be a Young diffeomorphism admitting a Young tower with base Λ, s-sets Λ_i^s, and inducing time $\tau = \{\tau_i\}$. Consider the associated Young tower given by (see Section 5.4)

$$\hat{Y} = \{(x, k) \in \Lambda \times \mathbb{N} : 0 \le k < \tau(x)\}.$$

For $k > 0$ let

$$\hat{M}_k = \{(x, \ell) \in \hat{Y} : 0 \le \ell \le \min\{k, \tau(x)\}\}.$$

Consider the projection $\pi : \hat{Y} \to M$ given by $\pi(x, \ell) = f^\ell(x)$ and let

$$Y_M = \pi(\hat{Y}), \quad M_k = \pi(\hat{M}_k).$$

The sets M_k are clearly nested and exhaust Y.

To establish upper and lower bounds on the decay of correlations we need the following result which is a corollary of results in [23] (see also [8] and [22]).

Proposition 10.2. *Assume that*

- *the greatest common divisor of numbers $\{\tau_i\}$, $gcd\{\tau_i\} = 1$, where τ_i are the values of the function τ;*
- *there is $C > 0$ such that for all $x, y \in \Delta_i^s$ and $0 \le j \le \tau_i$,*

$$d(f^j(x), f^j(y)) \le C \max\{d(x, y), d(f^{\tau_i}(x), f^{\tau_i}(y))\}; \tag{35}$$

- *there are $\nu > 0$ and $C_1 > 0$ such that*

$$m(\tau > n) \le \frac{C_1}{n^\nu}. \tag{36}$$

Then the following statements hold:

1. *There is $C_2 > 0$ such that $Cor_n(h_1, h_2) \le \frac{C_2}{n^{\nu-1}}$ for any $h_1, h_2 \in C^\rho(M)$.*

(2) *For any $h_1, h_2 \in C^\rho(M)$ supported in M_k for some $k > 0$, we have*

(37)

$$Cor_n(h_1, h_2) = \sum_{k=n+1}^{\infty} m(\{x\colon \tau(x) > k\}) \int_M h_1 dm \int_M h_2 dm + r_\nu(n),$$

where $r_\nu(n) = O(R_\nu(n))$ and

$$R_\nu(n) = \begin{cases} \frac{1}{n^\nu} & \text{if } \nu > 2, \\ \frac{\log n}{n^2} & \text{if } \nu = 2, \\ \frac{1}{n^{2\nu-2}} & \text{if } 1 < \nu < 2. \end{cases}$$

Moreover, if $\int_M h_1 \int_M h_2 = 0$, then $Cor_n(h_1, h_2) = O(1/n^\nu)$.

To apply Proposition 10.2 to the Young diffeomorphism f_M we need to verify the assumptions of this proposition.

To prove the first assumption that $\gcd\{\tau_i\} = 1$ observe that the maps φ_i, $i = 1, 2, 3$ and the map H are homeomorphisms. Hence, it suffices to prove this for the linear map A. This is well known (see for example, [24]).

To prove the second assumption that the map f_M satisfies (35) observe that this is true for the map $f_{\mathbb{T}^2}^{\tau_i}$ of the torus which is smoothly conjugate to the map $f_M^{\tau_i}$.

Finally, to prove the third assumption observe that by Lemmas 7.2 and 8.3, we have that

$$\frac{C_8}{n^{\gamma-1}} < m(\{x \in \Delta\colon \tau(x) > n\}) < \frac{C_{11}}{n^{\gamma'-1}}, \tag{38}$$

where γ and γ' are defined by (8). It is easy to see that $\gamma > \gamma' > 2$ for all $0 < \alpha < \frac{1}{4}$ and $0 < \mu < \frac{1}{2}$.

Since the homeomorphisms φ_1, φ_2, and φ_3 are measure preserving we also have the same estimates for the map f_M^τ. In particular, (36) holds with $\nu = \gamma' - 1$ and $C_1 = C_{11}$.

The upper bound on correlations is now an immediate corollary of Statement 1 of Proposition 10.2 where $\nu = \gamma' - 1$. In particular, we have $\gamma_1 = \gamma' - 2 > 0$.

To obtain the lower bound on correlations we apply Statement 2 of Proposition 10.2 with $\nu = \gamma' - 1$ and obtain for all $h_1, h_2 \in C^\rho(M)$ supported in M_k for some $k > 0$ that

(39)

$$\mathrm{Cor}_n(h_1, h_2) = \sum_{k=n+1}^{\infty} m(\{x\colon \tau(x) > k\}) \int_M h_1 dm \int_M h_2 dm + r_{\gamma'}(n),$$

where $r_{\gamma'}(n) = O(R_{\gamma'}(n))$ and

$$R_{\gamma'}(n) = \begin{cases} \frac{1}{n^{\gamma'-1}} & \text{if } \gamma' > 3, \\ \frac{\log n}{n^2} & \text{if } \gamma' = 3, \\ \frac{1}{n^{2\gamma'-4}} & \text{if } 2 < \gamma' < 3. \end{cases}$$

Consider the two cases $\gamma' \geq 3$ and $2 < \gamma' < 3$.

Assume first that $\gamma' > 3$, which is true if $\alpha < \frac{1}{6}$. By assumption, $\int_M h_1\, dm \int_M h_2\, dm > 0$ and hence, applying (38) and (39), we obtain

$$\mathrm{Cor}_n(h_1, h_2) > \frac{K_1}{n^{\gamma-2}} - \frac{K_2}{n^{\gamma'-1}},$$

where $K_1 > 0$ and $K_2 > 0$ are constants. Using definitions of γ and γ' (see (8)) and choosing any $0 < \mu < \frac{1}{2}$, one can show that $\gamma - 2 < \gamma' - 1$ for all $0 < \alpha < \frac{1}{6}$.[4] We conclude that for some $C > 0$,

$$\mathrm{Cor}_n(h_1, h_2) > \frac{C}{n^{\gamma-2}}.$$

Now, we consider the case when $\frac{1}{6} < \alpha < \frac{1}{4}$. This implies that $\gamma' > 2$. Depending on the value of μ, we may either have $\gamma' > 3$ or $\gamma' < 3$ and we assume the latter (otherwise we are back to the previous case). With this assumption we have

$$\mathrm{Cor}_n(h_1, h_2) > \frac{K_1}{n^{\gamma-2}} - \frac{K_3}{n^{2\gamma'-4}},$$

where $K_3 > 0$ is a constant. Choosing again $0 < \mu < \frac{1}{2}$, one can show that $\gamma - 2 < 2\gamma' - 4$ holds for all $0 < \alpha < \frac{1}{4}$.[5] Thus we have the desired estimate

$$\mathrm{Cor}_n(h_1, h_2) > \frac{C}{n^{\gamma-2}}$$

for some $C > 0$ and all $0 < \alpha < \frac{1}{4}$. In particular, we have $\gamma_2 = \gamma - 2 > 0$.

10.3. **The Central Limit Theorem.** By Statement 1 of Theorem 3.1, for any Hölder continuous function h satisfying $\int h dm = 0$ we have $\mathrm{Cor}_n(h, h) = O(\frac{1}{n^{\gamma'-1}})$. This implies that the correlation function is summable, when $\gamma' > 2$ that is when $0 < \alpha < \frac{1}{4}$. The desired result now follows from [14], Theorem 4.1 (see also [23], Theorem 3.1).

[4] One can use a computer assisted calculation to show that $\gamma - 2 < \gamma' - 1$ for all $0 < \alpha < 0.42\ldots$.

[5] Again a computer assisted calculation to show that $\gamma - 2 < 2\gamma' - 4$ holds for all $0 < \alpha < 0.36\ldots$.

10.4. **The Large Deviation property.** We consider the Young tower Y that represents the map f_M. The upper bound in (38) allows us to use Theorem 4.2 in [17] to obtain for $0 < \alpha < \frac{1}{4}$ that for all sufficiently small $a > 0$

$$m_M\Big(\Big|\frac{1}{n}\sum_{i=0}^{n-1} h(f_M^i(x)) - \int h\Big| > \varepsilon\Big) < C_{h,a}\varepsilon^{-2(\gamma'-2-a)}n^{-(\gamma'-2-a)}.$$

Moreover, for each such an $a > 0$ the constant $C_{h,a}$ depends on the Hölder norm of h continuously.

To get a lower bound we need to check the conditions of Theorem 4.3 in [17]. More precisely, for the set $\hat{Y}_k = \{(x,l) \in \hat{Y} : \tau(x) > k\}$ it must be true that for some k, $m_M(\pi(\hat{Y}_k)) < 1$ where $\pi : \hat{Y} \to M$ is given by $\pi(x,k) = f^k(x)$ as before.

Given k let us chose a partition element Δ_i in the base Δ of the tower for f_M with $\tau(\Delta_i) \le k$. Then $\Delta_i \subset \hat{Y} \setminus \hat{Y}_k$ and obviously $\hat{m}_M(\Delta_i) > 0$. Thus $\hat{m}_M(\hat{Y}_k) < 1$ and since π is measure preserving, we obtain $m_M(\pi(\hat{Y}_k)) < 1$. Thus, by Theorem 4.3 in [17], we obtain the lower bound

$$\frac{1}{n^{\gamma'-2+a}} < m_M\Big(\Big|\frac{1}{n}\sum_{i=0}^{n-1} h(f_M^i(x)) - \int h\Big| > \varepsilon\Big)$$

for small ε, open and dense subset of Hölder continuous observables h, and infinitely many n.

10.5. **The measure of maximal entropy (MME).** Recall that the diffeomorphism f_M of the surface M is a Young diffeomorphism and consider the corresponding collection $\{\Delta_i^s\}$ of s-sets. Denote by $\mathcal{S}_n = \{\Delta_i^s : \tau(\Delta_i^s) = n\}$. Since the map f_M is topologically conjugate to the toral automorphism A, the number $\mathcal{S}_n$ for f_M is equal to the number $\mathcal{S}_n$ for A. The latter is known to satisfy $\mathcal{S}_n \le e^{hn}$ with $h < h_{\text{top}}(A)$ (see [21]). It now follows from [20] (see Theorem 7.1) and [23] that the map f_M possesses a unique MME which has all the desired properties.

References

[1] D. Anosov and A. Katok, *New examples in smooth ergodic theory. Ergodic diffeomorphisms*, Trans. Moscow. Math. Soc., **23** (1970) 1–35.

[2] L. Barreira and Ya. Pesin, *Introduction to Smooth Ergodic Theory*, **148**, Graduate Studies in Mathematics, AM, Providence, RI, 2013.

[3] M. Brin, *Bernoulli diffeomorphims with nonzero exponents*, ETDS, **1** (1981), 453–465.

[4] M. Brin, J. Feldman, and A. Katok,*Bernoulli diffeomorphisms and group extensions of dynamical systems with non-zero characteristic exponents*, Ann. Math., **113:1** (1981) 159–179.

[5] H. Bruin and D. Terhesiu, *Regular variation and rates of mixing for infinite measure preserving almost Anosov diffeomorphisms*, ETDS, **40** (2020) 663–698.

[6] D. Dolgopyat and Ya. Pesin, *Every compact manifold carries a completely hyperbolic diffeomorphism*, ETDS **22** (2002), 409–435.

[7] P. Eslami and C. Liverani, *Mixing rates for symplectic almost Anosov maps*, Nonlinearity, **34** (2021) 3709–3731.

[8] S. Gouëzel, *Sharp polynomial estimates for the decay of correlations*, Israel J. Math, **139** (2004) 29–65.

[9] R. E. Greene and K. Shiohama, *Diffeomorphisms and volume-preserving embeddings of noncompact manifolds*, Trans. Amer. Math. Soc., **255** (1979) 403–414, .

[10] H. Hu, *Decay of correlations for piecewise smooth maps with indifferent fixed point*, ETDS, **24** (2004) 495–524.

[11] H. Hu and X. Zhang, *Polynomial decay of correlation for almost Anosov maps*, ETDS, 2017.

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

[13] A. Katok and B. Hasselblatt. *Introduction to the theory modern theory of dynamical systems*, **54**, Cambridge Univ. Press, 1995.

[14] C. Liverani, *Central limit theorem for deterministic systems*, International Conference on Dynamical Systems (Montevideo, 1995), **362**, Pitman Res. Notes Math. Ser., Longman-Harlow (1996) 56–75.

[15] C. Liverani, B. Saussol, and S. Vaienti. *A probabilistic approach to intermittency*, ETDS **19** (1999) 671–685.

[16] C. Liverani and M.Martens, *Convergence to equilibrium for intermittent symplectic maps*, Comm. Math. Phys., **260(3)**,(2005) 527–556.

[17] I. Melbourne and M. Nicol, *Large Deviations for nonuniformly hyperbolic systems*, Trans. AMS **360** (2008) 6661–6676.

[18] I. Melbourne and D. Terhesiu, *Decay of correlations for non-uniformly expanding systems with general return times*, ETDS, **34** (2014) 893–918.

[19] Yves Pomeau and Paul Manneville, *Intermittent transition to turbulence in dissipative dynamical systems*, Comm. Math. Phys., **74:2** (1980) 189–197.

[20] Y. Pesin, S. Senti, and K. Zhang, *Thermodynamics of towers of hyperbolic type*, Trans. Amer. Math. Soc., **368:12** (2016) 8519–8552.

[21] Y. Pesin, S. Senti, and K. Zhang, *Thermodynamics of the Katok map*, ETDS, **39:3** (2019) 764–794.

[22] O. Sarig, *Subexponential decay of correlations*, Inventiones Mathematicae, **150** (2002) 629–653.

[23] F. Shahidi and A. Zelerowicz, *Thermodynamics via inducing*, J. Stat. Phys. **175:2** (2019) 351–383.

[24] L.-S. Young, *Statistical properties of dynamical systems with some hyperbolicity*, Ann. Math., **147:2** (1998) 585-650.

[25] L.-S. Young, *Recurrence times and rates of mixing*, Israel J. Math., **110** (1999) 153–188.

Department of Mathematics, Pennsylvania State University, University Park, PA 16802, USA
Email address: pesin@math.psu.edu

Instituto de Matematica, Universidade Federal do Rio de Janeiro, C.P. 68 530, CEP 21945-970, R.J., Brazil
Email address: senti@im.ufrj.br

Department of Mathematics, Pennsylvania State University, University Park, PA, 16802, USA
Email address: fus144@psu.edu