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 $\mu$ 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 $C^\infty$ map with some additional metric properties. Katok [12] gave an example of area preserving $C^\infty$ diffeomorphism with non-zero Lyapunov exponents on any surface which is Bernoulli (see the definitions in the next section). Later Brin, Feldman, and Katok [4] and then Brin [3] extended this result by constructing a volume preserving $C^\infty$ diffeomorphism, which is Bernoulli,
on any Riemannian manifold of dimension $\geq 5$. In this example the map has all but one non-zero Lyapunov exponents. Finally, Dolgopyat and Pesin [9] constructed a volume preserving $C^\infty$ Bernoulli diffeomorphism with non-zero Lyapunov exponents on any Riemannian manifold of dimension $\geq 2$.

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 H"older 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 [7, 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, 6, 16, 15]. Also, starting from the work of Young, [21, 25], some general techniques for obtaining polynomial decay of correlations for maps admitting Young tower have been developed by Gou"ezel [7] 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.
We refer the reader to Remark 4 where we discuss upper bounds on the number $\beta$.

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 $C^\infty$), to ensure that $f_{T^2}$ has polynomial decay of correlations, we have to choose the slow down function to be polynomial at 0. This results in $f_{T^2}$ to be of class $C^{2+\kappa}$. 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, [7, 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 $C^{2+\kappa}$ 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 $C^\infty$ and is infinitely and arbitrarily flat at the boundary of the disk, 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 homeomorphism from the closed disk onto $M$ and is area preserving.

---

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.
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] which allow us to choose a 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 [7]) 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 [7, 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 result in Section 3. In Section 4 we construct the map 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 a diffeomorphism of a given surface. Finally, in section 10 we complete the proof the main result.

Acknowledgments. 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 CNP grant.
2. Definitions and notations

Let $X$ be a measurable space and $T : X \to X$ a measurable invertible transformation preserving a measure $\mu$. 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 $(T, \mu)$ has the Bernoulli property if it is metrically isomorphic to the Bernoulli shift $(\sigma, \kappa)$ associated to some Lebesgue space $(Y, \nu)$, so that $\nu$ is metrically isomorphic to the Lebesgue measure on an interval together with at most countably many atoms and $\kappa$ is given as the direct product of $Z$ copies of $\nu$ on $Y^Z$.

2.2. Decay of correlations. Let $H_1$ and $H_2$ be two classes of real-valued functions on $X$ called observables. For $h_1 \in H_1$ and $h_2 \in H_2$ define the correlation function
\[
\text{Cor}_n(h_1, h_2) := \int h_1(T^n(x))h_2(x) \, d\mu - \int h_1(x) \, d\mu \int h_2(x) \, d\mu.
\]
We say that $T$ has polynomial decay of correlations (more precisely, polynomial upper bound on correlations) with respect to classes $H_1$ and $H_2$ if there exists $\gamma_1 > 0$ such that for any $h_1 \in H_1$, $h_2 \in H_2$, and any $n > 0$,
\[
|\text{Cor}_n(h_1, h_2)| \leq Cn^{-\gamma_1},
\]
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 $H_1$ and $H_2$ of observables if there exists $\gamma_2 > 0$ such that for any $h_1 \in H_1$, $h_2 \in H_2$, and any $n > 0$,
\[
|\text{Cor}_n(h_1, h_2)| \geq C'n^{-\gamma_2},
\]
where $C' = C'(h_1, h_2) > 0$ is a constant.

We say that $T$ has exponential decay of correlations with respect to classes $H_1$ and $H_2$ if there exists $\gamma_3 > 0$ such that for any $h_1 \in H_1$, $h_2 \in H_2$, and any $n > 0$,
\[
|\text{Cor}_n(h_1, h_2)| \leq C''e^{-\gamma_3n},
\]
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 \textit{Polynomial Large Deviation} with respect to a class $\mathcal{H}$ of observables on $X$ if there is $\hat{\beta} > 0$ such that for any $h \in \mathcal{H}$, any $\varepsilon > 0$, and any sufficiently large $n > 0$
\[
\mu \left( \left| \frac{1}{n} \sum_{i=0}^{n-1} h(T^i(x)) - \int h \right| > \varepsilon \right) < Kn^{-\hat{\beta}},
\]
where $K = K(\varepsilon, \hat{\beta}, h) > 1$ is a constant.

2.5. Lyapunov exponents. Let $f : M \to M$ be a diffeomorphism of a compact smooth Riemannian manifold $M$. Given a point $x \in M$ and a vector $v \in T_x M$, the number
\[
\chi(x, v) := \limsup_{n \to \infty} \frac{1}{n} \log \| df^n_x v \|
\]
is called the \textit{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 $\chi_1(x) \leq \cdots \leq \chi_p(x)$, where $p = \dim M$. The functions $\chi_i(x)$, $i = 1, \ldots, p$ are Borel measurable and $f$-invariant.

If $\mu$ is an $f$-invariant measure, then for $\mu$-almost every $x \in M$ and any $v \in T_x M$,
\[
\chi(x, v) = \lim_{n \to \infty} \frac{1}{n} \log \| df^n_x v \|.
\]
We say that $f$ has \textit{nonzero Lyapunov exponents} with respect to $\mu$ or that $\mu$ is \textit{hyperbolic} if for $\mu$-almost every $x$ we have that $\chi_i(x) \neq 0$, $i = 1, \ldots, p$ and that $\chi_1(x) < 0$ while $\chi_p(x) > 0$.

Note that if $\mu$ is ergodic, then $\chi_i(x)$ is a constant for all $i = 1, \ldots, p$ and $\mu$-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 $C^\rho(M)$ be the class of all Hölder continuous functions (observables) on $M$ with exponent $\rho$.

Consider a nested sequence of subsets $\{M_j\}$ that exhausts $M$ that is $M_1 \subset M_2 \subset \ldots \subset M$ and $\bigcup_{j \geq 1} M_j = M$. Given such a sequence, let $\mathcal{G} = \mathcal{G}(\{M_j\})$ be the class of observables $h \subset C^\rho(M)$ for which there
is \( k = k(h) \) such that \( \text{supp}(h) \subset \overline{M_k} \) where \( \overline{M_k} \) is the closure of the set \( M_k \).

**Theorem 3.1.** Let \( M \) be a compact smooth connected and oriented surface. There are numbers \( \beta > 0, \hat{\beta} > 0, \rho > 0, \gamma_2 > \gamma_1 > 0 \), 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 \( h_i \in C^\rho(M), i = 1, 2 \)
   \[ |\text{Cor}_n(h_1, h_2)| \leq C_1 n^{-\gamma_1}, \]
   where \( C_1 = C_1(\|h_1\|_{C^\rho}, \|h_2\|_{C^\rho}) > 0; \)
   b. there is a nested sequence of subsets \( \{M_j\} \) which exhausts \( M \) such that for any \( h_i \in \mathcal{G}(\{M_j\}), i = 1, 2 \) we have
      i. if \( \int h_1 dm \int h_2 dm > 0 \), then
      \[ |\text{Cor}_n(h_1, h_2)| \geq C_2 n^{-\gamma_2}, \]
      where \( C_2 = C_2(\|h_1\|_{C^\rho}, \|h_2\|_{C^\rho}) > 0; \)
      ii. if \( \int h_1 dm \int h_2 dm = 0 \), then
      \[ C^{-1} n^{-\gamma_1} \leq |\text{Cor}_n(h_1, h_2)| \leq C n^{-\gamma_1} \]
      for some \( C > 0; \)
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 \( C^\rho \) of observables with the constant \( K \) of the form \( K = K(\|h\|_{C^\rho}) \varepsilon^{-2\hat{\beta}} \). In addition, for an open and dense subset of observables in \( C^\rho \) and any sufficiently small \( \varepsilon > 0 \)
   \[ n^{-\beta} < m\left( \left| \frac{1}{n} \sum_{i=0}^{n-1} h(f^i(x)) - \int h > \varepsilon \right| \right) \]
   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 $C^p$ of observables.

We will show (see Remark 4) that one can choose the number $\beta$ in this theorem to be any number in the interval $(0, 0.8)$.

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 $A := \begin{pmatrix} 5 & 8 \\ 8 & 13 \end{pmatrix}$. 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 $D^i_r = \{(s_1, s_2) : s_1^2 + s_2^2 \leq r^2\}$ 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

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

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

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

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

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

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

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

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

defines a homeomorphism of the torus $\mathbb{T}^2$, which is a $C^\infty$ 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

$$d\nu = q_0^{-1}q \, dm,$$

where $m$ is the area and the density $q$ is a positive $C^\infty$ function that is infinite at $x_i$ and is defined by

$$q(s_1, s_2) := \begin{cases} (\psi(s_1^2 + s_2^2))^{-1} & \text{if } (s_1, s_2) \in D_{r_1}^i, \\ 1 & \text{in } \mathbb{T}^2 \setminus D_{r_1} \end{cases}$$

and

$$q_0 := \int_{\mathbb{T}^2} q \, dm.$$ 

We further perturb the map $G$ by a coordinate change $\phi$ in $\mathbb{T}^2$ to obtain an area preserving map. To achieve this, define a map $\phi$ in $D_{r_1}^i$ by the formula

$$\phi(s_1, s_2) := \frac{1}{\sqrt{q_0(s_1^2 + s_2^2)}} \left( \int_0^{s_1^2 + s_2^2} \frac{du}{\psi(u)} \right)^{1/2} (s_1, s_2)$$

and set $\phi = \text{Id}$ in $\mathbb{T}^2 \setminus D_{r_1}$. Clearly, $\phi$ is a homeomorphism and is a $C^\infty$ diffeomorphism outside the points $x_1, x_2, x_3, x_4$. One can show that $\phi$ transfers the measure $\nu$ into the area and that the map $f_{\mathbb{T}^2} = \phi \circ G \circ \phi^{-1}$ is a homeomorphism and is a $C^\infty$ 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

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

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

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

**Proposition 4.2.** The map $f_{T^2}$ is of class of smoothness $C^{2+2\kappa}$, where $\kappa = \frac{\alpha_1 - \alpha}{1 - \alpha}$. 

*Proof.* Fix $i \in \{1, 2, 3, 4\}$ and consider the vector field in $D^r_{r_1}$ 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 $\nu$ (see (5)) and the same Hamiltonian function. The map $f_{T^2}$ is conjugate via $\varphi$ (see (6)) to the time-1 map of the flow generated by (3). Since $\phi^* \nu = m$, $f_{T^2}$ 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 $D^r_{r_1}$ 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 $f_{T^2}$ is of the desired class of smoothness we will show that the Hamiltonian $H_2$ has H"older continuous partial derivatives of second order with H"older exponent $2\kappa$. 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 H"older continuous partial derivatives of second order with H"older exponent $2\kappa$. Note that $g$ is of class $C^\infty$ except for $(x, y) = (0, 0)$, so we only need to show H"older 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 H"older continuous. Since $\frac{\partial g}{\partial x}(0, 0) = \lim_{\Delta x \to 0} \frac{g(\Delta x, 0) - g(0, 0)}{\Delta x} = 0$, 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,
\[
\frac{\partial^2 g}{\partial x \partial y} = \begin{cases} 
(1 + 2\kappa)(x^2 + y^2)^{\kappa} + 4(\kappa - 1)x^2 y(x^2 + y^2)^{\kappa-2} & (x, y) \neq (0, 0); \\
0, & (x, y) = (0, 0).
\end{cases}
\]

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
\[
|\frac{\partial^2 g}{\partial x \partial y}(x, y) - \frac{\partial^2 g}{\partial x \partial y}(0, 0)| \leq K(x^2 + y^2)^\kappa = Kd((x, y), (0, 0))^{2\kappa},
\]
where \(K > 0\) and \(d\) denotes the usual distance. Thus \(\frac{\partial^2 g}{\partial x \partial y}\) is Hölder continuous with Hölder exponent \(2\kappa\).

Now we consider \(\frac{\partial^2 g}{\partial x^2}\). Observe that \(\frac{\partial^2 g}{\partial x^2}(0, 0) = 0\) and
\[
\frac{\partial^2 g}{\partial x^2} = \begin{cases} 
6\kappa xy(x^2 + y^2)^{\kappa-1} + 4(\kappa - 1)x^3 y(x^2 + y^2)^{\kappa-2} & (x, y) \neq (0, 0); \\
0, & (x, y) = (0, 0).
\end{cases}
\]
It is easy to see that
\[
|\frac{\partial^2 g}{\partial x^2}(x, y) - \frac{\partial^2 g}{\partial x^2}(0, 0)| \leq K(x^2 + y^2)^\kappa = Kd((x, y), (0, 0))^{2\kappa}.
\]
Hence, \(\frac{\partial^2 g}{\partial x^2}\) is Hölder continuous with Hölder exponent \(2\kappa\).

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 \(f\).

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 \(d(f^{-n}(x), f^{-n}(y)) \to \)

\(^{3}\)Our requirements are slightly different than those in [24] since we do not assume the inducing domain \(\Lambda\) to be compact.
A continuous family of unstable disks if there exists a homeomorphism $\Phi : K^s \times D^u \to \bigcup \gamma^u$ satisfying:

- $K^s \subset M$ is a Borel subset and $D^u \subset \mathbb{R}^d$ is the closed unit disk for some $d < \dim M$;
- $x \to \Phi(x) \times D^u$ 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 $\gamma^u$ and a continuous family $\Gamma^s = \{\gamma^s\}$ of stable disks $\gamma^s$ such that

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

A subset $\Lambda_0 \subset \Lambda$ is called an $s$-subset if it has hyperbolic product structure and is defined by the same family $\Gamma^u$ of unstable disks as $\Lambda$ and a continuous subfamily $\Gamma^s_0 \subset \Gamma^s$ of stable disks. A $u$-subset is defined analogously.

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

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

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

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

Assume the map $f$ satisfies the following conditions:

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

$$\Lambda^s_i := \bigcup_{\gamma \in \Gamma^s_i} (\gamma \cap \Lambda) \subset \Lambda$$

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 \( \Lambda \) such that for all \( x \in \Lambda_i^s \)
\[
f^{-\tau_i}(\gamma_s(f^{\tau_i}(x)) \cap \Lambda_i^u) = \gamma_s(x) \cap \Lambda,
\]
\[
f^{\tau_i}(\gamma_u(x) \cap \Lambda_i^u) = \gamma_u(f^{\tau_i}(x)) \cap \Lambda.
\]

(Y2) The sets \( \Lambda_i^u \) 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 \( y \in \gamma_s(x) \),
\[
d(\tilde{f}(x), \tilde{f}(y)) \leq a d(x, y);
\]
(b) For \( x \in \Lambda_i^s \) and \( y \in \gamma_u(x) \cap \Lambda_i^s \),
\[
d(x, y) \leq a d(\tilde{f}(x), \tilde{f}(y)).
\]

For \( x \in \Lambda \) let \( Jac_f(x) = \det|Df|_{E^u(x)}| \) 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 \( n \geq 0 \), \( x \in \tilde{f}^{-n}(\bigcup_{i \in \mathbb{N}} \Lambda_i^s) \) and \( y \in \gamma_s(x) \) we have
\[
\left| \log \frac{Jac\tilde{f}(\tilde{f}^n(x))}{Jac\tilde{f}(\tilde{f}^n(y))} \right| \leq cb^n,
\]
(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 \( 0 \leq k \leq n \) and \( y \in \gamma_u(x) \) we have
\[
\left| \log \frac{Jac\tilde{f}(\tilde{f}^{n-k}(x))}{Jac\tilde{f}(\tilde{f}^{n-k}(y))} \right| \leq cb^k.
\]

(Y5) For every \( \gamma_u \in \Gamma^u \) one has
\[
\mu_{\gamma_u}(\gamma_u \cap \Lambda) > 0, \quad \mu_{\gamma_u} \left( (\Lambda \setminus \bigcup \Lambda_i^s) \cap \gamma_u \right) = 0,
\]
where \( \mu_{\gamma_u} \) is the leaf volume on \( \gamma_u \).
(Y6) There exists $\gamma^u \in \Gamma^u$ such that
\[
\sum_{i=1}^{\infty} \tau_i \mu_{\gamma^u}(\Lambda^s_i \cap \gamma^u) < \infty.
\]

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, $\tilde{\mathcal{P}} = \text{int} \mathcal{P}$ for any $\tilde{\mathcal{P}} \in \hat{\mathcal{P}}$. Let $\tilde{\mathcal{P}} \in \hat{\mathcal{P}}$ be a partition element which does not intersect any of the disks $D_{r_0}$, $i = 1, 2, 3, 4$. Given $\delta > 0$, we can always choose the Markov partition $\tilde{\mathcal{P}}$ in such a way that $\text{diam}(\tilde{\mathcal{P}}) < \delta$. For a point $x \in \tilde{\mathcal{P}}$ denote by $\tilde{\gamma}^s(x)$ (respectively, $\tilde{\gamma}^u(x)$) the connected component of the intersection of $\tilde{\mathcal{P}}$ 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 $x \in \tilde{\mathcal{P}}$, let $\tilde{\tau}(x)$ be the first return time of $x$ to $\text{int}\tilde{\mathcal{P}}$. 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 $\tilde{\mathcal{P}}$. 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 $y \in \tilde{\mathcal{P}}$ 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 $x, y \in \tilde{\mathcal{P}}$ 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 $f_{T^2}$.** 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 $\varepsilon$, there is $\delta > 0$ such that $\text{diam} (P) < \varepsilon$ for any $P \in \mathcal{P}$ provided $\text{diam} (\tilde{P}) < \delta$. 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. $\Lambda_s(x) = H(\tilde{\Lambda}_s)$ are $s$-sets;
2. $\Lambda_u(x) = H(\tilde{\Lambda}_u) = f_{T^2}^1(\Lambda_s)$ are $u$-sets;
3. $\tau_i = \tilde{\tau}_i$ – the inducing times – are the first return time of points in $\Lambda_s$ to $\Lambda$;
4. $\hat{f}(x) = f_{T^2}^{\tau(x)}(x)$ 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 $f_{T^2}$ such that

1. there is a partition element $P$ for which $f_{T^2}^j(x) \notin D_r$ for any $0 \leq j \leq Q$ and for any point $x$ for which either $x \in \Lambda$ or $x \notin f_{T^2}(D_r)$ while $f_{T^2}^{-1}(x) \in D_r$ 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 $x_i \in D_r \subset \text{Int} P_i$.

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 $\Lambda^s$ satisfies Conditions (Y1)-(Y6).

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

$$\hat{\mathcal{Y}} = \{(x, k) \in \Lambda \times \mathbb{N} : 0 \leq k < \tau(x)\}.$$
and the tower map \( \hat{f}_{T^2} : \hat{Y} \to \hat{Y} \) by \( \hat{f}_{T^2}(x, k) = (x, k + 1) \) if \( k < \tau(x) - 1 \) and \( \hat{f}_{T^2}(x, k) = (\hat{f}(x), 0) \) if \( k = \tau(x) - 1 \) where \( \hat{f} : \Lambda \to \Lambda \) is the induced map. The map \( \hat{f}_{T^2} \) 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 \( \hat{f}_{T^2} \) is a measurable bijection from \( \Lambda_s^\alpha \times \{k - 1\} \) to \( \Lambda_s^\alpha \times \{k\} \) for all \( 1 \leq k < \tau_i - 1 \) and from \( \Lambda_s^\alpha \times \{\tau_i - 1\} \) 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 \( D^i \). We also set \( f := f_{T^2} \).

Lemma 6.1 ([21], Lemma 5.1). For \( s = (s_1, s_2) \in D^i_T \) let

\[
\begin{aligned}
d_{i,j} &= d_{i,j}(s_1, s_2) := \frac{\partial^2}{\partial s_i \partial s_j} \psi(s_1^2 + s_2^2).
\end{aligned}
\]

Then

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

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 \( f^{-1}(s(0)) \notin D^i_T \) and \( f(s(T)) \notin D^i_T \) but \( s(t) \in D^i_T \) for all \( 0 \leq t \leq T \). In particular, \( s_1(t) \neq 0 \) and \( s_2(t) \neq 0 \). Setting \( T_1 = \frac{T}{2} \) we have that \( s_1(t) \leq s_2(t) \) for all \( 0 \leq t \leq T_1 \) and \( s_1(t) \geq s_2(t) \) for all \( T_1 \leq t \leq T \). The following statement provides effective lower and upper bounds on the functions \( s_1(t) \) and \( s_2(t) \).

Lemma 6.2. The following statements hold:

\[
\begin{aligned}
|s_2(t)| &\geq |s_2(a)| \left(1 + 2^\alpha C_1 s_2^{2\alpha}(a) (t - a)\right)^{-\frac{1}{2\alpha}}, & 0 \leq a \leq t \leq T_1; \\
|s_2(t)| &\leq |s_2(a)| \left(1 + C_1 s_2^{2\alpha}(a) (t - a)\right)^{-\frac{1}{2\alpha}}, & 0 \leq a \leq t \leq T; \\
|s_1(t)| &\geq |s_1(b)| \left(1 + 2^\alpha C_1 s_1^{2\alpha}(b) (b - t)\right)^{-\frac{1}{2\alpha}}, & T_1 \leq t \leq b \leq T; \\
|s_1(t)| &\leq |s_1(b)| \left(1 + C_1 s_1^{2\alpha}(b) (b - t)\right)^{-\frac{1}{2\alpha}}, & 0 \leq t \leq b \leq T; \\
|s_1(t)| &\leq |s_1(T_1)| \left(1 - 2^\alpha C_1 s_1^{2\alpha}(T_1) (t - T_1)\right)^{-\frac{1}{2\alpha}}, & T_1 \leq t \leq T.
\end{aligned}
\]

where \( C_1 = \frac{2^\alpha \log \Lambda}{r_0^2} \) is a constant.
Proof. The first four estimated are proven in [21] (see Lemma 5.2 in the erratum to the paper) and the fifth estimate can be easily shown in the same manner as in [21]. □

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 [21], Lemma 5.3 and the erratum to the paper.

**Lemma 6.3.** Fix \( 0 < \mu < 1 \) and assume that \( s_1(t) \neq 0, s_2(t) \neq 0 \) for all \( 0 < t < T \), and that
\begin{enumerate}
  \item \( \Delta s_2(t) > 0 \) and \( |\Delta s_1(t)| \leq \mu \Delta s_2(t) \) for \( t \in [0, T] \);
  \item \( \left| \frac{\Delta s_2(T_1)}{s_2(T_1)} \right| < \lambda \mu \),
\end{enumerate}
Then
\[
\Delta s_2(t) \leq \Delta s_2(0) s_2(t) \left( 1 + 2^\alpha C_1 s_2^{2\alpha}(0) t \right)^{-\beta'}, \quad 0 \leq t \leq T_1;
\]
\[
\Delta s_2(t) \leq \Delta s_2(T_1) s_1(t) \left( \frac{1 + 2^\alpha C_1 s_2^{2\alpha}(b)(b - t)}{1 + 2^\alpha C_1 s_2^{2\alpha}(b)(b - T_1)} \right)^{\beta'}, \quad T_1 \leq t \leq b \leq T,
\]
where \( \beta' = \frac{1 - \mu}{2^\alpha + 2} \) and \( C_1 \) is the constant in Lemma 6.2. In addition,
\[
\|\Delta s(T)\| \leq \sqrt{1 + \mu^2} \left( \frac{s_1(T)}{s_2(0)} \right) \|\Delta s(0)\|.
\]

Given \( 0 < \alpha < 1 \) and \( 0 < \mu < 1 \), denote by
\[
\gamma = \frac{1}{2\alpha} + 2^{\alpha - 1}(1 + \mu) + \frac{1 - \mu}{6}, \quad \gamma' = \frac{1}{2\alpha} + \frac{1 - \mu}{2^{\alpha+2}}.
\]
It is easy to see that \( \gamma > \gamma' \) for all \( 0 < \alpha < 1 \) and \( 0 < \mu < 1 \) and that
\[
\gamma > \gamma' \quad \text{for all } 0 < \alpha < 1 \text{ and all } 0 < \mu < 1,
\]
\[
\gamma' > 2 \quad \text{for all } 0 < \alpha < \frac{1}{4} \text{ and all } 0 < \mu < 1.
\]

**Lemma 6.4.** Under the assumptions of Lemma 6.3 for any \( 0 < t < T_1 \) we have that
\[
\Delta s_2(t) < C_2 \Delta s_2(0) t^{-\gamma'},
\]
where \( C_2 > 0 \).

Proof. By Lemma 6.2 (the second estimate), one has
\[
s_2(t) \leq s_2(0) \left( 1 + C_1 s_2^{2\alpha}(0) t \right)^{-\frac{1}{\alpha}}.
\]
Therefore, Lemma 6.3 implies that
\[
\Delta s_2(t) \leq \frac{\Delta s_2(0)}{s_2(0)} s_2(t) \left( 1 + 2^\alpha C_1 s_2^{2\alpha}(0) \right)^{\beta'} \\
\leq \Delta s_2(0) \left( 1 + C_1 s_2^{2\alpha}(0) t \right)^{-\gamma'}.
\]
Since \((1 + C_1 s_2^{2\alpha}(0)) t > C_1 s_2^{2\alpha}(0) t\), we have
\[
\Delta s_2(t) \leq \Delta s_2(0) \left( 1 + C_1 s_2^{2\alpha}(0) \right)^{-\gamma'} t^{-\gamma'}
\]
and the desired estimate follows, since \(s_2(0)\) is of order \(r_0\).

\[\square\]

**Lemma 6.5.** Under the assumptions of Lemma 6.3 there is \(C_3 > 0\) such that for all \(T_1 < t < T\) we have
\[
\Delta s_2(t) < C_3 \Delta s_2(T_1).
\]

**Proof.** Let \(s(t) = (s_1(t), s_2(t))\) be the solution of (3) with \(s_1(t)s_2(t) = \kappa_0\), where \(\kappa_0 > 0\) is a constant. Then we have
\[
\frac{ds_2}{dt} = -\log \lambda s_2(s_1^2 + s_2^2)\alpha \leq -\log \lambda s_2(2\kappa_0)\alpha.
\]
By integrating the above over \([0, T_1]\), we obtain
\[
\log \frac{s_2(T_1)}{s_2(0)} \leq -\log \lambda (2\kappa_0)^\alpha T_1.
\]
It follows that
\[
1 - 2^\alpha C_1 s_1^{2\alpha}(T_1)(t - T_1) = 1 - 2^\alpha C_1 s_2^{2\alpha}(T_1)(t - T_1) \\
\geq 1 - 2^\alpha C_1 s_2^{2\alpha}(0) \exp(-2\alpha \log \lambda (2\kappa_0)^\alpha T_1)(t - T_1) \\
\geq 1 - 2^\alpha C_1 s_2^{2\alpha}(0) \exp(-2\alpha \log \lambda (2\kappa_0)^\alpha T_1)T_1 \\
\geq 1 - C_0 > 0
\]
for some \(C_0 > 0\). Therefore, the second estimate in Lemma 6.3, the fifth inequality in Lemma 6.2, and the assumption that \(T_1 < t < T\) imply
\[
\frac{\Delta s_2(t)}{\Delta s_2(T_1)} \leq \left( 1 - 2^\alpha C_1 s_1^{2\alpha}(T_1)(t - T_1) \right)^{-\frac{1}{2\alpha}} \left( \frac{1 + 2^\alpha C_1 s_1^{2\alpha}(b)(b - t)}{1 + 2^\alpha C_1 s_1^{2\alpha}(b)(b - T_1)} \right)^{\beta'} \\
\leq (1 - C_0)^{-\frac{1}{2\alpha}},
\]
which proves the desired estimate. \[\square\]
Proof. Let $s_1 = s_1(t), s_2 = s_2(t), u := s_1^2 + s_2^2$, and $\bar{u} = s_1^2 + s_2^2$. 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

$$\frac{d}{dt} \Delta s_2(t) = \frac{d}{dt} \tilde{s}_2(t) - \frac{d}{dt} s_2(t) = -(\log \lambda) (\tilde{s}_2 \psi(\bar{u}) - s_2 \psi(u))$$

(10)

$$= -\log \lambda \left( \frac{\partial}{\partial s_1} (s_2 \psi(u)) \Delta s_1 + \frac{\partial}{\partial s_2} (s_2 \psi(u)) \Delta s_2 \right)$$

$$- \frac{\log \lambda}{2} \sum_{i,j=1,2} d_{i,j}(\xi_1, \xi_2)(\Delta s_i)(\Delta s_j)$$

for some $\xi = (\xi_1, \xi_2)$ for which $\xi_i$ 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

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

and hence,

$$\frac{d}{dt} \left( \frac{\Delta s_2}{s_2} \right) = \frac{1}{s_2} \left( \frac{d}{dt} \Delta s_2 \right) - \frac{\Delta s_2}{s_2^2} \left( \frac{ds_2}{dt} \right)$$

$$= -\log \lambda \left( 2\psi'(s_1 \Delta s_1 + s_2 \Delta s_2) + \frac{\Delta s_2}{s_2} \psi \right)$$

$$+ (\log \lambda) \frac{\Delta s_2}{s_2} \psi - \frac{\log \lambda}{2} \sum_{i,j=1,2} d_{i,j}(\xi_1, \xi_2) \frac{\Delta s_i \Delta s_j}{s_2}$$

$$= -\frac{2\alpha \log \lambda}{r_0} (s_1^2 + s_2^2)^{\alpha-1} (s_1 \Delta s_1 + s_2 \Delta s_2)$$

$$- \frac{\log \lambda}{2} \sum_{i,j=1,2} d_{i,j}(\xi_1, \xi_2) \frac{\Delta s_i \Delta s_j}{s_2}.$$
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

$$\sum_{i,j=1,2} d_{i,j}(\xi_1, \xi_2) \Delta s_i \Delta s_j \leq \frac{24 \alpha}{r_0^\alpha} \left( \xi_1^2 + \xi_2^2 \right)^{\alpha - \frac{1}{2}} (\Delta s_2)^2.$$ (11)

Therefore, we have

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

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

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

$$\frac{d}{dt} \left( \frac{\Delta s_2}{s_2} \right) \geq -(1 + \mu) \frac{\alpha \log \lambda}{r_0^\alpha} 2^\alpha (s_2^2) 2^\alpha \frac{s_2}{s_2} \Delta s_2$$

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

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

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

Following arguments in [21] (see page 17) one can derive from here that

$$\left( \frac{\xi_1^2 + \xi_2^2}{s_2^2} \right)^{\alpha - \frac{1}{2}} \leq 2$$

and that

$$\frac{d\chi}{dt} \geq -A \frac{\alpha \log \lambda}{r_0^\alpha} s_2 2^\alpha (t) \chi(t),$$
where \( A = (1 + \mu)2^\alpha + \frac{1 - \mu}{3} \). By Grönwall’s inequality (applied to \(-\chi(t)\)) and the second inequality in Lemma 6.2, we obtain that
\[
\chi(t) \geq \chi(0) \exp \left( -A \frac{\log \lambda}{r_0^\alpha} \int_0^t s_2^{2\alpha}(\tau) \, d\tau \right)
\geq \chi(0) \exp \left( -A \frac{\log \lambda}{r_0^\alpha} \int_0^t s_2^{2\alpha}(0)(1 + C_1 s_2^{2\alpha}(0) \tau)^{-1} \, d\tau \right)
= \chi(0) \exp \left( -A \frac{\log \lambda}{C_1} \log(1 + C_1 s_2^{2\alpha}(0) t) \right)
= \chi(0) \exp \left( -\frac{A}{2} \log(1 + C_1 s_2^{2\alpha}(0) t) \right)
= \chi(0) \left( 1 + C_1 s_2^{2\alpha}(0) t \right)^{-\beta},
\]
where \( \beta = \frac{A}{2} = (1 + \mu)2^{\alpha - 1} + \frac{1 - \mu}{6} \).

In the case \( s_2 < 0 \) one can show using an 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 (10), we obtain that
\[
\frac{d}{dt} \left( \frac{\Delta s_2}{s_1} \right) = -\log \lambda \left( 2s_2 \psi' \Delta s_1 + (2s_2^2 \psi' + \psi') \frac{\Delta s_2}{s_1} \right) - \log \psi \frac{\Delta s_2}{s_1}
- \frac{\log \lambda}{2} \sum_{i,j=1,2} d_{i,j}(\xi_1, \xi_2) \frac{\Delta s_i \Delta s_j}{s_1}.
\]

By the assumption \( |\Delta s_1| \leq \mu \Delta s_2 \) and positivity of \( s_1, s_2, \psi', \) and \( \Delta s_2 \), we obtain
\[
\frac{d}{dt} \left( \frac{\Delta s_2}{s_1} \right) \geq -\log \lambda (2\mu s_1 s_2 \psi' + 2s_2^2 \psi' + 2\psi) \frac{\Delta s_2}{s_1}
- \frac{\log \lambda}{2} \sum_{i,j=1,2} d_{i,j}(\xi_1, \xi_2) \frac{\Delta s_i \Delta s_j}{s_1}.
\]

Since \( s_2(t) \leq s_1(t) \) on \([T_1, T]\), we have
\[
2\mu s_1 s_2 \psi' + 2s_2^2 \psi' + 2\psi = 2\psi' (\mu s_1^2 + s_1^2 + \frac{1}{\alpha} 2s_1^2) \leq 2^{\alpha} (\frac{\alpha \mu + \alpha + 2}{r_0^\alpha}) s_1^{2\alpha}.
\]

Using this fact along with the estimate (11), we find that
\[
\frac{d}{dt} \left( \frac{\Delta s_2}{s_1} \right) \geq -\log \lambda 2^{\alpha} \frac{\alpha \mu + \alpha + 2}{r_0^\alpha} \frac{\Delta s_2}{s_1}
- \frac{12\alpha \log \lambda}{r_0^\alpha} s_1^{2\alpha} \left( \frac{\xi_1^2 + \xi_2^2}{s_1^2} \right)^{\alpha - \frac{1}{2}} \left( \frac{\Delta s_2}{s_1} \right)^2.
\]
Let \( \bar{\chi} = \bar{\chi}(t) = \frac{\Delta \alpha}{s_1}(t) \). Then the above inequality can be written as

\[
\frac{d\bar{\chi}}{dt} \geq -\frac{\log \lambda}{\tau_0^{\alpha}} s_1^{2\alpha} \bar{\chi} \left( 2^{\alpha}(\alpha\mu + \alpha + 2) + 12\alpha \left( \frac{\xi_1^2 + \xi_2^2}{s_1^2} \right)^{\alpha - \frac{1}{2}} \bar{\chi} \right).
\]

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

\[
\left( \frac{\xi_1^2 + \xi_2^2}{s_1^2} \right)^{\alpha - \frac{1}{2}} \leq \begin{cases} 
(1 - \bar{\chi})^{2\alpha - 1} & 0 < \alpha \leq \frac{1}{2}, \\
\frac{2^{\alpha - \frac{1}{2}}(1 + \bar{\chi})^{2\alpha - 1}}{\frac{1}{2} \leq \alpha < 1 
\end{cases}
\]

and that \( \bar{\chi} = \frac{\Delta \alpha}{s_2} \) 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

\[
0 \leq \bar{\chi}(T_1) = \frac{\Delta s_2(T_1)}{s_1(T_1)} = \frac{\Delta s_2(T_1)}{s_2(T_1)} \leq \frac{\Delta s_2(0)}{s_2(0)} < \frac{1 - \mu}{72}.
\]

\[
\bar{\chi}(t) = \frac{\Delta s_2(t)}{s_1(t)} \leq \frac{\Delta s_2(T_1)}{s_1(T_1)} (1 + 2^\alpha C_1 s_1^{2\alpha}(t)(t - T_1))^{-\beta} \leq \frac{\Delta s_2(T_1)}{s_1(T_1)}.
\]

Setting \( B = 2^{\alpha}(\alpha\mu + \alpha + 2) + \frac{1 - \mu}{\alpha} \) and combining [12], [13], and [14], we obtain that

\[
\frac{d\bar{\chi}}{dt} \bigg|_{t = T_1} \geq -\frac{B \log \lambda}{\tau_0^{\alpha}} s_1^{2\alpha}(T_1) \bar{\chi}(T_1).
\]

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

\[
\bar{\chi}(t) \geq \bar{\chi}(T_1) \exp \left( -B \frac{\log \lambda}{\tau_0^{\alpha}} \int_{T_1}^t s_1^{2\alpha}(\tau) d\tau \right)
\]

\[
\geq \bar{\chi}(T_1) \exp \left( -B \frac{\log \lambda}{\tau_0^{\alpha}} \int_{T_1}^t s_1^{2\alpha}(t)(1 + C_1 s_1^{2\alpha}(0)(t - \tau))^{-1} d\tau \right)
\]

\[
= \bar{\chi}(T_1) \exp \left( -B \frac{\log \lambda}{\tau_0^{\alpha}} \frac{1}{C_1} \log(1 + C_1 s_1^{2\alpha}(t)(t - T_1)) \right)
\]

\[
= \bar{\chi}(T_1) \left( 1 + C_1 s_1^{2\alpha}(t)(t - T_1) \right)^{-\beta_1},
\]

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

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

In the case \( s_1 < 0 \) one can show using argument similar to the above that the same estimate for \( \bar{\chi}(t) \) 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. \( \square \)
Lemma 6.7. Under the assumptions of Lemma 6.3 one has for some $C_4 > 0$ and all $0 < t < T_1$ that $\Delta s_2(t) \geq C_4 \Delta s_2(0)t^{-\gamma}$.

Proof. By assumption, Lemma 6.6 holds and the first estimate together with the first inequality in Lemma 6.2 yield

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

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

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

It remains to observe that $s_2(0)$ is of order $r_0$ and so $(1 + C_1s_2^{2\alpha}(0))^{-\left(\frac{1}{2\alpha} + \beta\right)}$ 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 6.8. Under the assumptions of Lemma 6.3 one has for some $C_5 > 0$ and all $T_1 < t < T$ that $\Delta s_2(t) \geq C_5 \Delta s_2(T_1)$.

Proof. The fifth inequality in Lemma 6.2 implies that $2^{\alpha}C_1s_1^{2\alpha}(T_1)(t - T_1) < 1$ and hence,

$$1 - C_1s_1^{2\alpha}(T_1)(t - T_1) > 1 - \frac{1}{2\alpha}.$$ 

In addition, since $s_1$ is increasing with respect to time $t$ we have $s_1(t) > s_1(T_1)$ By assumption, Lemma 6.6 holds and the second estimate of this lemma yields

$$\Delta s_2(t) \geq \frac{\Delta s_2(T_1)}{s_1(T_1)}s_1(t)(1 - C_1s_1^{2\alpha}(T_1)(t - T_1))^{\beta_1} \geq \Delta s_2(T_1)(1 - \frac{1}{2\alpha})^{\beta_1}.$$ 

This completes the proof of the lemma. □

Consider a set $\Lambda_s^i$ and note that it consists of full length $s$-curves. Let us fix one of these curves, say $\sigma$.

Lemma 6.9. Assume that $\sigma$ enters the slow down disk $D_\frac{2\alpha}{T}$ around the origin at time $n$, so that the intersection $f^n(\sigma) \cap D_\frac{2\alpha}{T}$ is not empty. Assume that $\sigma$ then exits $D_\frac{2\alpha}{T}$ at time $m$, $m > n > 1$. Then

$$C_6(m - n)^{-\gamma} \leq \frac{L(f^n(\sigma))}{L(f^m(\sigma))} \leq C_7(m - n)^{-\gamma},$$
where $C_6 > 0$, $C_7 > 0$, $\gamma$, $\gamma'$ are as in (8), and $L$ denotes the curve length.

Proof. Let $x$ and $y$ be the endpoints of the curve $\sigma$. For $k \geq 0$ 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 $k \geq 1$,

$$K_0^{-1}d(x_k, y_k) \leq L(f^k(\sigma)) \leq K_0d(x_k, y_k), \tag{16}$$

where $d$ denotes the usual distance.

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

$$K_1^{-1}\|\Delta s(j)\| \leq d(x_j, y_j) \leq K_1\|\Delta s(j)\|. \tag{17}$$

We wish to apply Lemmas 6.3–6.8 to the orbits of $x_n$ and $y_n$ and we need to check the assumptions of Lemma 6.3. 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 (17), 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 $f^j(\sigma) \subset D_{2n} \cap \{(s_1, s_2) : s_2 \geq s_1\}$ for $n \leq j \leq \frac{n+m}{2}$ and $f^j(\sigma) \subset D_{2n} \cap \{(s_1, s_2) : s_2 < s_1\}$ for $\frac{n+m}{2} < j < m$. Applying Lemmas 6.7 and 6.8 with $0 \leq t \leq \frac{n+m}{2}$ and $\frac{n+m}{2} < t < m$ and using (16), we obtain that

$$L(f^m(\sigma)) \geq K_0^{-1}d(x_m, y_m) \geq K_0^{-1}K_1^{-1}\|\Delta s(m-n)\|$$

$$\geq K_0^{-1}K_1^{-1}\Delta s_2(m-n) > K_0^{-1}K_1^{-1}C_5\Delta s_2\left(\frac{m-n}{2}\right)$$

$$\geq K_0^{-1}K_1^{-1}C_5C_4\Delta s_2(0)\left(\frac{m-n}{2}\right)^{-\gamma}$$

$$\geq K_0^{-1}K_1^{-1}C_5C_42^\gamma(m-n)^{-\gamma}\frac{1}{\sqrt{1+\mu^2}}\|\Delta s(0)\|$$

$$\geq K_0^{-2}K_1^{-2}C_5C_42^\gamma(m-n)^{-\gamma}\frac{1}{\sqrt{1+\mu^2}}L(f^n(\sigma)).$$

Therefore, for some $C_6 > 0$,

$$\frac{L(f^m(\sigma))}{L(f^n(\sigma))} \geq C_6(m-n)^{-\gamma}.$$
Using again Lemmas 6.4 and 6.5 and inequalities (16) and (16), similar arguments 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. \(\square\)

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 $\tilde{P}$ and $P$ for the automorphism $A$ and the map $f_{T^2}$ respectively and let $\tilde{P} \in \tilde{P}$ and $P \in P$ 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 $P$ and the number $r_0$ are chosen such that Proposition 5.1 holds and we set again $f := f_{T^2}$. 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 $\Lambda_1^s$ with $\tau(\Lambda_1^s) = n$ and numbers $0 < m_1 < m_2$ satisfying $m_1 < Q_1$, $n - m_2 < Q_1$ such that $f^{k}(\Lambda_1^s) \cap D^1_{r_0} = \emptyset$ for all $0 \leq k < m_1$ or $m_2 < k \leq n$ and $f^{k}(\Lambda_1^s) \cap D^1_{r_0} \neq \emptyset$ for all $m_1 \leq k \leq m_2$.

**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 
\[
P \overline{W}_1 P \overline{W}_2 P,
\]
where the words $\overline{W}_1$ and $\overline{W}_2$ are of length $l(\overline{W}_j) < Q_1$ 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 $\overline{P}_i$ consists of the symbol $P_i$ which is repeated $n - 2 - l(\overline{W}_1) - l(\overline{W}_2)$ times. Since the map $f$ is topologically conjugate to $A$, it is enough to find an admissible word of the form (18) which consists of the corresponding elements of the partition $\tilde{P}$.

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 (18) 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 $\gamma^s$ and $\gamma^u$ respectively. It takes finitely many iterates of $G$
and $G^{-1}$ for each of these curves to completely enter the disk $D_{r_0}^1$. Now for each sufficiently large $n > 0$ with $n \in \mathcal{N}$ there is an $s$-set $A^s_\ell$ with $\tau(A^s_\ell) = n$ which completely enters $D_{r_0}^1$ (under iterates of $G$ and $G^{-1}$) at the same time as $\gamma^s$ and $\gamma^n$ 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 $\gamma$ is defined in (8).

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

$$m(\{x \in \Lambda : \tau(x) > n\}) = \sum_{N=n+1}^{\infty} m(\{x \in \Lambda : \tau(x) = N\})$$

$$= \sum_{N=n+1}^{\infty} \sum_{A^s_\ell \cap \tau(A^s_\ell) = N} m(A^s_\ell) > \sum_{N=n+1}^{\infty} m(A^s_\ell),$$

where $A^s_\ell$ is the set constructed in Lemma 7.1. We wish to obtain a polynomial bound for the measure of the set $A^s_\ell$.

Given $x \in A^s_\ell$, denote by $\gamma^s_\ell(x) := \gamma^s(x) \cap A^s_\ell$ (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

$$m(A^s_\ell) = m(G^N(A^s_\ell)) = K_1 L(G^N(\gamma^s_\ell(x))),$$

where $L$ stands for the length of the curve.

Let $x_j = G^j(x)$ for $j = 0, \ldots, n$. 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 $0 \leq j < k_1$, or $k_2 < j \leq N$;
2. $G^j(x) \in D_{r_0}^1$ if $k_1 \leq j \leq k_2$.

Note that for $0 \leq j < k_1$ and $k_2 < j \leq N$ the curve $G^j(\gamma^s_\ell(x))$ lies in the stable cone for the automorphism $A$ at $x_j$ and indeed, is an admissible manifold for $A$ (i.e., for any $y \in \gamma^s_\ell(x)$ the line $T_y \gamma^s_\ell(x)$ lies in the stable cone at $y$). So the length of the curve $\gamma^s_\ell(x)$ 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

1. $L(\gamma^s_\ell(x)) = \lambda^{k_1} L(G^{k_1}(\gamma^s_\ell(x))) \leq \lambda^{Q_1} L(G^{k_1}(\gamma^s_\ell(x)))$ and
2. $L(G^N(\gamma^s_\ell(x))) = \lambda^{-(N-k_2)} L(G^{k_2}(\gamma^s_\ell(x))) \geq \lambda^{-Q_1} L(G^{k_2}(\gamma^s_\ell(x)))$,

where $\lambda$ is the largest eigenvalue of the matrix $A$. 

By Lemma 5.6 in [21], the time the trajectory spends in $D_{r_0}^1 \setminus D_{r_2}^1$ is uniformly bounded. Thus, by Lemma 6.9,

\begin{equation}
L(G^{k_2}(\gamma^s_t(x))) > C_6(k_2 - k_1)^{-\gamma} L(G^{k_1}(\gamma^s_t(x))).
\end{equation}

Since $k_2 - k_1 < N$, combining Equations (19)-(22) yields
\[ m(\Lambda^s_t) \geq K_2 L(G^N(\gamma^s_t(x))) = K_2 \lambda^{-Q_1} L(G^{k_2}(\gamma^s_t(x))) \geq K_2 C_6 \lambda^{-Q_1} (k_2 - k_1)^{-\gamma} L(G^{k_1}(\gamma^s_t(x))) \geq K_3 C_6 \lambda^{-2Q_1} (k_2 - k_1)^{-\gamma} L(\gamma^s_t(x)) \]
where $K_2 > 0$ is a constant and $K_3 = K_3 C_6 \lambda^{2Q_1} L(\gamma^s_t(x))$.

Note that $\gamma^s_t(x)$ is a full length stable curve in $P$ and hence, has length which is independent of $N$. It follows that
\[ m(\{x \in \Lambda : \tau(x) > n\}) > \sum_{N=n+1}^{\infty} m(\Lambda^s_t) > C_8 \frac{1}{n^{\gamma-1}}, \]
where $C_8 > 0$ is a constant. The desired lower bound follows. \qed

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 $D_{x_i}$ 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 $D_{r_0}^1 \subset P_i$. Using the conjugacy (6), it suffices to establish that upper bound for the map $G$.

Given an $s$-set $\Lambda^s_t \subset P$ with $\tau(\Lambda^s_t) = n$, choose any numbers $k = k(\Lambda^s_t)$, $p = p(\Lambda^s_t)$, and two finite collections of numbers $\{k_m \geq 0\}_{m=1,\ldots,p}$ and $\{l_m \geq 0\}_{m=1,\ldots,p}$ 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 $\Lambda^s_t$ 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
\[ S_{k,n,p} = \{ \Lambda^s_t \subset P : \tau(\Lambda^s_t) = n, k = k(\Lambda^s_t), p = p(\Lambda^s_t) \}. \]

Lemma 8.1. There are $0 < h < h_{top}(f)$, $\varepsilon_0 > 0$, and $C_9 > 0$ such that
\[ \varepsilon_0 < h_{top}(f) - h \text{ and } \]
\[ \text{Card } S_{k,n,p} \leq C_9 \frac{1}{p^2} e^{(h+\varepsilon_0)(n-k)}. \]
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 $\Lambda^s_i$ 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 $\Lambda^s_i$ 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 $\Lambda^s_i$ may enter any of the disks $D_{r_0}^i$, $i = 1, 2, 3, 4$, we obtain

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

Write

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

To prove the lemma we wish to estimate $p^2 A^p \binom{k-1}{p-1} \binom{n-k-1}{p}$ 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 $\Lambda^s_i$ 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 $p - 1 = \lceil \frac{k-1}{2} \rceil$ or $p - 1 = \lfloor \frac{k-1}{2} \rfloor + 1$. We may assume $p - 1 = \frac{k-1}{2}$. Then using the asymptotic formula $\binom{n}{l} \sim \left(\frac{ne}{l}\right)^l$, we obtain that

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

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

$$\binom{n-k}{p} < \binom{n-k}{\frac{n-k}{Q}} < \left(\frac{(n-k)e}{\frac{n-k}{Q}}\right)^{\frac{n-k}{Q}} < e^{\frac{n-k}{Q} \ln \left(\frac{(n-k)e}{\frac{n-k}{Q}}\right)} < e^{\frac{n-k}{Q} \ln(e \frac{n-k}{Q})}.$$
Finally, note that
\[ p^2 A^p < e^2 \ln p \ln 4 < e^{2p+p \ln 4} < e^{\frac{n-k}{4} (\ln 4 + 2)}. \]
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
\[ q^p \left( \frac{k-1}{p-1} \right) \left( \frac{n-k-1}{p} \right) e^{h(n-k)} < e^{(n-k) \varepsilon_0} e^{h(n-k)} = e^{(n-k) \varepsilon_0 + h} \]
and hence,
\[ \text{Card } S_{k,n,p} \leq C_9 \frac{1}{p^2} e^{(h+\varepsilon_0)(n-k)}. \]
This completes the proof of the lemma.

\[ \square \]

**Lemma 8.2.** There exists \( \varepsilon_0 > 0 \) such that for any \( \Lambda_i^* \in S_{k,n,p}, \)
\[ m(\Lambda_i^*) \leq C_{10} k^{-\gamma'} e^{(-\log \lambda + \varepsilon_0)(n-k)}, \]
where \( C_{10} > 0 \) is a constant and \( \gamma' \) is given by (8).

**Proof.** Note that by (19), \( m(\Lambda_i^*) = m(G^n(\Lambda_i^*)) = K_1 L(G^n(\gamma_i^*(x))) \) and that the length of the backward iterates of \( \gamma_i^* \) lying outside the region \( D_{r_0} \) are stretched by the largest eigenvalue \( \lambda \) of the matrix \( A \). Note also that every time the iterates of \( \gamma_i^* \) enter the region we have an upper estimate for its length according to Lemma 6.9 (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,
\[ m(\Lambda_i^*) = m(G^n(\Lambda_i^*)) = K_1 L(G^n(\gamma_i^*(x))) = K_1 \lambda^{-l_{p+1}} L(G^{n-l_{p+1}}(\gamma_i^*(x))) \]
\[ \leq K_1 C_7 \lambda^{-l_{p+1}} k_p^{-\gamma'} L(G^{-(l_{p}+1)}(\gamma_i^*(x))) \leq \cdots \]
\[ \leq K_1 C_7^{p} \lambda^{-((l_{p+1} + \cdots + l_1)} k_p^{-\gamma'} \cdots k_1^{-\gamma'} L(\gamma_i^*(x))). \]
Note that \( \gamma_i^*(x) \) 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
\[ k_1 k_2 \cdots k_p \geq k_{\text{max}} 2^{p-1} \geq k_{\text{max}} p \geq \sum_{i=1}^{p} k_i = k, \]
where \( k_{\text{max}} \) denotes the largest of \( k_i^\prime \)'s.
In addition, \( C_6^p = e^{p \ln c_6} < e^{n-k \ln c_6} < e^{\varepsilon_0 (n-k)} \) for sufficiently small \( \varepsilon_0 > 0 \) if one chooses \( Q \) large. Therefore,
\[ m(\Lambda_i^*) < K_1 e^{\varepsilon_0 (n-k)} \lambda^{-(n-k)} k^{-\gamma'} < C_{10} k^{-\gamma'} e^{(-\log \lambda + \varepsilon_0)(n-k)} . \]
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 $\gamma'$.

Note that

$$m(\{x \in \Lambda : \tau(x) = n\}) \leq \sum_{k=1}^{n} \sum_{p=1}^{k} \max_{\Lambda_{\gamma} \in S_{k,n,p}} \{m(\Lambda_{\gamma})\}\text{Card } S_{k,n,p}.$$

Therefore, by Lemmas 8.1 and 8.2, we have

$$m(\{x \in \Lambda : \tau(x) = n\})$$

$$\leq \sum_{k=1}^{n} \sum_{p=1}^{k} \frac{1}{p^2} C_{9} e^{(h+\varepsilon_0)(n-k)} C_{10} e^{(\varepsilon_0 - \log \lambda)(n-k) k^{-\gamma'}}$$

$$< C_{9} C_{10} \frac{n^2}{6} e^{-\delta n} \sum_{k=1}^{n} e^{\delta k} k^{-\gamma'},$$

where $\delta = 2\varepsilon_0 + \log \lambda - h > 0$ if $\varepsilon_0$ 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,

$$m(\{x \in \Lambda : \tau(x) = n\}) \leq C_{9} C_{10} e^{-\delta n} \sum_{k=1}^{n} e^{\delta k} k^{-\gamma'} < C_{9} C_{10} n^{-\gamma'}.$$

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:

\[
\begin{array}{ccc}
T^2 & \xrightarrow{\varphi_1} & S^2 & \xrightarrow{\varphi_2} & D^2 & \xrightarrow{\varphi_3} & M \\
\downarrow f_{T^2} & & \downarrow f_{S^2} & & \downarrow f_{D^2} & & \downarrow f_M \\
T^2 & \xrightarrow{\varphi_1} & S^2 & \xrightarrow{\varphi_2} & D^2 & \xrightarrow{\varphi_3} & M
\end{array}
\]

We stress that while our construction of the maps \( \varphi_1 \) and \( \varphi_2 \) follows [12], our construction of the map \( \varphi_3 \) 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 \( \varphi_1: T^2 \to S^2 \) satisfying:

1. \( \varphi_1 \) is a double branched covering, is one-to-one on each branch, and \( C^\infty \) everywhere except at the points \( x_i, i = 1, 2, 3, 4 \) where it branches;
2. \( \varphi_1 \circ I = \varphi_1 \) where \( I: T^2 \to T^2 \) is the involution map given by \( I(t_1, t_2) = (1 - t_1, 1 - t_2) \);
3. \( \varphi_1 \) 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 \( D^i \) such that
   \[
   \varphi_1(s_1, s_2) = \left( \frac{s_1^2 - s_2^2}{\sqrt{s_1^2 + s_2^2}}, \frac{2s_1s_2}{\sqrt{s_1^2 + s_2^2}} \right);
   \]
5. The map \( f_{S^2} := \varphi_1 \circ f_{T^2} \circ \varphi_1^{-1} \) 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 \( \varphi_2 \) 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 \( \varphi_2 \) to an area preserving \( C^\infty \) diffeomorphism (still denoted by \( \varphi_2 \)) between \( S^2 \setminus \{p_4\} \) and the interior of the unit disk \( D^2 \).
The map
\[
    f_D^2 := \begin{cases} 
    \varphi_2 \circ f_{S^2} \circ \varphi_2^{-1} & \text{on int} D^2 \\
    \text{Id} & \text{on } \partial D^2 
    \end{cases}
\]
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 \( 2\kappa \). 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
\[
    \frac{\partial^2 g}{\partial x^2}(x, y) = \begin{cases} 
    2\delta y(x^2 + y^2)^{1-\delta-2(1-\delta)}x^2y(x^2 + y^2)^{-\delta} & (x, y) \neq (0, 0); \\
    0 & (x, y) = (0, 0).
    \end{cases}
\]
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,
\[ |x^2y(x^2+y^2)^{\delta-2}| \leq d((x, y), (0, 0))^{2\delta-1}. \]
Hence, $\frac{\partial^2 g}{\partial x^2}$ is Hölder continuous with Hölder exponent
\[ 2\delta - 1 = 2 \left( \frac{1}{1 - \alpha} - \frac{1}{2} \right) - 1 = 2\kappa. \]
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
\begin{equation}
(25) \quad 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}
\end{equation}
and are 0 at the point $(x, y) = (0, 0)$. Arguing as above one can show that each function in (25) is Hölder continuous with Hölder exponent $2\kappa$. This completes the proof of the proposition. \hfill $\square$

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 $C^\infty$ diffeomorphism of the disk, which is identity on the boundary $\partial D^2$ 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
\begin{equation}
(26) \quad V_n \subset \overline{V}_n \subset V_{n+1} \text{ and } \bigcup_{n \geq 1} V_n = D^2,
\end{equation}
we have that for every $n \geq 1$,
\[ V_{n-1} \subset f(V_n) \subset V_{n+1} \text{ and } \|f - \text{Id}\|_{C^\infty(V_{n+1} \setminus V_{n-1})} \leq \rho_n. \]
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. $\varphi$ is of class $C^\infty$ in the interior of the disk;
2. $\varphi$ is area preserving, i.e., $h_* m_{D^2} = m_M$;
3. the map $\varphi \circ f \circ \varphi^{-1}$ is a $C^\infty$ 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 (26), such that for every $n \geq 1$ we have $V_{n-1} \subset f_{D^2}(V_n) \subset V_{n+1}$ and for every $0 < \beta < 2 + 2\kappa$,
\begin{equation}
(27) \quad \|f_{D^2} - \text{Id}\|_{C^{1+\beta}(V_{n+1} \setminus V_{n-1})} \leq (r_{n-1})^{2+2\kappa-\beta},
\end{equation}
where \( r_n = \text{dist}(V_n, \partial D^2) \). This requires us to develop a specific construction of the homeomorphism \( \varphi \) 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 \( M \) of genus \( \geq 2 \) and numbers \( 0 < \alpha < 1 \) and \( 0 < \mu < 1 \), there exist \( \beta = \beta(\alpha, \mu) > 0 \) and a continuous map \( \varphi_3 : \overline{D^2} \to M \) such that

1. the restriction \( \varphi_3|\text{int} \ D^2 \) is a diffeomorphic embedding;
2. \( \varphi_3(\overline{D^2}) = M \);
3. \( \varphi_3 \) 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;
5. \( f_M|(M \setminus \varphi_3(\text{int} \ D^2)) = \text{Id} \).

**Remark 1.** Let \( B_M = (M \setminus \varphi_3(\text{int} \ D^2)) \cup \bigcup_{i=1,2,3} q_i \),

where \( q_i = (\varphi_3 \circ \varphi_2 \circ \varphi_1)(x_i) \), \( i = 1,2,3 \) and \( x_i \) are the fixed points of the automorphism \( A \) in the construction of the Katok map (see Section 4.1). Clearly, \( q_i \) are the fixed points of the map \( f_M \). Therefore, by Statement 5 of Theorem 9.3, we have that \( f_M|B_M = \text{Id} \). We call the set \( B_M \) the boundary set for the map \( f_M \).

We proceed with the proof of Theorem 9.3. One can represent the surface \( M \) as a regular \( p \)-polygon \( P \) (the number \( p \) is even) whose angles are \( \hat{\alpha} = \frac{\pi(p-2)}{p} \). Let \( A_1, A_2, \ldots, A_p \) be vertices of the polygon and \( O \) its center. For each \( i = 2, \ldots, p \) denote by \( B_i, \tilde{B}_i \) the points on the segment \( A_iO \) for which \( \frac{|A_iB_i|}{|A_iO|} = \frac{1}{3} \) and \( \frac{|\tilde{B}_iO|}{|A_iO|} = \frac{1}{3} \). In what follows we assume that \( A_{p+1} = A_1 \), \( B_{p+1} = B_1 \), and \( \tilde{B}_{p+1} = A_1 \). Denote by

\[
(28) \quad P^* := A_1O \cup \left( \bigcup_{i=1}^p A_iA_{i+1} \right) \cup \left( \bigcup_{i=2}^p (A_iB_i \cup \tilde{B}_iO) \right).
\]

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 \( C^\infty \) diffeomorphisms \( h_n : U_n \to U_{n+1} \) for \( n \geq 0 \);
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

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

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 \( U_n \).

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 \( i = 1, \ldots, p \) consider the following collection of points associated with the point \( A_i \):

- \( K_{ti} \), the point that is determined uniquely by the requirements that the angle \( \angle(K_{ti}A_iA_{i+1}) = \frac{1}{8} \alpha \) and \( \text{dist}(K_{ti}, A_iA_{i+1}) = r(t) \);
- \( L_{ti} \), the point that is determined uniquely by the requirements that the angle \( \angle(\overline{OA_iL_{ti}}) = \frac{1}{8} \alpha \) and \( \text{dist}(\overline{L_{ti}}A_iO) = r(t) \);
- \( M_{ti} \), the image of \( K_{ti} \) 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 \( \text{dist}(E_{ti}, B_i) = r(t) \);
- \( F_{ti} \), the image of \( E_{ti} \) under the reflection about the line \( A_iB_i \).
- \( \tilde{L}_{ti} \), the point that is determined uniquely by the requirements that the angle \( \angle(A_iO\tilde{L}_{ti}) = \frac{1}{4}(\pi - \alpha) \) and \( \text{dist}(\tilde{L}_{ti}, O\tilde{B}_i) = r(t) \);
- \( \tilde{M}_{ti} \), the image of the point \( \tilde{L}_{ti} \) under the reflection about the line \( A_iO \);
- \( \tilde{E}_{ti} \), the point on the line through \( \tilde{B}_i \) which is perpendicular to the line \( O\tilde{B}_i \) and such that \( \text{dist}(\tilde{E}_{ti}, \tilde{B}_i) = r(t) \);
- \( \tilde{F}_{ti} \), the image of the point \( \tilde{E}_{ti} \) under the reflection about the line \( O\tilde{B}_i \).

We introduce the following curves: for \( i = 1, \ldots, p \) 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;
- \( \tilde{\gamma}_{ti}^{(2)} \) be a curve connecting the points \( \tilde{L}_{ti} \) and \( \tilde{M}_{t(i+1)} \) to be determined later in Step 2 where we assume that \( \tilde{M}_{t(p+1)} = \tilde{M}_{t1} \).

Now for \( i = 2, \ldots, p \) let
\( \gamma_{ti}^{(4)} \) be a line segment connecting the points \( L_{ti} \) and \( E_{ti} \);
\( \gamma_{ti}^{(5)} \) be a 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;
\( \tilde{\gamma}_{ti}^{(4)} \) be a line segment connecting the points \( \tilde{L}_{ti} \) and \( \tilde{E}_{ti} \);
\( \tilde{\gamma}_{ti}^{(5)} \) be a line segment connecting the points \( \tilde{M}_{ti} \) and \( \tilde{F}_{ti} \);
\( \tilde{\gamma}_{ti}^{(6)} \) be a curve connecting the points \( \tilde{E}_{ti} \) and \( \tilde{F}_{ti} \) to be determined later in Step 2.

Finally, we let \( \sigma_t^{(1)} \) to be a line segment connecting the points \( L_{t1} \) and \( \tilde{L}_{t1} \) and \( \sigma_t^{(2)} \) to be a line segment connecting the points \( M_{t1} \) and \( \tilde{M}_{tp} \).

Let \( \tau_t \) be the curve given by
\[
\tau_t = \sigma_t^{(1)} \cup \sigma_t^{(2)} \cup \left( \bigcup_{i=1}^{p} \left( \gamma_{ti}^{(1)} \right) \right) \cup \left( \bigcup_{i=2}^{p} \left( \bigcup_{j=2}^{6} \left( \gamma_{ti}^{(j)} \cup \tilde{\gamma}_{ti}^{(j)} \right) \right) \right).
\]

By construction, \( \tau_t, t \in [n, n + 1] \), 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 curves \( \gamma_{ti}^{(3)} \), \( \tilde{\gamma}_{ti}^{(2)} \) and \( \tilde{\gamma}_{ti}^{(6)} \) 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 \( 1 - \tan \frac{\hat{\alpha}}{4} < a < \cot \frac{\hat{\alpha}}{4} \) and \( \varphi(-1) = \varphi(1) = \cot \frac{\hat{\alpha}}{8}; \)
4. \( \varphi'(0) = 1 \) and \( \psi(-1) = \psi(1) = 0; \)
5. \( 0 < \varphi'(x) \leq \cot \frac{\hat{\alpha}}{4} \) for \( 0 < x \leq 1 \) and \( -\cot \frac{\hat{\alpha}}{4} \leq \varphi'(x) < 0 \) for \( -1 \leq x < 0; \)
6. \( \varphi'(-1) = -\cot \frac{\hat{\alpha}}{4} \) and \( \varphi'(1) = \cot \frac{\hat{\alpha}}{4}; \)
7. \( \psi \) is infinitely vertically flat at \(-1 \) and \( 1 \).

For \( i = 1, \ldots, p \) 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 \( \varphi_t(x) = r(t) \varphi\left( \frac{x}{r(t)} \right) \) where \( -r(t) \leq x \leq r(t) \). It is easy to see that \( \gamma_{ti}^{(2)} \) is a \( C^\infty \) 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 \( A_i B_i \). We let \( \gamma_{ti}^{(6)} \) be the graph of the
function $\psi_t(x) = r(t)\psi(x/r(t))$ where $-r(t) \leq x \leq r(t)$. It is easy to see that $\gamma^{(6)}_{ti}$ is a $C^\infty$ curve that connects the points $E_{ti}$ and $F_{ti}$ and is infinitely tangent to the lines $L_{ti}, E_{ti}$ and $M_{ti}, F_{ti}$.

It is easy to see that with the above choice of curves $\gamma^j_{ti}, j = 1, \ldots, 6$, the curve $\tau_t$ is of class $C^\infty$.

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

$$\varphi_{t_1}(0) = ar(t_1), \quad \varphi_{t_1}(r(t_1)) = r(t_1) \cot \frac{\hat{\alpha}}{8}, \quad \varphi'_{t_1}(r(t_1)) = \cot \frac{\hat{\alpha}}{4}. \quad (1)$$

Similarly,

$$\varphi_{t_2}(0) = ar(t_2), \quad \varphi_{t_2}(r(t_2)) = r(t_2) \cot \frac{\hat{\alpha}}{8}, \quad \varphi'_{t_2}(r(t_2)) = \cot \frac{\hat{\alpha}}{4}. \quad (2)$$

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

$$\varphi_{t_1}(r(t_2)) = r(t_1)\varphi\left(\frac{r(t_2)}{r(t_1)}\right) \geq \varphi_{t_2}(r(t_2)) = r(t_2) \cot \frac{\hat{\alpha}}{8}. \quad (3)$$

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

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

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

- $\hat{h}_st : \sigma^{(j)}_s \rightarrow \sigma^{(j)}_t$ is a linear map, given by $\hat{h}_st(z) = \frac{r(t)}{r(s)}(z)$, $z \in \sigma^{(j)}_s$ and $j = 1, 2$;
- $\hat{h}_st : \gamma^{(1)}_s \rightarrow \gamma^{(1)}_t$ is a linear map, given by $\hat{h}_st(z) = \frac{r(t)}{r(s)}(z)$, $z \in \gamma^{(1)}_s$ and $i = 1, \ldots, p$;
- $\hat{h}_st : \gamma^{(j)}_s \rightarrow \gamma^{(j)}_t$ is a map, given by $\hat{h}_st(z) = v$, where $z = (y, \varphi_s(y))$ and $v = \left(\frac{r(t)}{r(s)}y, \varphi_t(y)\right)$ for $-r(s) \leq y \leq r(s)$, $i = 2, \ldots, p$, $j = 2, \ldots, 6$;
- $\hat{h}_st : \tilde{\gamma}^{(j)}_s \rightarrow \tilde{\gamma}^{(j)}_t$ is a map, given by $\hat{h}_st(z) = v$, where $z = (y, \psi_s(y))$ and $v = \left(\frac{r(t)}{r(s)}y, \psi_t(y)\right)$ for $-r(s) \leq y \leq r(s)$, $i = 2, \ldots, p$, $j = 2, \ldots, 6$. 

Now given $n - 1 \leq s \leq n$, define the map $\hat{h}_s := \hat{h}_a$ 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 $C^\infty$ 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 $C^\infty$ diffeomorphism from $\text{int}D^2$ onto $P \setminus P^*$ where $P^*$ is given by \[28\]. It also follows from the construction of the map $h$ that there is $C > 0$ such that

$$
\|h\|_{C_1} \leq C, \quad \|h^{-1}\|_{C_1} \leq C.
$$

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

$$
\|f_P - \text{Id}\|_{C^{1+\beta}(U_{n+1} \setminus U_n)} \to 0
$$
as $n \to \infty$.

First, we will prove the following lemma.

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

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

Then for sufficiently large $n$ the open sets $V_n = h^{-1}(U_n)$ satisfy \[26\] 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

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

Furthermore, in view of \[29\] there are $C_4 \geq C_3 > 0$ such that

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

Since the map $f_{D^2}$ is identity on $\partial D^2$ 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 $\delta > 0$

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

Therefore, setting $\delta = 2\kappa - \delta$ and assuming that $\delta > 0$ for any $x \in V_n$ we obtain that

$$
C_3 r(n) - r(n)^{2+\delta} \leq \text{dist}(f_{D^2}(x), \partial D^2) \leq C_4 r(n) + r(n)^{2+\delta}.
$$
To prove the desired inclusion (32) we will show that

\[ C_4 r(n + 2) < C_3 r(n) - r(n)^{2+\delta} \leq C_4 r(n) + r(n)^{2+\delta} < C_3 r(n - 2). \]

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

\[ C_4 r(n + 2) < C_3 r(n) - r(n)^{2+\delta} \iff \]

\[ C_4 \frac{r(n + 2)}{r(n)} < C_3 - r(n)^{1+\delta} \iff \]

\[ C_4 r(n)^3 < C_3 - r(n)^{1+\delta}. \]

For large values of \( n \) both \( C_4 r(n)^3 \) and \( r(n)^{1+\delta} \) are small. Hence, the last inequality holds for all sufficiently large \( n \).

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

\[ C_4 r(n) + r(n)^{2+\delta} < C_3 r(n - 2) \iff \]

\[ C_4 + r(n)^{1+\delta} < C_3 \frac{r(n - 2)}{r(n)} \iff \]

\[ C_4 + r(n)^{1+\delta} < C_3 r(n)^{-\frac{2}{3}}. \]

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.

Fix \( \beta > 0 \) whose choice will be specified later (see (37)). The above lemma allows us to write

\[ \| h \circ (f_{D^2} - \text{Id}) \circ h^{-1} \|_{C^{1+\beta}(U_{n+1} \setminus U_n)} \leq \]

\[ \| h \|_{C^{1+\beta}(V_{n+2} \setminus V_{n-1})} \| f_{D^2} - \text{Id} \|_{C^{1+\beta}(V_{n+1} \setminus V_n)} \| h^{-1} \|_{C^{1+\beta}(U_{n+1} \setminus U_n)}. \]

Observe that \( \| f_{D^2} - \text{Id} \|_{C^{1+\beta}} \) admits Estimate (27) 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 \( n = 0, 1, 2, \ldots \).

Further, it suffices to estimate the norm of \( h \) restricted to the boundary of the sets \( U_n \). Recall that \( \tau_n \), the boundary of \( U_n \), is a union of the curves \( \sigma^{(1)}_n, \sigma^{(2)}_n, \gamma^{(j)}_n \), and \( \bar{\gamma}^{(j)}_n \), \( j = 1, \ldots, 6 \). Note that the map \( h \) acts linearly on the curves \( \sigma^{(1)}_n, \sigma^{(2)}_n \), and \( \gamma^{(1)}_n \), and hence, the norm of \( h \) restricted to these parts of the curve \( \tau_n \) are bounded.

The curves \( \gamma^{(j)}_n \) and \( \bar{\gamma}^{(j)}_n \), \( j \geq 2 \), are the graphs of the functions \( \varphi_n \) and \( \psi_n \) respectively. We shall only give an estimate of the norm of \( h \) restricted to the curves \( \gamma^{(j)}_n \), since the estimates of the norm of \( h \) restricted to the curves \( \bar{\gamma}^{(j)}_n \) is similar. Also, we can assume that \( i \) and \( j > 1 \) are fixed.
Now for a fixed curve $\gamma^{(j)}_{ni}$ 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

$$h_n : (x, \varphi_n(x)) \rightarrow \left( \frac{r(n + 1)}{r(n)} x, \varphi_{n+1}(x) \right), \quad -r(n) \leq x \leq r(n).$$

Since $\varphi_n$ is isometric, we can further assume that $x > 0$. We have that

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

where

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

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 $h_n = (h_n^{(1)}, h_n^{(2)}), \ y = \varphi(x)$ where

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

Since

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

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

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

$$\partial_y h_n^{(2)}(x, y) = \frac{\partial}{\partial y} \left( \frac{r(n + 1)}{r(n)} x \right) = \frac{\partial}{\partial y} \left( \frac{r(n + 1)}{r(n)} \varphi_n^{-1}(y) \right)
= r(n + 1)(\varphi^{-1} \left( \frac{x}{r(n)} \right) \frac{1}{r(n)} < 1.$$  

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

$$\partial h_n^{(2)}(x, y) = \frac{\partial}{\partial y} \left( \varphi'_{n+1} \left( \frac{r(n)}{r(n + 1)} \frac{y}{r(n)} \right) \right)
= r(n + 1) \varphi' \left( \frac{r(n)}{r(n + 1)} \varphi^{-1} \left( \frac{y}{r(n)} \right) \right) \left( \varphi^{-1} \left( \frac{y}{r(n)} \right) \frac{1}{r(n)} \right) < K_2.$$  

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

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

(34)
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

\[
\| h^{-1} \|_{C^{1+\beta}(U_{n+1}\backslash U_n)} = \| h_n^{-1} \circ h_{n-1}^{-1} \circ \cdots \circ h_0^{-1} \|_{C^{1+\beta}(U_{n+1}\backslash U_n)}
\]

\[
\leq \prod_{i=0}^{n} \frac{r(n)^{1-\beta}}{r(n+1)} = \prod_{i=0}^{n} r(n)^{-1-\beta}.
\]

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

\[
\prod_{i=0}^{n} r(n)^{-1-\beta} = r(n)^{-\left(1+\beta\right)+\frac{1}{2}+\frac{1}{2^{n+1}}+\cdots+\frac{1}{2n+1}} = r(n)^{2\left(1+\beta\right)-\frac{1}{2n+2}}.
\]

Finally, using (33), we find that

\[
\| h \circ (f_{D^2} - \text{Id}) \circ h^{-1} \|_{C^{1+\beta}(U_{n-1})} \leq r(n)^{-2\beta} r(n)^{2+2\kappa-\beta} r(n)^{-2\left(1+\beta\right)-\frac{1}{2n+2}} = r(n)^{2\kappa-5\beta+\frac{1}{2n+1}+\frac{\beta}{2n+1}}.
\]

For any \( 0 < \alpha < 1 \) one can choose \( \beta = \beta(\alpha) \) such that

\[
\beta < \frac{2\alpha}{5(1-\alpha)}.
\]

This implies that for any sufficiently large \( n \),

\[
2\kappa - 5\beta + \frac{1}{2n+1} - \frac{\beta}{2n+1} > 0
\]

and conclude that

\[
\| h \circ (f_{D^2} - \text{Id}) \circ h^{-1} \|_{C^{1+\beta}(U_{n-1})} \to 0
\]

as \( n \to \infty \). Therefore, \( f_P \) is tangent to \( \text{Id} \) near \( \partial P \) 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 \( C^{\infty} \) 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 $C^\infty$ diffeomorphism $\psi : D^2 \to D^2$ that can be continuously extended to $\partial D^2$ such that $\psi_* \mu = m_{D^2}$.

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

(i) $\mu_n \in C^\infty(D^2)$ that is the measure $\mu_n$ is absolutely continuous with respect to $m_{D^2}$ with density function of class $C^\infty$;
(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 [8], Lemma 1).

Lemma 9.6. Let $\omega$ and $\mu$ 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 $C^\infty$ 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 $\mu_{n+1}|(U_n \setminus U_n)$ and $\mu_n|(\overline{U}_n \setminus U_n)$, we obtain a $C^\infty$ 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

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

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 (31) and using (34), (35), and (36), we find that $\text{diam } \hat{\psi}_n^{-1}(U_n) \leq C d_n$ 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 $d(x, \hat{\psi}_n(x)) \leq C d_n$ for any $x \in D^2$. It follows that for any $x \in D^2$ and $n > j > 0$,

$$d(\psi_j(x), \psi_n(x)) \leq \sum_{i=j}^{n-1} d(\psi_i(x), \psi_{i+1}(x))$$
$$\leq \sum_{i=j}^{n-1} d(\psi_i(x), \hat{\psi}_i(x)) \leq C \sum_{i=j}^{n-1} d_i.$$
This implies that the sequence $\psi_n$ is uniformly Cauchy and hence, $\psi$ is well defined and continuous on $D^2$. We can also get that $\psi: D^2 \to D^2$ is a $C^\infty$ diffeomorphism.

By construction, we know that $(\psi_n)_*\mu_{n+1} = \mu_1 = m_{D^2}$. Note that $D^2 = \bigcup_{n \geq 1} h^{-1}(U_n)$. 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 genus $g \geq 0$ 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 genus $g \geq 1$. Consider the collection of $s$-subsets $\Lambda^s_i$ and the return time $\tau: \Lambda \to \mathbb{N}$ for the map $f_{T^2}$ defined in Section 5.3. Define $\Delta^s_i := \varphi_3(\varphi_2(\varphi_1(\Lambda^s_i)))$ with the return time on $M$ (again denoted by $\tau$) given by $\tau(\varphi_3(\varphi_2(\varphi_1(x)))) = \tau(x)$, $x \in T^2$. Let $\Delta = \bigcup_i \Delta^s_i$. We claim that $f_M$ is a Young diffeomorphism with respect to the collection of $s$-subsets $\Delta^s_i$.

To prove this we need to check Conditions (Y1)–(Y6). Since the maps $\varphi_i$, $i = 1, 2, 3$ are homeomorphisms and (Y1) and (Y2) are satisfied for the map $f_{T^2}$, then these conditions are also satisfied for $f_M$. In addition, (Y5) and (Y6) hold true for $f_M$ since the maps $\varphi_i$, $i = 1, 2, 3$ 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 $f_{T^2}$ 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 $\hat{f}_M = f_{M}^{\hat{f}_M}$ is defined on $\Delta \subset P$, (Y4) follows as it hold true for the map $f_{T^2}$ by Proposition 5.2. Note that (Y3) holds for the map $f_{T^2}$ 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 $\Lambda$, $s$-sets $\Lambda^s_i$, and inducing time $\tau = \{\tau_i\}$. Consider the associated Young tower given by (see Section 5.4)

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

$$\hat{M}_k = \{(x,\ell) \in \hat{Y} : 0 \leq \ell \leq \min\{k, \tau(x)\}\}.$$ Consider the projection $\pi : \hat{Y} \to M$ given by $\pi(x,\ell) = f^\ell(x)$ and let $Y_M = \pi(\hat{Y}), M_k = \pi(\hat{M}_k)$. The sets $M_k$ are clearly nested and exhaust $Y_M$.

Remark 2. For every $k > 0$ the complement of the set $\overline{M}_k$ (the closure of the set $M_k$) is a neighborhood of the boundary set $B|M$ on which $f_M$ acts as the identity (see Remark 7).

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 [7] and [22]).

Proposition 10.2. Assume that

- the greatest common divisor of numbers $\{\tau_i\}$, $\gcd\{\tau_i\} = 1$, where $\tau_i$ are the values of the function $\tau$;
- there is $C > 0$ such that for all $x, y \in \Delta^s_i$ and $0 \leq j \leq \tau_i$,

$$d(f^j(x), f^j(y)) \leq C \max\{d(x,y), d(f^{\tau_i}(x), f^{\tau_i}(y))\};$$
- there are $\nu > 0$ and $C_1 > 0$ such that

$$m(\tau > n) \leq \frac{C_1}{n^\nu}.$$ Then the following statements hold:

1. There is $C_2 > 0$ such that $\text{Cor}_n(h_1, h_2) \leq \frac{C_2}{n^\nu \tau}$ for any $h_1, h_2 \in C^\rho(M)$.
2. For any $h_1, h_2 \in C^\rho(M)$ supported in $\overline{M}_k$ for some $k > 0$, we have

$$\text{Cor}_n(h_1, h_2) = \sum_{k=n+1}^{\infty} m(\{x : \tau(x) > k\}) \int_M h_1 dm \int_M h_2 dm + r_\nu(n),$$
where \( r_\nu(n) = O(R_\nu(n)) \) and

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

Moreover, if \( \int_M h_1 \int_M h_2 = 0 \), then \( \text{Cor}_n(h_1,h_2) = O(1/n^\nu) \).

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 \( \varphi_i, \ i = 1,2,3 \) 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 (38) observe that this is true for the map \( f_\tau^{\tau_2} \) of the torus which is smoothly conjugate to the map \( f_M^{\tau_1} \).

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

\[
C_8 \frac{1}{n^{\gamma - 1}} < m(\{x \in \Delta: \tau(x) > n\}) < C_{11} \frac{1}{n^{\gamma' - 1}},
\]

where \( \gamma \) and \( \gamma' \) are defined by (8). By (9), \( \gamma > \gamma' > 2 \) for all \( 0 < \alpha < \frac{1}{4} \) and \( 0 < \mu < 1 \).

Since the homeomorphisms \( \varphi_1, \varphi_2, \) and \( \varphi_3 \) are measure preserving we also have the same estimates for the map \( f_M^{\tau_1} \). In particular, (39) 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 \( \nu = \gamma' - 1 \) and obtain for all \( h_1, h_2 \in C^\alpha(M) \) supported in \( M_k \) for some \( k > 0 \) that

\[
\text{Cor}_n(h_1,h_2) = \sum_{k=n+1}^{\infty} m(\{x: \tau(x) > k\}) \int_M h_1 dm \int_M h_2 dm + r_{\gamma'}(n),
\]

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

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

Assume first that $\gamma' > 3$, which is true if $0 < \alpha \leq \frac{1}{6}$ and any $0 < \mu < 1$. By assumption, $\int_M h_1 \, dm \int_M h_2 \, dm > 0$ and hence, applying (41) and (42), we obtain

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

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

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

Now, we consider the case when $\frac{1}{6} < \alpha < \frac{1}{4}$. For such $\alpha$ we have that $\gamma' > 2$ (see [9]). Depending on the value of $\mu$, 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

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

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 $0 < \alpha < \frac{1}{4}$. Thus we have the desired estimate

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

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

In the case $\int_M h_1 \, dm \int_M h_2 \, dm = 0$ the desired result follows immediate from Statement 2 of Proposition 10.2. In particular, again we have $\gamma_2 = \gamma - 2 > 0$.

**Remark 3.** By Remark 2, any observable $h \in \mathcal{G}([M_k])$ (that is $h$ is Hölder continuous and is supported on $M_k$ for some $k > 0$) vanishes in a small neighborhood of the boundary set $\partial B \setminus M$ for $f_M$ (see Remark [4]).

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 $\text{Cor}_n(h, h) = O\left(\frac{1}{n^{\gamma-1}}\right)$. 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 [13], Theorem 4.1 (see also [20], Theorem 3.1).
10.4. The Large Deviation property. We consider the Young tower $Y$ that represents the map $f_M$. The upper bound in (41) allows us to use Theorem 4.2 in [17] to obtain for $0 < \alpha < \frac{1}{4}$ that for all sufficiently small $a > 0$

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

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 $\Delta_i$ in the base $\Delta$ of the tower for $f_M$ with $\tau(\Delta_i) \leq k$. 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 $\pi$ is measure preserving, we obtain $m_M(\pi(\hat{Y}_k)) < 1$. Thus, by Theorem 4.3 in [17], we obtain the lower bound

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

for small $\varepsilon$, 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 $\{\Delta^s_i\}$ of $s$-sets. Denote by $S_n = \{\Delta^s_i : \tau(\Delta^s_i) = n\}$. Since the map $f_M$ is topologically conjugate to the toral automorphism $A$, the number $S_n$ for $f_M$ is equal to the number $S_n$ for $A$. The latter is known to satisfy $S_n \leq e^{hn}$ 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.

Remark 4. Given a surface $M$, the diffeomorphism $f_M$ constructed in Theorem 3.1 is of class $C^{1+\beta}$ where the number $\beta < \frac{2\alpha}{5(1-\alpha)}$ (see [37]) and hence, it can be made arbitrary large if the number $\alpha$ is chosen sufficiently close to 1.

However, to guarantee that $f_M$ admits polynomial lower and upper bounds for the decay of correlations with respect to volume we require the number $\alpha$ to satisfy $0 < \alpha < \frac{1}{4}$ implying that we can only choose $\beta$ from the small interval $(0, 0.8)$. We note that the upper bound 0.8 on $\beta$ can be slightly increased by replacing the upper bound $\frac{1}{4}$ on the number...
\(\alpha\) in (9) with a sharper one and slightly improving the argument in the proof of Proposition 10.2. This however, will still allow only small values of the number \(\beta\).

On the other hand, to guarantee that \(f_M\) admits exponential decay of correlations with respect to MME we do not need to impose any additional requirements on \(\beta\) besides \(\beta < \frac{2\alpha}{5(1-\alpha)}\). Thus \(\beta\) can be made arbitrary large if the number \(\alpha\) is chosen sufficiently close to 1.

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 diffeomorphisms 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] P. Eslami and C. Liverani, Mixing rates for symplectic almost Anosov maps, Nonlinearity, 34 (2021) 3709–3731.
[7] S. Gouëzel, Sharp polynomial estimates for the decay of correlations, Israel J. Math, 139 (2004) 29–65.
[8] R. E. Greene and K. Shiohama, Diffeomorphisms and volume-preserving embeddings of noncompact manifolds, Trans. Amer. Math. Soc., 255 (1979) 403–414.
[9] D. Dolgopyat and Ya. Pesin, Every compact manifold carries a completely hyperbolic diffeomorphism, ETDS 22 (2002), 409–435.
[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).

[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