Annals of Mathematics **185** (2017), 285–310
http://dx.doi.org/10.4007/annals.2017.185.1.5

Lyapunov exponents for random perturbations of some area-preserving maps including the standard map

By Alex Blumenthal, Jinxin Xue, and Lai-Sang Young

Abstract

We consider a large class of 2D area-preserving diffeomorphisms that are not uniformly hyperbolic but have strong hyperbolicity properties on large regions of their phase spaces. A prime example is the *standard map*. Lower bounds for Lyapunov exponents of such systems are very hard to estimate, due to the potential switching of "stable" and "unstable" directions. This paper shows that with the addition of (very) small random perturbations, one obtains with relative ease Lyapunov exponents reflecting the geometry of the deterministic maps.

1. Introduction

A signature of chaotic behavior in dynamical systems is sensitive dependence on initial conditions. Mathematically, this is captured by the *positivity of Lyapunov exponents*: a differentiable map F of a Riemannian manifold M is said to have a positive Lyapunov exponent (LE) at $x \in M$ if $\|dF^n_x\|$ grows exponentially fast with n. This paper is about volume-preserving diffeomorphisms, and we are interested in behaviors that occur on positive Lebesgue measure sets. Though the study of chaotic systems occupies a good part of smooth ergodic theory, the hypothesis of positive LE is extremely difficult to verify when one is handed a concrete map defined by a specific equation — except where the map possesses a continuous family of invariant cones.

An example that has come to symbolize the enormity of the challenge is the *standard map*, a mapping $\Phi = \Phi_L$ of the 2-torus given by

$$\Phi(I,\theta) = (I + L\sin\theta,\ \theta + I + L\sin\theta),$$

where both coordinates I, θ are taken modulo 2π and $L \in \mathbb{R}$ is a parameter. For $L \gg 1$, the map Φ_L has strong expansion and contraction, their directions separated by clearly defined invariant cones on most of the phase space

The research of J. X. was supported by NSF Grant DMS-1500897.
The research of L.-S. Y was supported in part by NSF Grant DMS-1363161.

— except on two narrow strips near $\theta = \pm\pi/2$ on which vectors are rotated violating cone preservation. As the areas of these "critical regions" tend to zero as $L \to \infty$, one might expect LE to be positive, but this problem has remained unresolved: *no one has been able to prove, or disprove, the positivity of Lyapunov exponents for* Φ_L *for any one* L, *however large*, in spite of considerable effort by leading researchers. The best result known [Gor12] is that the LE of Φ_L is positive on sets of Hausdorff dimension 2 (which are very far from having positive Lebesgue measure). The presence of elliptic islands, which has been shown for a residual set of parameters [Dua94], [Dua08], confirms that the obstructions to proving the positivity of LE are real.

In this paper, we propose that this problem can be more tractable if one accepts that dynamical systems are inherently noisy. We show, for a class of 2D maps F that includes the standard map, that by adding a very small, independent random perturbation at each step, the resulting maps have a positive LE that correctly reflects the rate of expansion of F — provided that F has sufficiently large expansion to begin with. More precisely, if $\|dF\| \sim L$, $L \gg 1$, on a large portion of the phase space, then random perturbations of size $O(e^{-L^{2-\varepsilon}})$ are sufficient for guaranteeing an LE $\sim \log L$.

Our proofs for these results, which are very short compared to previous works on establishing nonuniform hyperbolicity for deterministic maps (e.g., [Jak81], [BC85], [BC91], [WY01], [WY06], [WY08]), are based on the following idea: We view the random process as a Markov chain on the projective bundle of the manifold on which the random maps act and represent LE as an integral. Decomposing this integral into a "good part" and a "bad part," we estimate the first leveraging the strong hyperbolicity of the unperturbed map and obtain a lower bound for the second provided the stationary measure is not overly concentrated in certain "bad regions." We then use a large enough random perturbation to make sure that the stationary measure is sufficiently diffused.

We expect that with more work, this method can be extended both to higher dimensions and to situations where conditions on the unperturbed map are relaxed.

Relation to existing results. Closest to the present work are the unpublished results of Carleson and Spencer [CS], [Spe], who showed for very carefully selected parameters $L \gg 1$ of the standard map that LE are positive when the map's derivatives are randomly perturbed. For comparison, our first result applies to all $L \gg 1$ with a slightly larger perturbation than in [CS], and our second result assumes additionally a finite condition on a finite set; we avoid the rather delicate parameter selection by perturbing the maps themselves, not just their derivatives.

Parameter selections similar to those in [CS] were used — without random perturbations — to prove the positivity of LE for the Hénon maps [BC91], quasi-periodic cocycles [You97], and rank-one attractors [WY08], building on earlier techniques in 1D; see, e.g., [Jak81], [Ryc88], [BC85], [WY06]. See also [SS15], which estimates LE from below for Schrödinger cocycles over the standard map. Relying on random perturbation alone — without parameter deletion — are [LS12], which contains results analogous to ours in 1D, and [LSSW03], which applied random rotations to twist maps. We mention also [BC14], which uses hyperbolic toral automorphisms in lieu of random perturbations.

Farther from our setting, the literature on LE is vast. Instead of endeavoring to give reasonable citation of individual papers, let us mention several *categories* of results in the literature that have attracted much attention, together with a small sample of results in each. Furstenberg's work [Fur63] in the early 60's initiated extensive research on criteria for the LE of random matrix products to be distinct (see, e.g., [GM89], [GR86], [Vir80]). Similar ideas were exploited to study LE of cocycles over hyperbolic and partially hyperbolic systems (see, e.g., [BV04], [BGMV03]), with a generalization to deterministic maps [AV10]. Unlike the results in the first two paragraphs, these results do not give quantitative estimates; they assert only that LE are simple, or nonzero.

We mention as well the formula of Herman [Her83], [Kni92] and the related work [AB02], which use subharmonicity to estimate Lyapunov exponents, and the substantial body of work on 1D Schrödinger operators (e.g., [Kot84], [Bou13], [Pui04], [AJ09]). We also note the C^1 genericity of zero Lyapunov exponents of volume-preserving surface diffeomorphisms away from Anosov [Boc02] and its higher-dimensional analogue [BV05]. Finally, we acknowledge results on the continuity or stability of LE, as in, e.g., [Rue79], [Hen84], [Kif82], [BNV10], [LY91].

This paper is organized as follows. We first state and prove two results in a relatively simple setting: Theorem 1, which contains the core idea of this paper, is proved in Sections 3 and 4, while Theorem 2, which shows how perturbation size can be decreased if some mild conditions are assumed, is proved in Section 5. We also describe a slightly more general setting, which includes the standard map, and observe in Section 6 that the proofs given earlier in fact apply, exactly as written, to this broader setting.

2. Results and remarks

2.1. *Statement of results.* We let $\psi : \mathbb{S}^1 \to \mathbb{R}$ be a C^3 function for which the following hold:

(H1) $C'_\psi = \{\hat{x} \in \mathbb{S}^1 : \psi'(\hat{x}) = 0\}$ and $C''_\psi = \{\hat{z} \in \mathbb{S}^1 : \psi''(\hat{z}) = 0\}$ have finite cardinality;

(H2) $\min_{\hat{x} \in C'_\psi} |\psi''(\hat{x})| > 0$ and $\min_{\hat{z} \in C''_\psi} |\psi'''(\hat{z})| > 0$.

For $L > 1$ and $a \in [0, 1)$, we define

$$f = f_{L,a} : \mathbb{S}^1 \to \mathbb{R} \qquad \text{by} \qquad f(x) = L\psi(x) + a.$$

Let $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$ be the 2-torus. The deterministic map to be perturbed is

$$F = F_{L,a} : \mathbb{T}^2 \to \mathbb{T}^2, \qquad \text{where} \qquad F(x, y) = \begin{pmatrix} f(x) - y \pmod 1 \\ x \end{pmatrix}. \tag{1}$$

We have abused notation slightly in equation (1): We have made sense of $f(x) - y$ by viewing $y \in \mathbb{S}^1$ as belonging in $[0, 1)$, and have written "$z \pmod 1$" instead of $\pi(z)$ where $\pi : \mathbb{R} \to \mathbb{S}^1 \cong \mathbb{R}/\mathbb{Z}$ is the usual projection. Observe that F is an area-preserving diffeomorphism of $\mathbb{T}^2$.

We consider compositions of random maps

$$F^n_{\underline{\omega}} = F_{\omega_n} \circ \cdots \circ F_{\omega_1} \qquad \text{for} \quad n = 1, 2, \ldots,$$

where

$$F_\omega = F \circ S_\omega, \qquad S_\omega(x, y) = (x + \omega \pmod 1, y),$$

and the sequence $\underline{\omega} = (\omega_1, \omega_2, \ldots)$ is chosen independent and identically distributed with respect to the uniform distribution ν^ϵ on $[-\epsilon, \epsilon]$ for some $\epsilon > 0$. Thus our sample space can be written as $\Omega = [-\epsilon, \epsilon]^{\mathbb{N}}$, equipped with the probability $\mathbb{P} = \left(\nu^\epsilon\right)^{\mathbb{N}}$.

Throughout, we let Leb denote Lebesgue measure on $\mathbb{T}^2$.

THEOREM 1. *Assume ψ obeys* (H1) *and* (H2), *and fix $a \in [0, 1)$. Then*

(a) *for every $L > 0$ and $\epsilon > 0$,*

$$\lambda_1^\epsilon = \lim_{n \to \infty} \frac{1}{n} \log \|(dF^n_{\underline{\omega}})_{(x,y)}\| \tag{2}$$

exists and is independent of $(x, y, \underline{\omega})$ for every $(x, y) \in \mathbb{T}^2$ and $\mathbb{P}$-a.e. $\underline{\omega} \in \Omega$;

(b) *given $\alpha, \beta \in (0, 1)$, there is a constant $C = C_{\alpha,\beta} > 0$ such that for all L, ϵ where L is sufficiently large (depending on ψ, α, β) and $\epsilon \geq L^{-CL^{1-\beta}}$, we have*

$$\lambda_1^\epsilon \geq \alpha \log L.$$

Theorem 1 assumes no information whatsoever on dynamical properties of F beyond its definition in equation (1). Our next result shows, under some minimal, easily checkable, condition on the first iterates of F, that the bound above on λ_1^ϵ continues to hold for a significantly smaller ϵ. Let $\mathcal{N}_c(C'_\psi)$ denote the c-neighborhood of C'_ψ in $\mathbb{S}^1$. We formulate the following condition on $f = f_{L,a}$:

(H3)(c) For any $\hat{x}, \hat{x}' \in C'_\psi$, we have that $f\hat{x} - \hat{x}' \pmod 1 \notin \mathcal{N}_c(C'_\psi)$.

Observe that for L large, the set of a for which (H3)(c) is satisfied tends to 1 as $c \to 0$.

THEOREM 2. *Let ψ be as above, and fix an arbitrary $c_0 > 0$. Then given $\alpha, \beta \in (0,1)$, there is a constant $C = C_{\alpha,\beta} > 0$ such that for all L, a, ϵ where*

— *L is sufficiently large (depending on ψ, c_0, α, β);*
— *$a \in [0,1)$ is chosen so that $f = f_{L,a}$ satisfies (H3)(c_0); and*
— *$\epsilon \geq L^{-CL^{2-\beta}}$,*

then we have

$$\lambda_1^\epsilon \geq \alpha \log L.$$

A slight extension. Let $\psi : \mathbb{S}^1 \to \mathbb{R}$ be as above. For $L > 0$ and $a \in [0,1)$, we write $f_0 = f_{\psi,L,a} = L\psi + a$, and for $\varepsilon > 0$, we define

$$\mathcal{U}_{\varepsilon,L}(f_0) = \{f : \mathbb{S}^1 \to \mathbb{R} \text{ such that } \|f - f_0\|_{C^3} < L\varepsilon\}.$$

We let C'_f and C''_f denote the zeros of f' and f''. Below, (H3)(c) is to be read with C'_f in the place of C'_ψ. We write $F_f(x,y) = (f(x) - y \pmod 1, x)$ for $f \in \mathcal{U}_{\varepsilon,L}(f_0)$.

THEOREM 3. *Let $\psi : \mathbb{S}^1 \to \mathbb{R}$ satisfy* (H1) *and* (H2) *as before. For $a \in [0,1)$ and $L > 1$, let f_0 be as defined above. Then there exists $\varepsilon > 0$ sufficiently small so that*

(1) *Theorems 1 and 2 hold for $F = F_f$ for all $L > 0$ sufficiently large and $f \in \mathcal{U}_{\varepsilon,L}(f_0)$;*
(2) *L depends only on ψ as before but a in Theorem 2 depends on f.*

The Chirikov standard map is defined as follows: a parameter $L > 0$ is fixed, and the map $(I, \theta) \mapsto (\bar{I}, \bar{\theta})$, sending $[0, 2\pi)^2$ into itself, is defined by

$$\bar{I} = I + 2\pi L \sin\theta,$$
$$\bar{\theta} = \theta + \bar{I} = \theta + I + 2\pi L \sin\theta,$$

where both coordinates I, θ are taken modulo 2π.

COROLLARY 4 (The standard map). *Let L be sufficiently large. Then,*

- *Theorem 1 holds for the standard map;*
- *if additionally the map $f(x) = L\sin(2\pi x) + 2x$ satisfies* (H3)(c) *for some $c > 0$, then Theorem 2 holds for the standard map for this value of L.*

Theorem 3 and Corollary 4 are proved in Section 6. *All discussions prior to Section 6 pertain to the setting described at the beginning of this section.*

2.2. *Remarks.*

Remark 1: *Uniform hyperbolicity on large but noninvariant regions of the phase space.* An important property of the deterministic map F is that cone fields can be defined on all of $\mathbb{T}^2$ in such a way that they are preserved by

$dF_{(x,y)}$ for (x,y) in a large but noninvariant region in $\mathbb{T}^2$. For example, let $C_{\frac{1}{5}} = \{v = (v_x, v_y) : |v_y/v_x| \leq \frac{1}{5}\}$. Then for $(x,y) \notin \{|f'| < 10\}$, which by (H1) and (H2) is comprised of a finite number of very narrow vertical strips in $\mathbb{T}^2$ for L large, one checks easily that $dF_{(x,y)}$ maps $C_{\frac{1}{5}}$ into $C_{\frac{1}{5}}$, and expands vectors in these cones uniformly. It is just as easy to see that this cone invariance property cannot be extended across the strips in $\{|f'| < 10\}$ and that F is not uniformly hyperbolic.

These "bad regions" where the invariant cone property fails shrink in size as L increases. More precisely, let $K_1 > 1$ be such that $|\psi'(x)| \geq K_1^{-1} d(x, C'_\psi)$; that such a K_1 exists follows from (H1) and (H2) in Section 2.1. It is easy to check that for any $\eta \in (0,1)$,

$$d(x, C'_\psi) \geq \frac{K_1}{L^{1-\eta}} \qquad \Longrightarrow \qquad |f'(x)| \geq L^\eta,$$

and this strong expansion in the x-direction is reflected in $dF_{(x,y)}$ for any y.

We must stress, however, that regardless of how small these "bad regions" are, the positivity of Lyapunov exponents is not guaranteed for the deterministic map F — except for the Lebesgue measure zero set of orbits that never venture into these regions. In general, tangent vectors that have expanded in the good regions can be rotated into contracting directions when the orbit visits a bad region. This is how elliptic islands are formed.

Remark 2: *Interpretation of condition* (H3). We have seen that visiting neighborhoods of $V_{\hat{x}} := \{x = \hat{x}\}$ for $\hat{x} \in C'_\psi$ can lead to a loss in hyperbolicity, yet at the same time it is unavoidable that the "typical" orbit will visit these "bad regions." Intuitively, it is logical to expect the situation to improve if we do not permit orbits to visit these bad regions two iterates in a row — except that such a condition is impossible to arrange: since $F(V_{\hat{x}'}) = \{y = \hat{x}'\}$, it follows that $F(V_{\hat{x}'})$ meets $V_{\hat{x}}$ for every $\hat{x}, \hat{x}' \in C'_\psi$. In Theorem 2, we assert that in the case of random maps, to reduce the size of ϵ it suffices to impose the condition that no orbit can be in $C'_\psi \times \mathbb{S}^1$ for *three consecutive iterates*. That is to say, suppose $F(x_i, y_i) = (x_{i+1}, y_{i+1}),\ i = 1, 2, \ldots$. If $x_i, x_{i+1} \in C'_\psi$, then x_{i+2} must stay away from C'_ψ. This is a rephrasing of (H3). Such a condition is both realizable and checkable, as it involves only a finite number of iterates for a finite set of points.

Remark 3: *Potential improvements*. Condition (H3) suggests that one may be able to shrink ϵ further by imposing similar conditions on one or two more iterates of F. Such conditions will cause the combinatorics in Section 5 to be more involved, and since our ϵ, which is $\sim L^{-L^{2-\beta}}$, is already extremely small for large L, we will not pursue these possibilities here.

3. **Preliminaries**

The results of this section apply to all $L, \epsilon > 0$ unless otherwise stated.

3.1. *Relevant Markov chains.* Our random maps system $\{F^n_{\underline{\omega}}\}_{n\geq 1}$ can be seen as a time-homogeneous Markov chain $\mathcal{X} := \{(x_n, y_n)\}$ given by

$$(x_n, y_n) = F^n_{\underline{\omega}}(x_0, y_0) = F_{\omega_n}(x_{n-1}, y_{n-1}).$$

That is to say, for fixed ϵ, the transition probability starting from $(x, y) \in \mathbb{T}^2$ is

$$P((x, y), A) = P^\epsilon((x, y), A) = \nu^\epsilon\{\omega \in [-\epsilon, \epsilon] : F_\omega(x, y) \in A\}$$

for Borel $A \subset \mathbb{T}^2$. We write $P^{(k)}((x, y), \cdot)$ (or $P^{(k)}_{(x,y)}$) for the corresponding k-step transition probability. It is easy to see that for this chain, Lebesgue measure is *stationary*, meaning for any Borel set $A \subset \mathbb{T}^2$,

$$\mathrm{Leb}(A) = \int P((x, y), A)\, d\mathrm{Leb}(x, y).$$

Ergodicity of this chain is easy and we dispose of it quickly.

Lemma 5. *Lebesgue measure is ergodic.*

Proof. For any $(x, y) \in \mathbb{T}^2$ and $\omega_1, \omega_2 \in [-\epsilon, \epsilon]$,

$$F_{\omega_2} \circ F_{\omega_1}(x, y) = F \circ F \circ S'_{\omega_1, -\omega_2}(x, y), \tag{3}$$

where $S'_{\omega,\omega'}(x, y) = \big(x + \omega\,(\mathrm{mod}\ 1), y + \omega'(\mathrm{mod}\ 1)\big)$. That is to say, $P^{(2)}_{(x,y)}$ is supported on the set $F^2([x - \epsilon, x + \epsilon] \times [y - \epsilon, y + \epsilon])$, on which it is equivalent to Lebesgue measure. From this one deduces immediately that

(i) every ergodic stationary measure of $\mathcal{X} = \{(x_n, y_n)\}$ has a density, and
(ii) all nearby points in $\mathbb{T}^2$ are in the same ergodic component. Thus there can be at most one ergodic component. □

Part (a) of Theorem 1 follows immediately from Lemma 5 together with the Multiplicative Ergodic Theorem for random maps.

Next we introduce a Markov chain $\hat{\mathcal{X}}$ on $\mathbb{P}\mathbb{T}^2$, the projective bundle over $\mathbb{T}^2$. Associating $\theta \in \mathbb{P}^1 \cong [0, \pi)$ with the unit vector $u_\theta = (\cos\theta, \sin\theta)$, F_ω induces a mapping $\hat{F}_\omega : \mathbb{P}\mathbb{T}^2 \to \mathbb{P}\mathbb{T}^2$ defined by

$$\hat{F}_\omega(x, y, \theta) = (F_\omega(x, y), \theta'), \qquad \text{where} \qquad u_{\theta'} = \pm\frac{(dF_\omega)_{(x,y)} u_\theta}{\|(dF_\omega)_{(x,y)} u_\theta\|}.$$

Here $\pm$ is chosen to ensure that $\theta' \in [0, \pi)$. The Markov chain $\hat{\mathcal{X}} := \{(x_n, y_n, \theta_n)\}$ is then defined by

$$(x_n, y_n, \theta_n) = \hat{F}_{\omega_n}(x_{n-1}, y_{n-1}, \theta_{n-1}).$$

We write $\hat{P}$ for its transition operator, $\hat{P}^{(n)}$ for the n-step transition transition operator, and use Leb to denote also Lebesgue measure on $\mathbb{P}\mathbb{T}^2$.

For any stationary probability measure $\hat{\mu}$ of the Markov chain (x_n, y_n, θ_n), define

$$\lambda(\hat{\mu}) = \int \log \|(dF_\omega)_{(x,y)} u_\theta\| \; d\hat{\mu}(x,y,\theta) \; d\nu^\epsilon(\omega).$$

LEMMA 6. *For any stationary probability measure $\hat{\mu}$ of the Markov chain $\hat{\mathcal{X}}$, we have*

$$\lambda_1^\epsilon \geq \lambda(\hat{\mu}).$$

Proof. By the additivity of the cocycle $(x, y, \theta) \mapsto \log \|d(F_\omega)_{(x,y)} u_\theta\|$, we have, for any $n \in \mathbb{N}$,

$$\begin{aligned}\lambda(\hat{\mu}) &= \int \frac{1}{n} \log \|(dF^n_{\underline{\omega}})_{(x,y)} u_\theta\| \; d\hat{\mu}(x,y,\theta) \; d(\nu^\epsilon)^n(\underline{\omega}) \\ &\leq \int \frac{1}{n} \log \|(dF^n_{\underline{\omega}})_{(x,y)}\| \; d\mathrm{Leb}(x,y) \; d(\nu^\epsilon)^n(\underline{\omega}).\end{aligned}$$

That $\hat{\mu}$ projects to Lebesgue measure on $\mathbb{T}^2$ is used in passing from the first to the second line, and the latter converges to λ_1^ϵ as $n \to \infty$ by the Multiplicative Ergodic Theorem. □

Thus to prove part (b) of Theorem 1, it suffices to prove that $\lambda(\hat{\mu}) \geq \alpha \log L$ for some $\hat{\mu}$. Uniqueness of $\hat{\mu}$ is not required. On the other hand, once we have shown that $\lambda_1^\epsilon > 0$, it will follow that there can be at most one $\hat{\mu}$ with $\lambda(\hat{\mu}) > 0$. Details are left to the reader.

We remark also that while Theorems 1–3 hold for arbitrarily large values of ϵ, we will treat only the case $\epsilon \leq \frac{1}{2}$, leaving the very minor modifications needed for the $\epsilon > \frac{1}{2}$ case to the reader.

Finally, we will omit from time to time the notation "(mod 1)" when the meaning is obvious, e.g., instead of the technically correct but cumbersome $f(x + \omega \pmod 1) - y \pmod 1$, we will write $f(x + \omega) - y$.

3.2. *A 3-step transition.* In anticipation for later use, we compute here the transition probabilities $\hat{P}^{(3)}((x, y, \theta), \cdot)$, also denoted $\hat{P}^{(3)}_{(x,y,\theta)}$. Let $(x_0, y_0, \theta_0) \in \mathbb{PT}^2$ be fixed. We define

$$H = H^{(3)}_{(x_0,y_0,\theta_0)} : [-\epsilon, \epsilon]^3 \to \mathbb{PT}^2$$

by

$$H(\omega_1, \omega_2, \omega_3) = \hat{F}_{\omega_3} \circ \hat{F}_{\omega_2} \circ \hat{F}_{\omega_1}(x_0, y_0, \theta_0).$$

Then $\hat{P}^{(3)}_{(x_0,y_0,\theta_0)} = H_*((\nu^\epsilon)^3)$, the pushforward of $(\nu^\epsilon)^3$ on $[-\epsilon, \epsilon]^3$ by H. Write $(x_i, y_i, \theta_i) = \hat{F}_{\omega_i}(x_{i-1}, y_{i-1}, \theta_{i-1})$, $i = 1, 2, 3$.

LEMMA 7. *Let $\epsilon \in (0, \frac{1}{2}]$. Let $(x_0, y_0, \theta_0) \in \mathbb{PT}^2$ be fixed, and let $H = H^{(3)}_{(x_0,y_0,\theta_0)}$ be as above. Then*

(i) *we have*

$$\det dH(\omega_1,\omega_2,\omega_3) = \sin^2(\theta_3)\tan^2(\theta_2)\tan^2(\theta_1) f''(x_0+\omega_1); \tag{4}$$

(ii) *assuming* $\theta_0 \neq \pi/2$, *we have that* $\det dH \neq 0$ *on* V *where* $V \subset [-\epsilon,\epsilon]^3$ *is an open and dense set having full Lebesgue measure in* $[-\epsilon,\epsilon]^3$;

(iii) H *is at most* $\#(C''_\psi)$*-to-one, i.e., no point in* $\mathbb{PT}^2$ *has more than* $\#C''_\psi$ *preimages.*

Proof of Lemma 7. The projectivized map $\hat{F}_\omega$ can be written as

$$\hat{F}_\omega(x,y,\theta) = \left(f(x+\omega) - y, x+\omega, \ \arctan \frac{1}{f'(x+\omega) - \tan\theta} \right), \tag{5}$$

where arctan is chosen to take values in $[0,\pi)$.

(i) It is convenient to write $k_i = \tan\theta_i$, so that $k_{i+1} = (f'(y_{i+1}) - k_i)^{-1}$. Note as well that $x_{i+1} = f(y_{i+1}) - y_i$. Then

(6)

$$\begin{aligned} dx_3 \wedge dy_3 \wedge d\theta_3 & \\ &= (f'(y_3)dy_3 - dy_2) \wedge dy_3 \wedge \left(\frac{\partial\theta_3}{\partial y_3} dy_3 + \frac{\partial \theta_3}{\partial k_2} dk_2 \right) \\ &= -dy_2 \wedge dy_3 \wedge \left(\frac{\partial\theta_3}{\partial k_2} dk_2 \right) \\ &= -dy_2 \wedge (d\omega_3 + f'(y_2)dy_2 - dy_1) \wedge \left(\frac{\partial\theta_3}{\partial k_2} \right) \left(\frac{\partial k_2}{\partial y_2} dy_2 + \frac{\partial k_2}{\partial k_1} dk_1 \right) \\ &= -dy_2 \wedge d(\omega_3 - y_1) \wedge \left(\frac{\partial\theta_3}{\partial k_2} \frac{\partial k_2}{\partial k_1} dk_1 \right) \\ &= -(d\omega_2 + f'(y_1)dy_1) \wedge d(\omega_3 - y_1) \wedge \left(\frac{\partial\theta_3}{\partial k_2} \frac{\partial k_2}{\partial k_1} \frac{\partial k_1}{\partial y_1} dy_1 \right) \\ &= -d\omega_2 \wedge d\omega_3 \wedge \left(\frac{\partial\theta_3}{\partial k_2} \frac{\partial k_2}{\partial k_1} \frac{\partial k_1}{\partial y_1} d\omega_1 \right). \end{aligned}$$

It remains to compute the parenthetical term. The second two partial derivatives are straightforward. The first partial derivative is computed by taking the partial derivative of the formula $\cot\theta_3 = f'(y_3) - k_2$ with respect to k_2 on both sides. As a result, we obtain

$$\frac{\partial\theta_3}{\partial k_2} \frac{\partial k_2}{\partial k_1} \frac{\partial k_1}{\partial y_1} = -\sin^2\theta_3 \tan^2\theta_2 \tan^2\theta_1 f''(x_0+\omega_1).$$

(ii) For $x \in [0,1)$ and $\theta \in [0,\pi) \setminus \{\pi/2\}$, define $U(x,\theta) = \{\omega \in [-\epsilon,\epsilon] : f'(x+\omega) - \tan\theta \neq 0\}$. Note that $U(x,\theta)$ has full Lebesgue measure in $[-\epsilon,\epsilon]$ by (H1). We define

$$\begin{aligned} V = \{(\omega_1,\omega_2,\omega_3) \in [-\epsilon,\epsilon]^3 : & \omega_1 \in U(x_0,\theta_0), \omega_2 \in U(x_1,\theta_1), \\ & \omega_3 \in U(x_2,\theta_2), \text{ and } f''(x_0+\omega_1) \neq 0\}. \end{aligned}$$

By (H1) and Fubini's Theorem, V has full measure in $[-\epsilon, \epsilon]^3$, and it is clearly open and dense. To show $\det dH \neq 0$, we need $\theta_i \neq 0$ for $i = 1, 2, 3$ on V. This follows from the fact that for $\theta_{i-1} \neq \pi/2$, if $\omega_i \in U(x_{i-1}, \theta_{i-1})$, then $\theta_i \neq 0, \pi/2$.

(iii) Given (x_3, y_3, θ_3), we solve for $(\omega_1, \omega_2, \omega_3)$ so that $H(\omega_1, \omega_2, \omega_3) = (x_3, y_3, \theta_3)$. Letting (x_i, y_i, θ_i), $i = 1, 2$, be the intermediate images, we note that y_2 is uniquely determined by $x_3 = f(y_3) - y_2$, and θ_2 is determined by $\cot\theta_3 = f'(y_3) - \tan\theta_2$, as is θ_1 once θ_2 and y_2 are fixed. This in turn determines $f'(x_0 + \omega_1)$, but here uniqueness of solutions breaks down.

Let $\omega_1^{(i)} \in [-\epsilon, \epsilon], i = 1, \ldots, n$, give the required value of $f'(x_0 + \omega_1^{(i)})$. We observe that each $\omega_1^{(i)}$ determines uniquely $y_1^{(i)} = x_0 + \omega_1^{(i)}$, $x_1^{(i)} = f(y_1^{(i)}) - y_0$, $\omega_2^{(i)} = y_2 - x_1^{(i)}$, $x_2^{(i)} = f(y_2) - y_1^{(i)}$, and finally $\omega_3^{(i)} = y_3 - x_2^{(i)}$. Thus the number of H-preimages of any one point in $\mathbb{PT}^2$ cannot exceed n. Finally, we have $n \leq 2$ for ϵ small and $n \leq \#(C''_\psi)$ for ϵ as large as $\frac{1}{2}$. $\square$

COROLLARY 8. *For any stationary probability $\hat{\mu}$ of $\hat{\mathcal{X}}$, we have $\hat{\mu}(\mathbb{T}^2 \times \{\pi/2\}) = 0$, and for any (x_0, y_0, θ_0) with $\theta_0 \neq \pi/2$ and any $(x_3, y_3, \theta_3) \in \mathbb{PT}^2$, the density of $\hat{P}^{(3)}_{(x_0,y_0,\theta_0)}$ at (x_3, y_3, θ_3) is given by*

$$\frac{1}{(2\epsilon)^3}\left(\sum_{\omega_1 \in \mathcal{E}(x_3,y_3,\theta_3)} \frac{1}{|f''(x_0+\omega_1)|}\right)\frac{1}{\rho(x_3,y_3,\theta_3)}, \tag{7}$$

where

$$\mathcal{E}(x_3, y_3, \theta_3) = \{\omega_1 : \exists \omega_2, \omega_3 \text{ such that } H(\omega_1, \omega_2, \omega_3) = (x_3, y_3, \theta_3)\}$$

and

$$\rho(x, y, \theta) = \sin^2(\theta)\left[f'(f(y) - x)(f'(y) - \cot\theta) - 1\right]^2.$$

Proof. To show $\hat{\mu}(\mathbb{T} \times \{\pi/2\}) = 0$, it suffices to show that given any $x \in [0, 1)$ and any $\theta \in [0, \pi)$, $\nu^\epsilon\{\omega \in [-\epsilon, \epsilon] : f'(x + \omega) = \tan\theta\} = 0$, and that is true because C''_ψ is finite by (H1). The formula in (7) follows immediately from the proof of Lemma 7, upon expressing $\tan^2(\theta_2)\tan^2(\theta_1)$ in terms of (x_3, y_3, θ_3) as was done in the proof of Lemma 7(iii). $\square$

4. Proof of Theorem 1

The idea of our proof is as follows: Let $\hat{\mu}$ be any stationary probability of the Markov chain $\hat{\mathcal{X}}$. To estimate the integral in $\lambda(\hat{\mu})$, we need to know the distribution of $\hat{\mu}$ in the θ-direction. Given that the maps F_ω are strongly uniformly hyperbolic on a large part of the phase space with expanding directions well aligned with the x-axis (see Remark 1), one can expect that under $dF^N_{\underline{\omega}}$ for large N, $\hat{\mu}$ will be pushed toward a neighborhood of $\{\theta = 0\}$ on much of $\mathbb{T}^2$, and that is consistent with $\lambda_1^\epsilon \approx \log L$. This reasoning, however, is predicated on $\hat{\mu}$ not being concentrated, or stuck, on very small sets far away from $\{\theta \approx 0\}$, a scenario not immediately ruled out as the densities of transition probabilities are not bounded.

We address this issue directly by proving in Lemma 9 an *a priori* bound on the extent to which $\hat{\mu}$-measure can be concentrated on (arbitrary) small sets. This bound is used in Lemma 10 to estimate the $\hat{\mu}$-measure of the set in $\mathbb{PT}^2$ not yet attracted to $\{\theta = 0\}$ in N steps. The rest of the proof consists of checking that these bounds are adequate for our purposes. In the rest of the proof, let $\hat{\mu}$ be an arbitrary invariant probability measure of $\hat{\mathcal{X}}$.

LEMMA 9. *Let $A \subset \{\theta \in [\pi/4, 3\pi/4]\}$ be a Borel subset of $\mathbb{PT}^2$. Then for L large enough,*

$$\hat{\mu}(A) \leq \frac{\hat{C}}{L^{\frac{1}{4}}}\left(1 + \frac{1}{\epsilon^3 L^2}\operatorname{Leb}(A)\right) \tag{8}$$

for all $\epsilon \in (0, \frac{1}{2}]$, where $\hat{C} > 0$ is a constant independent of L, ϵ or A.

Proof. By the stationarity of $\hat{\mu}$, we have, for every Borel set $A \subset \mathbb{PT}^2$,

$$\hat{\mu}(A) = \int_{\mathbb{PT}^2} \hat{P}^{(3)}_{(x_0,y_0,\theta_0)}(A)\, d\hat{\mu}(x_0, y_0, \theta_0). \tag{9}$$

Our plan is to decompose this integral into a main term and "error terms," depending on properties of the density of $\hat{P}^{(3)}_{(x_0,y_0,\theta_0)}$. The decomposition is slightly different depending on whether $\epsilon \leq L^{-\frac{1}{2}}$ or $\geq L^{-\frac{1}{2}}$.

The case $\epsilon \leq L^{-\frac{1}{2}}$. Let $K_2 \geq 1$ be such that $|\psi''(x)| \geq K_2^{-1} d(x, C''_\psi)$; such a K_2 exists by (H1) and (H2). Define $B'' = \{(x, y) : d(x, C''_\psi) \leq 2K_2 L^{-1/2}\}$. Then splitting the right side of (9) into

$$\int_{B''\times[0,\pi)} + \int_{\mathbb{PT}^2 \setminus (B''\times[0,\pi))}, \tag{10}$$

we see that the first integral is $\leq \operatorname{Leb}(B'') \leq \frac{4K_2 M_2}{\sqrt{L}}$, where $M_2 = \#C''_\psi$. As for $(x_0, y_0) \notin B''$, since $|f''(x_0 + \omega)| \geq L^{\frac{1}{2}}$, the density of $\hat{P}^{(3)}_{(x_0,y_0,\theta_0)}$ is $\leq [(2\epsilon)^3 M_2^{-1} L^{\frac{1}{2}} \rho]^{-1}$ by Corollary 8.

To bound the second integral in (10), we need to consider the zeros of ρ. As $A \subset \mathbb{T}^2 \times [\pi/4, 3\pi/4]$, we have $\sin^2(\theta_3) \geq 1/2$. The form of ρ in Corollary 8 prompts us to decompose A into

$$A = (A \cap \hat{G}) \cup (A \setminus \hat{G}),$$

where $\hat{G} = G \times [0, \pi)$ and

$$G = \{(x, y) : d(y, C'_\psi) > K_1 L^{-\frac{1}{2}},\ d(f(y) - x, C'_\psi) \geq K_1 L^{-\frac{1}{2}}\}.$$

Then on $\hat{G} \cap A$, we have $\rho \geq \frac{1}{2}(\frac{1}{2}L)^2$ for L sufficiently large. This gives

$$\int_{\mathbb{PT}^2 \setminus (B'' \times [0,\pi))} \hat{P}^{(3)}_{(x_0,y_0,\theta_0)}(A \cap \hat{G})\, d\hat{\mu} \leq \frac{C}{\epsilon^3 \sqrt{L}} \frac{1}{L^2} \operatorname{Leb}(A).$$

Finally, by the invariance of $\hat{\mu}$,

$$\int_{\mathbb{PT}^2 \setminus (B'' \times [0,\pi))} \hat{P}^{(3)}_{(x_0,y_0,\theta_0)}(A \setminus \hat{G})\, d\hat{\mu} \ \leq \ \hat{\mu}(A \setminus \hat{G}) = \operatorname{Leb}(\mathbb{T}^2 \setminus G).$$

We claim that this is $\lesssim L^{-\frac{1}{2}}$. Clearly, $\operatorname{Leb}\{d(y, C'_\psi) \leq K_1 L^{-\frac{1}{2}}\}) \approx L^{-\frac{1}{2}}$. As for the second condition,

$$\{y : f(y) \in (z - K_1 L^{-\frac{1}{2}}, z + K_1 L^{-\frac{1}{2}})\} = \{y : \psi(y) \in (z' - K_1 L^{-\frac{3}{2}}, z' + K_1 L^{-\frac{3}{2}})\},$$

which in the worst case has Lebesgue measure $\lesssim L^{-\frac{3}{4}}$ by (H1) and (H2).

The case $\epsilon \geq L^{-\frac{1}{2}}$. Here we let $\widetilde{B}'' = \{(x,y) : d(x, C''_\psi) \leq K_2 L^{-3/4}\}$, and we decompose the right side of (9) into

$$\int \left(\hat{P}^{(3)}_{(x_0,y_0,\theta_0)}\right)_1 (A)\, d\hat{\mu} \ + \ \int \left(\hat{P}^{(3)}_{(x_0,y_0,\theta_0)}\right)_2 (A)\, d\hat{\mu}$$

where, in the notation in Section 3.2,

$$\left(\hat{P}^{(3)}_{(x_0,y_0,\theta_0)}\right)_1 = H_* \left((\nu^\epsilon)^3|_{\{x_0+\omega_1 \in \widetilde{B}''\}}\right)$$

and

$$\left(\hat{P}^{(3)}_{(x_0,y_0,\theta_0)}\right)_2 = H_* \left((\nu^\epsilon)^3|_{\{x_0+\omega_1 \notin \widetilde{B}''\}}\right).$$

Then the first integral is bounded above by

$$\sup_{x_0 \in \mathbb{S}^1} \nu^\epsilon\{\omega_1 \in \widetilde{B}'' - x_0\} \lesssim \epsilon^{-1} \operatorname{Leb}(\widetilde{B}'') \leq \mathrm{Const} \cdot L^{-1/4},$$

while the density of $(\hat{P}^{(3)}_{(x_0,y_0,\theta_0)})_2$ is $\leq [(2\epsilon)^3 M_2^{-1} L^{1/4} \cdot \rho(x_3,y_3,\theta_3)]^{-1}$. The second integral is treated as in the case of $\epsilon \leq L^{-\frac{1}{2}}$. $\square$

As discussed above, we now proceed to estimate the Lebesgue measure of the set that remains far away from $\{\theta = 0\}$ after N steps, where N is arbitrary for now. For fixed $\underline{\omega} = (\omega_1, \ldots, \omega_N)$, we write $(x_i, y_i) = F^i_{\underline{\omega}}(x_0, y_0)$ for $1 \leq i \leq N$ and define $G_N = G_N(\omega_1, \ldots, \omega_N)$ by

$$G_N = \{(x_0, y_0) \in \mathbb{T}^2 : d(x_i + \omega_{i+1}, C'_\psi) \geq K_1 L^{-1+\beta} \text{ for all } 0 \leq i \leq N-1\}.$$

We remark that for $(x_0, y_0) \in G_N$, the orbit $F^i_{\underline{\omega}}(x_0, y_0), i \leq N$, passes through uniformly hyperbolic regions of $\mathbb{T}^2$, where invariant cones are preserved and $|f'(x_i + \omega_{i+1})| \geq L^\beta$ for each $i < N$; see Remark 1 in Section 2. We further define $\hat{G}_N = \{(x_0, y_0, \theta_0) : (x_0, y_0) \in G_N\}$.

LEMMA 10. *Let $\beta > 0$ be given. We assume L is sufficiently large (depending on β). Then for any $N \in \mathbb{N}$, $\epsilon \in (0, \frac{1}{2}]$ and $\omega_1, \ldots, \omega_N \in [-\epsilon, \epsilon]$,*

$$\hat{\mu}(\hat{G}_N \cap \{|\tan\theta_N| > 1\}) \leq \frac{\hat{C}}{L^{\frac{1}{4}}}\left(1 + \frac{1}{\epsilon^3 L^{2+\beta N}}\right).$$

Proof. For $(x_0, y_0) \in G_N$, consider the singular value decomposition of $(dF^N_{\underline{\omega}})_{(x_0,y_0)}$. Let ϑ^-_0 denote the angle corresponding to the most contracted direction of $(dF^N_{\underline{\omega}})_{(x_0,y_0)}$ and ϑ^-_N its image under $(dF^N_{\underline{\omega}})_{(x_0,y_0)}$, and let $\sigma > 1 > \sigma^{-1}$ denote the singular values of $(dF^N_{\underline{\omega}})_{(x_0,y_0)}$. A straightforward computation gives

$$\frac{1}{2}L^\beta \leq |\tan\vartheta^-_0|, |\tan\vartheta^-_N| \quad \text{and} \quad \sigma \geq \left(\frac{1}{3}L^\beta\right)^N.$$

It follows immediately that for fixed (x_0, y_0), $\{\theta_0 : |\tan\theta_N| > 1\} \subset [\pi/4, 3\pi/4]$ and

$$\text{Leb}\{\theta_0 : |\tan\theta_N| > 1\} < \text{ const } L^{-\beta N}.$$

Applying Lemma 9 with $A = \hat{G}_N \cap \{|\tan\theta_N| > 1\}$, we obtain the asserted bound. $\square$

By the stationarity of $\hat{\mu}$, it is true for any N that

$$\lambda(\hat{\mu}) = \int\left(\int \|(dF_{\omega_{N+1}})_{(x_N,y_N)} u_{\theta_N}\| d(\hat{F}_{\omega_N} \circ \cdots \circ \hat{F}_{\omega_1})_*\hat{\mu}\right) d\nu^\epsilon(\omega_1) \cdots d\nu^\epsilon(\omega_{N+1}).$$

We have chosen to estimate $\lambda(\hat{\mu})$ one sample path at a time because we have information from Lemma 10 on $(\hat{F}_{\omega_N} \circ \cdots \circ \hat{F}_{\omega_1})_*\hat{\mu}$ for each sequence $\omega_1, \ldots, \omega_N$.

PROPOSITION 11. *Let $\alpha, \beta \in (0,1)$. Then, there are constants $C = C_{\alpha,\beta} > 0$ and $C' = C'_{\alpha,\beta} > 0$ such that for any L sufficiently large, we have the following. Let $N = \lfloor C' L^{1-\beta} \rfloor$, $\epsilon \in [L^{-CL^{1-\beta}}, \frac{1}{2}]$, and fix arbitrary $\omega_1, \cdots, \omega_{N+1} \in [-\epsilon, \epsilon]$. Then,*

$$I := \int_{\mathbb{PT}^2} \log \|(dF_{\omega_{N+1}})_{(x_N,y_N)} u_{\theta_N}\| \, d\hat{\mu}(x_0, y_0, \theta_0) \geq \alpha \log L. \tag{11}$$

Integrating (11) over $(\omega_1, \ldots, \omega_{N+1})$ gives $\lambda(\hat{\mu}) \geq \alpha \log L$. As $\lambda^\epsilon_1 \geq \lambda(\hat{\mu})$, part (b) of Theorem 1 follows immediately from this proposition.

Proof. The number N will be determined in the course of the proof, and L will be enlarged a finite number of times as we go along. As usual, we will split I, the integral in (11), to one on a good and a bad set. The good set is essentially the one in Lemma 10, with an additional condition on (x_N, y_N), where dF will be evaluated. Let

$$G^*_N = \{(x_0, y_0) \in G_N : d(x_N + \omega_{N+1}, C'_\psi) \geq K_1 m\},$$

where $m > 0$ is a small parameter to be specified later. As before, we let $\hat{G}^*_N = G^*_N \times [0, 2\pi)$. Then $\mathcal{G} := \hat{G}^*_N \cap \{|\tan\theta_N| \leq 1\}$ is the good set; on $\mathcal{G}$, the integrand in (11) is $\geq \log\left(\frac{mL}{4}\right)$. Elsewhere we use the worst lower bound $-\log(2\|\psi'\|_{C_0}L)$. Altogether we have

$$I \geq \log(\frac{1}{4}mL) - \log\frac{m\|\psi'\|L^2}{2}\ \hat{\mu}(\mathcal{B}), \tag{12}$$

where

$$\mathcal{B} = \mathbb{PT}^2 \setminus \mathcal{G} = (\mathbb{PT}^2 \setminus \hat{G}_N^*) \cup (\hat{G}_N^* \cap \{|\tan\theta_N| > 1\}). \tag{13}$$

We now bound $\hat{\mu}(\mathcal{B})$. First,

$$\hat{\mu}(\mathbb{PT}^2 \setminus \hat{G}_N^*) = 1 - \mathrm{Leb}(G_N^*) \leq K_1 M_1(m + NL^{-1+\beta}), \tag{14}$$

where $M_1 = \#C'_\psi$. Letting $N = \lfloor C'L^{1-\beta}\rfloor$ and $m = C' = \frac{p}{4K_1M_1}$, where p is a small number to be determined, we obtain $\hat{\mu}(\mathbb{PT}^2 \setminus \hat{G}_N^*) \leq \frac{1}{2}p$. From Lemma 10,

$$\hat{\mu}(\hat{G}_N^* \cap \{|\tan\theta_N| > 1\}) \leq \frac{\hat{C}}{L^{\frac{1}{4}}}\left(1 + \frac{1}{(\epsilon L^{\frac{1}{3}\beta N})^3}\right). \tag{15}$$

For N as above and ϵ in the designated range (with $C = \frac{\beta}{3}C'$), the right side of (15) is easily made $< \frac{1}{2}p$ by taking L large, so we have $\hat{\mu}(\mathcal{B}) \leq p$. Plugging into (12), we see that

$$I \ \geq\ (1-2p)\log L\ - \{\text{ terms involving } \log p,\ p\log p \text{ and constants }\}.$$

Setting $p = \frac{1}{4}(1-\alpha)$ and taking L large enough, one ensures that $I > \alpha \log L$. $\square$

5. **Proof of Theorem 2**

We now show that with the additional assumption (H3), the same result holds for $\epsilon \geq L^{-CL^{-2+\beta}}$.

5.1. *Proof of the theorem modulo the main proposition.* As the idea of the proof of Theorem 2 closely parallels that of Proposition 11, it is useful to recapitulate the main ideas:

(1) the main Lyapunov exponent estimate is carried on the subset $\{(x_0, y_0, \theta_0) : (x_0, y_0) \in G_N, |\tan\theta_N| < 1\}$ of $\mathbb{PT}^2$, where G_N consists of points whose orbits stay $\gtrsim L^{-1+\beta}$ away from $C'_\psi \times \mathbb{S}^1$ in their first N iterates;
(2) since $\mathrm{Leb}(G_N^c) \sim NL^{-1+\beta}$, we must take $N \lesssim L^{1-\beta}$;
(3) by the uniform hyperbolicity of $F_{\underline{\omega}}^N$ on G_N, $\mathrm{Leb}\{|\tan\theta_N| > 1\} \sim L^{-cN}$;
(4) for $\hat{\mu}\{|\tan\theta_N| > 1\}$ to be small, we must have $\frac{1}{\epsilon^3}L^{-cN} \ll 1$ (Lemma 10).

Items (2)–(4) together suggest that we require $\epsilon > L^{-\frac{1}{3}cN} \geq L^{-c'L^{-1+\beta}}$, and we checked that for this ϵ, the proof goes through.

The proof of Theorem 2 we now present differs from the above in the following way: The set G_N, which plays the same role as in Theorem 1, will be different. It will satisfy

(A) $\mathrm{Leb}(G_N^c) \sim NL^{-2+\beta}$, and
(B) the composite map $dF_{\underline{\omega}}^N$ is uniformly hyperbolic on G_N.

The idea is as follows. To decrease ϵ, we must increase N, while keeping the set G_N^c small. This can be done by allowing the random orbit to come closer to $C'_\psi \times \mathbb{S}^1$, but with that, one cannot expect uniform hyperbolicity in each of the first N iterations, so we require only (B). This is the main difference between Theorems 1 and 2. Once G_N is properly identified and properties (A) and (B) are proved, the rest of the proof follows that of Theorem 1: Property (A) permits us to take $N \sim L^{2-\beta}$ in item (2), and item (3) is valid by property (B). Item (4) is general and therefore unchanged, leading to the conclusion that it suffices to assume $\epsilon > L^{-c'L^{-2+\beta}}$. As the arguments follow those in Theorem 1 *verbatim* modulo the bounds above and accompanying constants, we will not repeat the proof. The rest of this section is focused on producing G_N with the required properties.

It is assumed from here on that (H3)(c_0) holds, and L, a and ϵ are as in Theorem 2. Having proved Theorem 1, we may assume $\epsilon \le L^{-1}$. In light of the discussion above, $\omega_1, \cdots, \omega_N, \omega_{N+1} \in [-\epsilon, \epsilon]$ will be fixed throughout, and $(x_i, y_i) = F_{\underline{\omega}}^i(x_0, y_0)$ as before.

Definition of G_N. For arbitrary N, we define G_N to be

$$
\begin{aligned}
G_N = \{(x_0, y_0) \in \mathbb{T}^2 :& \\
&\text{(a) for all } 0 \le i \le N-1, \\
&\qquad \text{(i) } d(x_i + \omega_{i+1}, C'_\psi) \ge K_1 L^{-2+\beta}, \\
&\qquad \text{(ii) } d(x_i + \omega_{i+1}, C'_\psi) \cdot d(x_{i+1} + \omega_{i+2}, C'_\psi) \ge K_1^2 L^{-2+\beta/2}, \\
&\text{(b) } d(x_0 + \omega_1, C'_\psi),\ d(x_{N-1} + \omega_N, C'_\psi) \ge p/(16M_1)\},
\end{aligned}
$$

where $M_1 = \# C'_\psi$ and $p = p(\alpha)$ is a small number to be determined. Notice that (a)(i) implies only $|f'(x_i + \omega_{i+1})| \ge L^{-1+\beta}$, not enough to guarantee expansion in the horizontal direction. We remark also that even though (a)(ii) implies $|f'(x_i + \omega_{i+1}) f'(x_{i+1} + \omega_{i+2})| \ge L^{\beta/2}$, hyperbolicity does not follow without control of the *angles* of the vectors involved.

LEMMA 12 (Property (A)). *There exists $C_2 \ge 1$ such that for all N,*

$$Leb(G_N^c) \le C_2 N L^{-2+\beta} + \frac{p}{4}.$$

Proof. Let

$$
\begin{aligned}
A_1 &= \{x \in [0,1) : d(x, C'_\psi) \ge K_1 L^{-2+\beta}\}, \\
A_2 &= \{(x,y) \in \mathbb{T}^2 : x \in A_1,\ \text{and } d(x, C'_\psi) \cdot d(fx - y, C'_\psi) \ge K_1^2 L^{-2+\beta/2}\}.
\end{aligned}
$$

We begin by estimating $\mathrm{Leb}(A_2)$. Note that $\mathrm{Leb}(A_1^c) \le 2M_1K_1L^{-2+\beta}$, and for each fixed $x \in A_1$,

$$\text{(16)} \qquad \mathrm{Leb}\left\{y \in [0,1) : d(fx - y, C'_\psi) < \frac{K_1^2 L^{-2+\beta/2}}{d(x, C'_\psi)}\right\} \le \frac{2M_1K_1^2L^{-2+\beta/2}}{d(x, C'_\psi)},$$

hence

$$\mathrm{Leb}\, A_2^c \le \mathrm{Leb}\, A_1^c + \int_{x\in A_1} \frac{2M_1K_1^2L^{-2+\beta/2}}{d(x, C'_\psi)} dx.$$

Let $\hat{c} = \frac{1}{2}\min\{d(\hat{x}, \hat{x}') : \hat{x}, \hat{x}' \in C'_\psi, \hat{x} \ne \hat{x}'\}$. We split the integral above into $\int_{d(x,C'_\psi)>\hat{c}} + \int_{K_1L^{-2+\beta}\le d(x,C'_\psi)\le\hat{c}}$. The first one is bounded from above by $2M_1K_1^2\hat{c}^{-1}L^{-2+\beta/2}$ and the second by

$$\text{(17)} \qquad 4M_1^2K_1^2L^{-2+\beta/2}\int_{K_1L^{-2+\beta}}^{\hat{c}} \frac{du}{u} \le 4(2-\beta)M_1^2K_1^2L^{-2+\beta/2}\log L$$

(having used that $-\log K_1$ and $\log \hat{c}$ are < 0). So on taking L large enough so that $L^{\beta/2} \ge \log L$, it follows that $\mathrm{Leb}(A_2^c) \le C_2L^{-2+\beta}$, where $C_2 = C_{2,\psi}$ depends on ψ alone.

Let $\widetilde{G}_N$ be equal to G_N with condition (b) removed. Then

$$\widetilde{G}_N = \bigcap_{i=0}^{N-1} (F^i_{\underline{\omega}})^{-1}\big(A_2 - (\omega_{i+1}, 0)\big),$$

so $\mathrm{Leb}(\widetilde{G}_N) \ge 1 - C_2NL^{-2+\beta}$. The rest is obvious. $\square$

Proposition 13 (Property (B)). *For any $N \ge 2$, $(dF^N_{\underline{\omega}})_{(x_0,y_0)}$ is hyperbolic on G_N with the following uniform bounds. The larger singular value σ_1 of $(dF^N_{\underline{\omega}})_{(x_0,y_0)}$ satisfies*

$$\sigma_1\big((dF^N_{\underline{\omega}})_{(x_0,y_0)}\big) \ge L^{\frac{\beta}{15}N},$$

and if $\vartheta_0^- \in [0,\pi)$ denotes the most contracting direction of $(dF^N_{\underline{\omega}})_{(x_0,y_0)}$ and $\vartheta_N^- \in [0,\pi)$ its image, then

$$|\vartheta_0^- - \pi/2|, |\vartheta_N^- - \pi/2| \le L^{-\beta}.$$

The bulk of the work in the proof of Theorem 2 goes into proving this proposition.

5.2. *Proof of Property* (B) *modulo technical estimates.* Let $c = c_\psi \ll c_0$, where c_0 is as in (H3); we stipulate additionally that $c \le p/16M_1$, where $p = p_\alpha$ and M_1 are as before. First we introduce the following symbolic encoding of $\mathbb{T}^2$. Let

$$B = \mathcal{N}_{\sqrt{\frac{c}{L}}}(C'_\psi) \times \mathbb{S}^1, \quad I = \mathcal{N}_c(C'_\psi) \times \mathbb{S}^1 \setminus B, \quad \text{and} \quad G = \mathbb{T}^2 \setminus (B \cup I).$$

To each $(x_0, y_0) \in \mathbb{T}^2$ we associate a symbolic sequence

$$(x_0, y_0) \mapsto \bar{W} = W_{N-1} \cdots W_1 W_0 \in \{B, I, G\}^N,$$

where $(x_i + \omega_{i+1}, y_i) \in W_i$. We will refer to any symbolic sequence of length ≥ 1, e.g., $\bar{V} = GBBG$, as a *word*, and we use $\mathrm{Len}(\bar{V})$ to denote the length of $\bar{V}$, i.e., the number of letters it contains. We also write G^k as shorthand for a word consisting of k copies of G. Notice that symbolic sequences are to be read from right to left.

The following is a direct consequence of (H3).

LEMMA 14. *Assume that $\epsilon < L^{-1}$. Let $(x_0, y_0) \in \mathbb{T}^2$ be such that $(x_0 + \omega_1, y_0) \in B \cup I$ and $(x_1 + \omega_2, y_1) \in B$. Then $(x_2 + \omega_3, y_2) \in G$.*

Proof. Let $\hat{x}_0, \hat{x}_1 \in C'_\psi$ (possibly $\hat{x}_0 = \hat{x}_1$) be such that $d(x_0, \hat{x}_0) < c$ and $d(x_1, \hat{x}_1) < \sqrt{\frac{c}{L}}$. Since $(f(\hat{x}_1) - \hat{x}_0) \pmod 1 \notin \mathcal{N}_{c_0}(C'_\psi)$ by (H3), it suffices to show $|x_2 - (f(\hat{x}_1) - \hat{x}_0) \pmod 1| \ll c_0$:

$$\begin{aligned} |x_2 - (f(\hat{x}_1) - \hat{x}_0) \pmod 1| &= |(f(x_1 + \omega_2) - y_1) - (f(\hat{x}_1) - \hat{x}_0) \pmod 1| \\ &\leq |f(x_1 + \omega_2) - f(\hat{x}_1)| + d(y_1, \hat{x}_0). \end{aligned}$$

To see that this is $\ll c_0$, observe that for large L, we have

$$|f(x_1 + \omega_2) - f(\hat{x}_1)| < \frac{1}{2} L \|\psi''\| \left(\sqrt{\frac{c}{L}} + L^{-1} \right)^2 < \|\psi''\| c$$

and $d(y_1, \hat{x}_0) = d(x_0 + \omega_1, \hat{x}_0) < 2c$. □

Next we apply Lemma 14 to put constraints on the set of all possible words $\bar{W}$ associated with $(x_0, y_0) \in G_N$.

LEMMA 15. *Let $\bar{W}$ be associated with $(x_0, y_0) \in G_N$. Then $\bar{W}$ must have the following form:*

$$\bar{W} = G^{k_M} \bar{V}_M G^{k_{M-1}} \bar{V}_{M-1} \cdots G^{k_1} \bar{V}_1 G^{k_0}, \tag{18}$$

where $M \geq 0$, $k_0, k_1, \cdots, k_M \geq 1$, and if $M > 0$, then each $\bar{V}_i$ is one of the words in

$$\mathcal{V} = \{B, BB, \text{ or } BI^kB, I^kB, I^k, BI^k \text{ for some } k \geq 1\}.$$

Proof. The sequence $\bar{W}$ starts and ends with G by the definition of G_N and the stipulation that $c \leq p/(16M_1)$; thus a decomposition of the form (18) is obtained with words $\{\bar{V}_i\}_{i=1}^M$ formed from the letters $\{I, B\}$. To show that the words $\{\bar{V}_i\}_{i=1}^M$ must be of the proscribed form, observe that

- BB occurs only as a subword of $GBBG$,
- BI only occurs as a subword of GBI,
- IB only occurs as a subword of IBG.

Each of these constraints follows from Lemma 14; for the third, G is the only letter that can precede IB. It follows from the last two bullets that all the Is must be consecutive, and B can appear at most twice. $\square$

With respect to the representation in (18), we view each $\bar{V}_i$ as representing an excursion away from the "good region" G. In what follows, we will show that G_N and (H3) are chosen so that for $(x_0, y_0) \in G_N$, vectors are not rotated by too much during these excursions, and hyperbolicity is restored with each visit to G. To prove this, we introduce the following cones in tangent space:

$$\mathcal{C}_n = \mathcal{C}(L^{-1+\beta/4}), \qquad \mathcal{C}_1 = \mathcal{C}(1), \quad \text{and} \quad \mathcal{C}_w = \mathcal{C}(L^{1-\beta/4}),$$

where $\mathcal{C}(s)$ refers to the cone of vectors whose slopes have absolute value $\leq s$. The letters n, w stand for "narrow" and "wide," respectively.

Let $(x_0, y_0) \in G_N$, and for some m and l, suppose that

$$\{(x_{m+i-1} + \omega_{m+i}, y_{m+i-1})\}_{i=1}^{l}$$

corresponds to the word $\bar{V} = V_l \cdots V_1 \in \mathcal{V}$. To simplify notation, we write

$$(\widetilde{x}_i, \widetilde{y}_i) = (x_{m+i-1} + \omega_{m+i}, y_{m+i-1}) \qquad \text{and} \qquad d\widetilde{F}^l = dF_{(\widetilde{x}_l, \widetilde{y}_l)} \circ \cdots \circ dF_{(\widetilde{x}_1, \widetilde{y}_1)}.$$

Proposition 16. *Let $\{(\widetilde{x}_i, \widetilde{y}_i)\}_{i=1}^{l}$ and $\bar{V} = V_l \cdots V_1 \in \mathcal{V}$ be as above. Then*

$$d\widetilde{F}^l(\mathcal{C}_n) \subset \mathcal{C}_w, \quad \left(d\widetilde{F}^l\right)^*(\mathcal{C}_n) \subset \mathcal{C}_w \quad \textit{and} \quad \min_{u \in \mathcal{C}_n, \|u\|=1} \|d\widetilde{F}^l u\| \ \geq \ \frac{1}{2} L^{\frac{\beta}{5} n_I(\bar{V})},$$

where $n_I(\bar{V})$ is the number of appearances of the letter I in the word $\bar{V}$.

We defer the proof of Proposition 16 to the next subsection.

Proof of Proposition 13 *assuming Proposition* 16. For $(x_0, y_0) \in G_N$, let $\bar{W}$ be as in (18). It is easy to check that if $(x_m + \omega_{m+1}, y_m) \in G$, then

(19)
$$(dF_{\omega_{m+1}})_{(x_m, y_m)}(\mathcal{C}_w) \subset \mathcal{C}_n \quad \text{with} \quad \min_{u \in \mathcal{C}_w, \|u\|=1} \|(dF_{\omega_{m+1}})_{(x_m, y_m)} u\| \geq \frac{1}{4} L^{\beta/4}.$$

Applying (19) and Proposition 16 alternately, we obtain

$$(dF^N_{\underline{\omega}})_{(x_0, y_0)}(\mathcal{C}_w) \subset \mathcal{C}_n.$$

Identical considerations for the adjoint yield the cones relation $(dF^N_{\underline{\omega}})^*_{(x_0, y_0)} \mathcal{C}_w \subset \mathcal{C}_n$. We now use the following elementary fact from linear algebra: if M is a 2×2 real matrix with distinct real eigenvalues $\eta_1 > \eta_2$ and corresponding eigenvectors $v_1, v_2 \in \mathbb{R}^2$, and if $\mathcal{C}$ is any closed convex cone with nonempty interior for which $M\mathcal{C} \subset \mathcal{C}$, then $v_1 \in \mathcal{C}$.

We conclude that the maximal expanding direction ϑ_0^+ for $(dF^N_{\underline{\omega}})_{(x_0, y_0)}$ and its image ϑ_N^+ both belong to $\mathcal{C}_n$. The estimates for $\vartheta_0^-, \vartheta_N^-$ now follow on recalling that $\vartheta_0^- = \vartheta_0^+ + \pi/2 \,(\text{mod } \pi), \vartheta_N^- = \vartheta_N^+ + \pi/2 \,(\text{mod } \pi)$.

It remains to compute $\sigma_1\Big((dF^N_{\underline{\omega}})_{(x_0,y_0)}\Big)$. From (19) and the derivative bound in Proposition 16, we obtain

$$\min_{u\in\mathcal{C}_w,\|u\|=1}\|(dF^N_{\underline{\omega}})_{(x_0,y_0)}u\|\geq L^{\frac{\beta}{5}\left((k_0-1)+(k_1-1)+\cdots+(k_{M-1}-1)+k_M+\sum_{i=1}^{M}(n_I(\bar{V}_i)+1)\right)}$$

As there cannot be more than two copies of B in each $\bar{V}\in\mathcal{V}$, we have

$$\frac{n_I(\bar{V})+1}{\mathrm{Len}(\bar{V})+1}\geq\frac{1}{3},$$

and the asserted bound follows. □

5.3. *Proof of Proposition* 16. Cones relations for adjoints are identical to those of the original and so are omitted; hereafter, we work exclusively with the original (unadjointed) derivatives. We will continue to use the notation in Proposition 16. Additionally, in each of the assertions below, if $d\widetilde{F}^l$ is applied to the cone $\mathcal{C}$, then min refers to the minimum taken over all unit vectors $u\in\mathcal{C}$.

The proof consists of enumerating all cases of $\bar{V}\in\mathcal{V}$. We group the estimates as follows.

LEMMA 17.
(a) *For* $\bar{V}=I$: $d\widetilde{F}(\mathcal{C}_1)\subset\mathcal{C}_1$ *and* $\min\|d\widetilde{F}u\|\geq\frac{1}{2}K_1\sqrt{c}\sqrt{L}\gg L^{1/4}$.
(b) *For* $\bar{V}=B$: $d\widetilde{F}(\mathcal{C}_n)\subset\mathcal{C}_w$ *and* $\min\|d\widetilde{F}u\|\geq\frac{1}{2}$.

The next group consists of two-letter words, the treatment of which will rely on condition (a)(ii) in the definition of G_N.

LEMMA 18.
(c) *For* $\bar{V}=BB$: $d\widetilde{F}^2(\mathcal{C}_n)\subset\mathcal{C}_w$ *and* $\min\|d\widetilde{F}^2u\|\geq L^{\beta/3}$.
(d) *For* $\bar{V}=BI$: $d\widetilde{F}^2(\mathcal{C}_1)\subset\mathcal{C}_w$ *and* $\min\|d\widetilde{F}^2u\|\geq\min\{\frac{1}{2}K_1\sqrt{c}\sqrt{L},L^{\beta/3}\}\geq L^{\beta/5}$.
(e) *For* $\bar{V}=IB$: $d\widetilde{F}^2(\mathcal{C}_n)\subset\mathcal{C}_1$ *and* $\min\|d\widetilde{F}^2u\|\geq L^{\beta/3}$.

This leaves us with the following most problematic case.

LEMMA 19.
(f) *For* $\bar{V}=BIB$: $d\widetilde{F}^3(\mathcal{C}_n)\subset\mathcal{C}_w$ *and* $\min\|d\widetilde{F}^3u\|\geq L^{\beta/5}$.

Proof of Proposition 16 *assuming Lemmas* 17–19. We go over the following checklist:

- $\bar{V}=B$ or BB was covered by (b) and (c); total growth on $\mathcal{C}_n$ is $\geq\frac{1}{2}$.

For $k\geq1$,

- $\bar{V}=I^k$ follows from (a); total growth on $\mathcal{C}_n$ is $\geq L^{k/4}\gg L^{\frac{\beta}{5}k}$.

- $\bar{V} = I^k B = I^{k-1}(IB)$ follows from concatenating (e) and (a); total growth on $\mathcal{C}_n$ is $\geq L^{(k-1)/4} \cdot L^{\beta/3} \gg L^{\frac{\beta}{5}k}$.
- $\bar{V} = BI^k = (BI)I^{k-1}$ follows from concatenating (a) and (d); total growth on $\mathcal{C}_n$ is $\geq L^{\beta/5} \cdot L^{(k-1)/4} \gg L^{\frac{\beta}{5}k}$.

Lastly,

- $\bar{V} = BIB$ follows from (f); total growth on $\mathcal{C}_n$ is $\geq L^{\beta/5}$; and
- for $k \geq 2$, $\bar{V} = BI^kB = (BI)I^{k-2}(IB)$ follows by concatenating (e), followed by (a) then (d); total growth on $\mathcal{C}_n$ is $\geq L^{\beta/5} \cdot L^{(k-2)/4} \cdot L^{\beta/3} \gg L^{\frac{\beta}{5}k}$.

This completes the proof. □

Lemma 17 is easy and left to the reader; it is a straightforward application of the formulae

$$\tan\theta_1 = \frac{1}{f'(\widetilde{x}_1) - \tan\theta_0}, \quad \|d\widetilde{F}u_\theta\| = \sqrt{(f'(\widetilde{x}_1)\cos\theta_0 - \sin\theta_0)^2 + \cos^2\theta_0},$$

where $\theta_1 \in [0,\pi)$ denotes the angle of the image vector $d\widetilde{F}u_{\theta_0}$.

Below we let K be such that $|f'| \leq KL$.

Proof of Lemma 18. We write $u = u_{\theta_0}$ and $\theta_1, \theta_2 \in [0, 2\pi)$ for the angles of the images $d\widetilde{F}u, d\widetilde{F}^2u$ respectively. Throughout, we use the following "two step" formulae:

$$\tan\theta_2 = \frac{f'(\widetilde{x}_1) - \tan\theta_0}{f'(\widetilde{x}_1)f'(\widetilde{x}_2) - f'(\widetilde{x}_2)\tan\theta_0 - 1}, \tag{20}$$

$$\|d\widetilde{F}^2u_\theta\| \geq |(f'(\widetilde{x}_1)f'(\widetilde{x}_2) - 1)\cos\theta_0| - |f'(\widetilde{x}_2)\sin\theta_0|. \tag{21}$$

The estimate $|f'(\widetilde{x}_1)f'(\widetilde{x}_2)| \geq L^{\beta/2}$ (condition (a)(ii) in the definition of G_N) will be used repeatedly throughout.

We first handle the vector growth estimates. For (c) and (e), as $u = u_{\theta_0} \in \mathcal{C}_n$, the right side of (21) is $\geq \frac{1}{2}L^{\beta/2} - 2KL^{\beta/4} \gg L^{\beta/3}$.

For (d), we break into the cases

(d.i) $|f'(\widetilde{x}_2)| \geq L^{\beta/4}$ and
(d.ii) $|f'(\widetilde{x}_2)| < L^{\beta/4}$.

In case (d.i), by (a) we have that $u_{\theta_1} \in \mathcal{C}_1$ and $\|dF_{(\widetilde{x}_1,\widetilde{y}_1)}u_{\theta_0}\| \geq \frac{1}{2}K_1\sqrt{c}\sqrt{L}$. Thus $|\tan\theta_2| \leq 2L^{-\beta/4} \ll 1$ and $\|dF_{(\widetilde{x}_2,\widetilde{y}_2)}u_{\theta_1}\| \geq \frac{1}{2}L^{\beta/4} \gg 1$, completing the proof. In case (d.ii), the right side of (21) is

$$\geq \frac{1}{\sqrt{2}}(L^{\beta/2} - 1) - \frac{1}{\sqrt{2}}L^{\beta/4} \gg L^{\beta/3}.$$

We now check the cones relations for (c)–(e). For (c),

$$\begin{aligned}|\tan\theta_2| &\le \frac{|f'(\widetilde{x}_1)| + |\tan\theta_0|}{|f'(\widetilde{x}_1)f'(\widetilde{x}_2)| - |f'(\widetilde{x}_2)\tan\theta_0| - 1}\\ &\le \frac{KL + L^{-1+\beta/4}}{L^{\beta/2} - KL^{\beta/4} - 1} \le 2KL^{1-\beta/2} \ll L^{1-\beta/4},\end{aligned}$$

so that $u_{\theta_2} \in \mathcal{C}_w$ as advertised. The case (d.i) has already been treated. For (d.ii), the same bound as in (c) gives

$$|\tan\theta_2| \le \frac{KL+1}{L^{\beta/2} - L^{\beta/4} - 1} \le 2KL^{1-\beta/2} \ll L^{1-\beta/4},$$

hence $u_{\theta_2} \in \mathcal{C}_w$.

For (e) we again distinguish the cases

(e.i) $|f'(\widetilde{x}_1)| \ge L^{\beta/4}$ and

(e.ii) $|f'(\widetilde{x}_1)| < L^{\beta/4}$.

In case (e.i), one easily checks that $dF_{(\widetilde{x}_1,\widetilde{y}_1)}(\mathcal{C}_n) \subset \mathcal{C}_1$ and then $dF_{(\widetilde{x}_2,\widetilde{y}_2)}(\mathcal{C}_1) \subset \mathcal{C}_1$ by (a). In case (e.ii) we compute directly that

$$|\tan\theta_2| \le \frac{|f'(x_1)| + |\tan\theta_0|}{|f'(x_1)f'(x_2)| - |f'(x_2)\tan\theta_0| - 1} \le \frac{L^{\beta/4} + L^{-1+\beta/4}}{L^{\beta/2} - KL^{\beta/4} - 1} \ll 1,$$

hence $u_{\theta_2} \in \mathcal{C}_1$. □

Proof of Lemma 19. We let $u = u_{\theta_0} \in \mathcal{C}_n$ (i.e., $|\tan\theta_0| \le L^{-1+\beta/4}$) and write $\theta_1, \theta_2, \theta_3 \in [0,\pi)$ for the angles associated to the subsequent images of u. We break into two cases:

(I) $|f'(\widetilde{x}_3)| \ge |f'(\widetilde{x}_1)|$ and

(II) $|f'(\widetilde{x}_3)| < |f'(\widetilde{x}_1)|$.

In case (I), we compute

$$\begin{aligned}|\tan\theta_2| &\le \frac{|f'(\widetilde{x}_1)| + |\tan\theta_0|}{|f'(\widetilde{x}_1)f'(\widetilde{x}_2)| - |f'(\widetilde{x}_2)\tan\theta_0| - 1}\\ &\le \frac{2|f'(\widetilde{x}_1)|}{L^{\beta/2} - 2KL^{\beta/4} - 1} \le 4|f'(\widetilde{x}_1)|L^{-\beta/2},\end{aligned}$$

having used that $|f'(\widetilde{x}_1)| \ge L^{-1+\beta}$ and $|\tan\theta_0| \le L^{-1+\beta/4}$ in the second inequality. Now,

$$\begin{aligned}|\tan\theta_3| &\le \frac{1}{|f'(\widetilde{x}_3)| - |\tan\theta_2|} \le \frac{1}{|f'(\widetilde{x}_1)| - 4|f'(\widetilde{x}_1)|L^{-\beta/2}}\\ &\le \frac{2}{|f'(\widetilde{x}_1)|} \le 2L^{1-\beta} \ll L^{1-\beta/4}.\end{aligned}$$

In case (II), we use

$$|\tan\theta_1| \le \frac{1}{|f'(\widetilde{x}_1)| - |\tan\theta_0|} \le \frac{1}{|f'(\widetilde{x}_3)| - L^{-1+\beta/4}} \le \frac{2}{|f'(\widetilde{x}_3)|},$$

again using that $|f'(\widetilde{x}_3)| \geq L^{-1+\beta}$, and then

$$\begin{aligned}|\tan\theta_3| &\leq \frac{|f'(\widetilde{x}_2)| + |\tan\theta_1|}{|f'(\widetilde{x}_2)f'(\widetilde{x}_3)| - |f'(\widetilde{x}_3)\tan\theta_1| - 1}\\ &\leq \frac{KL + 2|f'(\widetilde{x}_3)|^{-1}}{L^{\beta/2}-3} \leq \frac{KL+2L^{1-\beta}}{L^{\beta/2}-3} \leq 2KL^{1-\beta/2} \ll L^{1-\beta/4}.\end{aligned}$$

For vector growth, observe that from (e) we have $d\widetilde{F}^2(\mathcal{C}_n) \subset \mathcal{C}_1$ and $\min \|d\widetilde{F}^2 u\| \geq L^{\beta/3}$. So, if $|f'(\widetilde{x}_3)| \geq L^{\beta/12}$, then

$$\|dF_{(\widetilde{x}_3,\widetilde{y}_3)} u_{\theta_2}\| \geq \frac{1}{\sqrt{2}}(L^{\beta/12}-1) \gg 1.$$

Conversely, if $|f'(\widetilde{x}_3)| < L^{\beta/12}$, then we can use the crude estimate $\|(dF_{(\widetilde{x},\widetilde{y})})^{-1}\| \leq \sqrt{|f'(\widetilde{x})|^2+1}$ applied to $(\widetilde{x},\widetilde{y}) = (\widetilde{x}_3,\widetilde{y}_3)$, yielding

$$\|(dF_{(\widetilde{x}_3,\widetilde{y}_3)})^{-1}\| \leq \sqrt{L^{2\beta/12}+1} \leq 2L^{\beta/12},$$

hence $\|d\widetilde{F}^3 u\| \geq \frac{1}{2}L^{\beta/3-\beta/12} = \frac{1}{2}L^{\beta/4} \gg L^{\beta/5}$, completing the proof. □

6. The standard map

Let ψ and $f_0 = f_{\psi,L,a} = L\psi + a$ be as defined in Section 2.1.

LEMMA 20. *There exists $\varepsilon > 0$ and $K_0 > 1$, depending only on ψ, for which the following holds: for all $L > 0$ and $f \in \mathcal{U}_{\varepsilon,L}(f_0)$,*

(a) $\max\{\|f'\|_{C^0}, \|f''\|_{C^0}, \|f'''\|_{C^0}\} \leq K_0 L$;

(b) *the cardinalities of C'_f and C''_f are equal to those of f_0 (equivalently those of ψ);*

(c) $\min_{\hat{x}\in C'_f} |f''(\hat{x})|,\ \min_{\hat{z}\in C''_f} |f'''(\hat{z})| \geq K_0^{-1}L$; *and*

(d) $\min_{\hat{x},\hat{x}'\in C'_f} d(\hat{x},\hat{x}'),\ \min_{\hat{z},\hat{z}'\in C''_f} d(\hat{z},\hat{z}') \geq K_0^{-1}$.

The proof is straightforward and is left to the reader.

Proof of Theorem 3. We claim — and leave it to the reader to check — that the proofs in Sections 3–5 (with C'_f, C''_f replacing C'_ψ, C''_ψ) use only the form of the maps $F = F_f$ as defined in Section 2.1 and the four properties above. Thus they prove Theorem 3 as well. □

Proof of Corollary 4. Under the (linear) coordinate change $x = \frac{1}{2\pi}\theta, y = \frac{1}{2\pi}(\theta - I)$, the standard map conjugates to the map

$$(x,y) \mapsto (L\sin(2\pi x) + 2x - y, x)$$

defined on $\mathbb{T}^2$, with both coordinates taken modulo 1. This map is of the form F_f, with $f(x) = f_0(x) + 2x$ and $f_0(x) := L\sin(2\pi x)$; here $a = 0$ and $\psi(x) = \sin(2\pi x)$. Let $\varepsilon > 0$ be given by Theorem 3 for this choice of ψ. Then f clearly belongs in $\mathcal{U}_{\varepsilon,L}(f_0)$ for large enough L. □

References

[AB02] A. Avila and J. Bochi, A formula with some applications to the theory of Lyapunov exponents, *Israel J. Math.* **131** (2002), 125–137. MR 1942304. Zbl 1022.37019. http://dx.doi.org/10.1007/BF02785853.

[AJ09] A. Avila and S. Jitomirskaya, The Ten Martini Problem, *Ann. of Math.* **170** (2009), 303–342. MR 2521117. Zbl 1166.47031. http://dx.doi.org/10.4007/annals.2009.170.303.

[AV10] A. Avila and M. Viana, Extremal Lyapunov exponents: an invariance principle and applications, *Invent. Math.* **181** (2010), 115–189. MR 2651382. Zbl 1196.37054. http://dx.doi.org/10.1007/s00222-010-0243-1.

[BC85] M. Benedicks and L. Carleson, On iterations of $1 - ax^2$ on $(-1, 1)$, *Ann. of Math.* **122** (1985), 1–25. MR 0799250. Zbl 0597.58016. http://dx.doi.org/10.2307/1971367.

[BC91] M. Benedicks and L. Carleson, The dynamics of the Hénon map, *Ann. of Math.* **133** (1991), 73–169. MR 1087346. Zbl 0724.58042. http://dx.doi.org/10.2307/2944326.

[BC14] P. Berger and P. D. Carrasco, Non-uniformly hyperbolic diffeomorphisms derived from the standard map, *Comm. Math. Phys.* **329** (2014), 239–262. MR 3207003. Zbl 06312185. http://dx.doi.org/10.1007/s00220-014-1932-1.

[Boc02] J. Bochi, Genericity of zero Lyapunov exponents, *Ergodic Theory Dynam. Systems* **22** (2002), 1667–1696. MR 1944399. Zbl 1023.37006. http://dx.doi.org/10.1017/S0143385702001165.

[BV05] J. Bochi and M. Viana, The Lyapunov exponents of generic volume-preserving and symplectic maps, *Ann. of Math.* **161** (2005), 1423–1485. MR 2180404. Zbl 1101.37039. http://dx.doi.org/10.4007/annals.2005.161.1423.

[BNV10] C. Bocker-Neto and M. Viana, Continuity of Lyapunov exponents for random 2D matrices, 2010. arXiv 1012.0872.

[BGMV03] C. Bonatti, X. Gómez-Mont, and M. Viana, Généricité d'exposants de Lyapunov non-nuls pour des produits déterministes de matrices, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **20** (2003), 579–624. MR 1981401. Zbl 1025.37018. http://dx.doi.org/10.1016/S0294-1449(02)00019-7.

[BV04] C. Bonatti and M. Viana, Lyapunov exponents with multiplicity 1 for deterministic products of matrices, *Ergodic Theory Dynam. Systems* **24** (2004), 1295–1330. MR 2104587. Zbl 1087.37017. http://dx.doi.org/10.1017/S0143385703000695.

[Bou13] J. Bourgain, On the Lyapunov exponents of Schrödinger operators associated with the standard map, in *Asymptotic Geometric Analysis*, *Fields Institute Communications* **68**, Springer-Verlag, New York, 2013, pp. 39–44. MR 3076146. Zbl 1273.37025. http://dx.doi.org/10.1007/978-1-4614-6406-8_3.

[CS] L. Carleson and T. Spencer, personal communication.

[Dua08] P. Duarte, Elliptic isles in families of area-preserving maps, *Ergodic Theory Dynam. Systems* **28** (2008), 1781–1813. MR 2465600. Zbl 1179.37007. http://dx.doi.org/10.1017/S0143385707000983.

[Dua94] P. Duarte, Plenty of elliptic islands for the standard family of area preserving maps, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **11** (1994), 359–409. MR 1287238. Zbl 0838.58030.

[Fur63] H. Furstenberg, Noncommuting random products, *Trans. Amer. Math. Soc.* **108** (1963), 377–428. MR 0163345. Zbl 0203.19102. http://dx.doi.org/10.2307/1993589.

[GM89] I. Y. Gol′dsheĭd and G. A. Margulis, Lyapunov exponents of a product of random matrices, *Uspekhi Mat. Nauk* **44** (1989), 13–60. MR 1040268. Zbl 0687.60008. http://dx.doi.org/10.1070/RM1989v044n05ABEH002214.

[Gor12] A. Gorodetski, On stochastic sea of the standard map, *Comm. Math. Phys.* **309** (2012), 155–192. MR 2864790. Zbl 06005038. http://dx.doi.org/10.1007/s00220-011-1365-z.

[GR86] Y. Guivarc'h and A. Raugi, Products of random matrices: convergence theorems, in *Random Matrices and their Applications* (Brunswick, Maine, 1984), *Contemp. Math.* **50**, Amer. Math. Soc., Providence, RI, 1986, pp. 31–54. MR 0841080. Zbl 0592.60015. http://dx.doi.org/10.1090/conm/050/841080.

[Hen84] H. Hennion, Loi des grands nombres et perturbations pour des produits réductibles de matrices aléatoires indépendantes, *Z. Wahrsch. Verw. Gebiete* **67** (1984), 265–278. MR 0762081. Zbl 0529.60025. http://dx.doi.org/10.1007/BF00535004.

[Her83] M.-R. Herman, Une méthode pour minorer les exposants de Lyapounov et quelques exemples montrant le caractère local d'un théorème d'Arnol′d et de Moser sur le tore de dimension 2, *Comment. Math. Helv.* **58** (1983), 453–502. MR 0727713. Zbl 0554.58034. http://dx.doi.org/10.1007/BF02564647.

[Jak81] M. V. Jakobson, Absolutely continuous invariant measures for one-parameter families of one-dimensional maps, *Comm. Math. Phys.* **81** (1981), 39–88. MR 0630331. Zbl 0497.58017. Available at http://projecteuclid.org/euclid.cmp/1103920159.

[Kif82] Y. Kifer, Perturbations of random matrix products, *Z. Wahrsch. Verw. Gebiete* **61** (1982), 83–95. MR 0671245. Zbl 0479.60019. http://dx.doi.org/10.1007/BF00537227.

[Kni92] O. Knill, Positive Lyapunov exponents for a dense set of bounded measurable $\mathrm{SL}(2, \mathbf{R})$-cocycles, *Ergodic Theory Dynam. Systems* **12** (1992), 319–331. MR 1176626. Zbl 0759.58025. http://dx.doi.org/10.1017/S0143385700006775.

[Kot84] S. Kotani, Ljapunov indices determine absolutely continuous spectra of stationary random one-dimensional Schrödinger operators, in *Stochastic*

Analysis (Katata/Kyoto, 1982), *North-Holland Math. Library* **32**, North-Holland, Amsterdam, 1984, pp. 225–247. MR 0780760. Zbl 0549.60058. http://dx.doi.org/10.1016/S0924-6509(08)70395-7.

[LSSW03] F. Ledrappier, M. Shub, C. Simó, and A. Wilkinson, Random versus deterministic exponents in a rich family of diffeomorphisms, *J. Statist. Phys.* **113** (2003), 85–149. MR 2012975. Zbl 1031.37028. http://dx.doi.org/10.1023/A:1025770720803.

[LY91] F. Ledrappier and L.-S. Young, Stability of Lyapunov exponents, *Ergodic Theory Dynam. Systems* **11** (1991), 469–484. MR 1125884. Zbl 0759.58026. http://dx.doi.org/10.1017/S0143385700006283.

[LS12] Z. Lian and M. Stenlund, Positive Lyapunov exponent by a random perturbation, *Dyn. Syst.* **27** (2012), 239–252. MR 2926701. Zbl 1293.37023. http://dx.doi.org/10.1080/14689367.2012.667787.

[Pui04] J. Puig, Cantor spectrum for the almost Mathieu operator, *Comm. Math. Phys.* **244** (2004), 297–309. MR 2031032. Zbl 1075.39021. http://dx.doi.org/10.1007/s00220-003-0977-3.

[Rue79] D. Ruelle, Analycity properties of the characteristic exponents of random matrix products, *Adv. in Math.* **32** (1979), 68–80. MR 0534172. Zbl 0426.58018. http://dx.doi.org/10.1016/0001-8708(79)90029-X.

[Ryc88] M. R. Rychlik, Another proof of Jakobson's theorem and related results, *Ergodic Theory Dynam. Systems* **8** (1988), 93–109. MR 0939063. Zbl 0671.58019. http://dx.doi.org/10.1017/S014338570000434X.

[SS15] M. Shamis and T. Spencer, Bounds on the Lyapunov exponent via crude estimates on the density of states, *Comm. Math. Phys.* **338** (2015), 705–720. MR 3351055. Zbl 06450940. http://dx.doi.org/10.1007/s00220-015-2324-x.

[Spe] T. Spencer, Einstein Chair Lecture at CUNY, videotape # 329.

[Vir80] A. D. Virtser, On products of random matrices and operators, *Theory Prob. Appl.* **24** (1980), 367–377. Zbl 0452.60015. http://dx.doi.org/10.1137/1124040.

[WY01] Q. Wang and L.-S. Young, Strange attractors with one direction of instability, *Comm. Math. Phys.* **218** (2001), 1–97. MR 1824198. Zbl 0996.37040. http://dx.doi.org/10.1007/s002200100379.

[WY06] Q. Wang and L.-S. Young, Nonuniformly expanding 1D maps, *Comm. Math. Phys.* **264** (2006), 255–282. MR 2212223. Zbl 1107.37028. http://dx.doi.org/10.1007/s00220-005-1485-4.

[WY08] Q. Wang and L.-S. Young, Toward a theory of rank one attractors, *Ann. of Math.* **167** (2008), 349–480. MR 2415378. Zbl 1181.37049. http://dx.doi.org/10.4007/annals.2008.167.349.

[You97] L.-S. Young, Lyapunov exponents for some quasi-periodic cocycles, *Ergodic Theory Dynam. Systems* **17** (1997), 483–504. MR 1444065. Zbl 0873.28013. http://dx.doi.org/10.1017/S0143385797079170.

(Received: February 16, 2016)

Courant Institute of Mathematical Sciences, New York University,
New York, NY
E-mail: alex@cims.nyu.edu

University of Chicago, Chicago, IL
E-mail: jxue@math.uchicago.edu

Courant Institute of Mathematical Sciences, New York University,
New York, NY
E-mail: lsy@cims.nyu.edu