TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 369, Number 2, February 2017, Pages 783–803
http://dx.doi.org/10.1090/tran/6761
Article electronically published on May 6, 2016

EXPONENTIAL MIXING FOR SKEW PRODUCTS WITH DISCONTINUITIES

OLIVER BUTTERLEY AND PEYMAN ESLAMI

ABSTRACT. We consider the 2D skew product $F : (x, u) \mapsto (f(x), u + \tau(x))$, where the base map f is a piecewise $\mathscr{C}^2$, covering and uniformly expanding the map of the circle, and the fibre map τ is piecewise $\mathscr{C}^2$. We show that this system mixes exponentially when τ is not cohomologous (via a Lipschitz function) to a piecewise constant.

1. Introduction and results

In the study of dynamical systems, establishing the rate of mixing of a given system is of foremost importance. It is a fundamental property describing the rate at which information about the system is lost. More importantly the rate of mixing (or typically slightly stronger information which is obtained whilst proving the rate of mixing) can be used to prove many other statistical properties (see, for example, [19, §9] and [7, Chapter 7]). Furthermore, of physical relevance, these strong results associated to good rates of mixing are crucial when studying weakly coupled systems [11, 22].

Rate of mixing results were first obtained for expanding maps and for hyperbolic maps (see [23] and references within), then also for slower mixing, nonuniformly hyperbolic systems (e.g., [30, 32, 33]). In the case of hyperbolic flows or skew products like the one studied here one direction is completely neutral, with no expansion or contraction. These systems are not hyperbolic but merely partially hyperbolic. In these situations there is a mechanism at work, different from hyperbolicity, but which is nonetheless sufficient for producing good statistical properties including exponential rate of mixing. Dolgopyat [10], extending work of Chernov [8], succeeded in developing technology for studying this neutral mechanism and consequently proved exponential mixing for mixing Anosov flows when the stable and unstable invariant foliations are both $\mathscr{C}^1$. Using and developing these ideas, various results followed [2, 5, 26, 31]. However all the above systems were rather smooth or at least Markov.

Our knowledge concerning this same neutral mechanism in systems with discontinuities is less than satisfactory at present. Baladi and Liverani [4] prove exponential mixing for piecewise smooth 3D hyperbolic flows which preserve a contact

Received by the editors June 26, 2014 and, in revised form, January 22, 2015.
2010 *Mathematics Subject Classification.* Primary 37A25; Secondary 37C30, 37D50.
Key words and phrases. Exponential mixing, skew product, oscillatory cancellation, transfer operator, partially hyperbolic.
The first author was supported by the Austrian Science Fund, Lise Meitner position M1583.
The second author was supported by an INdAM-COFUND Marie Curie fellowship.
This research was partially supported by the Stiftung Aktion Österreich Ungarn (AÖU), Projekt Nr. 87öu6.

structure; Obayashi [27] proves exponential mixing for certain suspension semiflows over expanding interval maps with discontinuities; Gouëzel [17] proves exponential mixing for certain skew products where the base map is nonuniformly expanding with discontinuities. In both the last works, the base map is required to admit an induced map which is Markov, uniformly expanding, and has exponential tails (i.e., a Young tower).

In this article we study the 2D skew product map of $\mathbb{T}^2$ which we denote by F. It is defined as follows. The base map f is a map on $\mathbb{T}^1$ which is required to be piecewise $\mathscr{C}^2$. By this we mean that there exists a finite set of disjoint open intervals $\{\mathbb{I}_k\}_k$ which covers $\mathbb{T}^1$ except for a finite number of points and that $f : \bigcup_k \mathbb{I}_k \to \mathbb{T}^1$ is $\mathscr{C}^2$ on each connected component of the domain and admits a $\mathscr{C}^2$ extension to the closure of each $\mathbb{I}_k$. Also f is required to be uniformly expanding and covering.[1] The fibre map $\tau : \bigcup_k \mathbb{I}_k \to \mathbb{R}$ is similarly required to be $\mathscr{C}^2$ on each connected component of the domain and admit a $\mathscr{C}^2$ extension to the closure of each $\mathbb{I}_k$. Let $\mathbb{J}_k := \mathbb{I}_k \times \mathbb{T}^1$. The skewproduct $F : \bigcup_k \mathbb{J}_k \to \mathbb{T}^2$ is defined by[2]

$$F : (x, u) \mapsto (f(x), u + \tau(x)).$$

At no stage do we require the map to be Markov, nor do we work with tower constructions to reduce to the Markov case. Since the map f is piecewise $\mathscr{C}^2$ and uniformly expanding it is known that there exists ν, an f-invariant probability measure which is absolutely continuous with respect to Lebesgue. Since the dynamics in the fibres is nothing more than a rigid rotation this means that $\mu := \nu \times \mathrm{Leb}$ is an F-invariant probability measure on $\mathbb{T}^2$. Given observables $g, h : \mathbb{T}^2 \to \mathbb{C}$ the correlation is defined as usual by $\mathrm{Cor}_{g,h}(n) := \mu(g \cdot h \circ F^n) - \mu(g) \cdot \mu(h)$. We say that $F : \mathbb{T}^2 \to \mathbb{T}^2$ *mixes exponentially* if, for each $\alpha \in (0, 1)$, there exists $\zeta > 0$, $C > 0$ such that[3] $|\mathrm{Cor}_{g,h}(n)| \leq C \|g\|_{\mathscr{C}^\alpha} \|h\|_{\mathscr{C}^\alpha} e^{-n\zeta}$ for every $g, h \in \mathscr{C}^\alpha(\mathbb{T}^2, \mathbb{C})$ and $n \in \mathbb{N}$. We say that τ is *cohomologous to a piecewise constant* if there exists Lipschitz $\theta : \mathbb{T}^1 \to \mathbb{R}$ and piecewise constant $\chi : \mathbb{T}^1 \to \mathbb{R}$ such that $\tau = \theta \circ f - \theta + \chi$. Moreover the discontinuities of χ occur only at points where either f or τ is discontinuous.

Our main result is the following.

Theorem 1. *Let $F : \mathbb{T}^2 \to \mathbb{T}^2$ be a piecewise-$\mathscr{C}^2$ skew product over an expanding map as described above. If τ is not cohomologous to a piecewise constant, then F mixes exponentially.*

The remainder of this document is devoted to the proof of the above theorem. The basic idea is from Dolgopyat [10]. However we combine the best technology from the subsequent articles [2, 5, 26, 31] in order to deal with the present difficulties, in particular the problems arising from the discontinuities.

In the case where τ is Lipschitz cohomologous to a piecewise constant function, the skew product might be mixing or might not. In the mixing case arbitrarily slow mixing rates are possible just as for suspension semiflows with piecewise constant

[1]Covering means that for every subinterval $\omega \subset \mathbb{T}^1$ there exists some n such that $f^n\omega$ covers a full measure subset of $\mathbb{T}^1$. Covering implies that the unique absolutely continuous invariant probability density is bounded away from zero [24].

[2]Here and throughout the document, if $u \in \mathbb{T}^1$, $s \in \mathbb{R}$, then we consider $u + s \in \mathbb{T}^1$ in the natural sense that $\mathbb{T}^1 = \mathbb{R}/\mathbb{Z}$.

[3]The space of α-Hölder functions on $\mathbb{T}^2$ taking values in $\mathbb{C}$ is denoted by $\mathscr{C}^\alpha(\mathbb{T}^2, \mathbb{C})$ and the Hölder norm defined as $\|h\|_{\mathscr{C}^\alpha} := \sup_{x \neq y} \frac{|h(x)-h(y)|}{|x-y|^\alpha} + \sup_x |h(x)|$.

return time functions [29]. See [16] for the case of skew products over Markov expanding maps where τ takes just two values. A polynomial rate of mixing is proven dependent on the Diophantine ratio of these two values. In an alternative direction, slow rates of mixing can be the result of a nonuniformly expanding base map, as opposed to the uniform expansion required in our setting: If the base map fails to mix exponentially, then the skew product will also fail to mix exponentially.

We note that the issue of the discontinuities can be approached by using a tower construction and so reducing to the case of a base map which is Markov. This is what has been done by Obayashi [27] and Gouëzel [17]. However one problem with such tower constructions is that the tower is very sensitive to changes in the underlying system, a phenomenon which could cause problems when one is interested in determining the behaviour of statistical properties under perturbation of the original system. Nevertheless some results on the stability of the system are possible, even using methods which involve tower constructions [1, 3, 14]. However many other stability results are of interest, for example, the following question: Is it possible to choose the constants $\zeta, C > 0$ which appear in the definition of exponential mixing such that an open set of skew products mixes at this uniform rate? There are suggestions [23] that answering this question could be possible with the direct strategy used in this article, whereas the possible sensitivity of tower constructions to perturbations suggest that such an approach could be problematic. Moreover, it is unclear how to use a tower structure when studying a hyperbolic base map, as opposed to an expanding base map, whereas the approach used in the present work has a chance to be extended to such a setting. We note that in both the papers [17, 27], the condition of not being cohomologous to a locally constant function is put on the induced fibre map (or return time), with respect to the induced base map. In this present work we are able to put this condition directly on the original fibre map with respect to the original base map.

From a technical point of view we are forced in two opposing directions in the proof. To deal with discontinuities we are forced to consider densities of rather low regularity. However we also need to take advantage of Dolgopyat's oscillatory cancellation argument, which requires some good degree of regularity for the density.

The result for skew products is closely related to the analogous result for suspension semiflows. At a techincal level this can be seen from the twisted transfer operator (introduced below (2.1)), which is the same object used when studied in the context of skew products or flows (see, for example, [6]), with exactly the same estimates being required.

In a separate work [12], the second author proves a stretched-exponential mixing rate for skew products where the base map is $\mathscr{C}^{1+\alpha}$ on a countable partition and the fibre map is piecewise $\mathscr{C}^1$. The weaker assumptions mean that $\sup |f'|$ is not necessarily finite and the invariant density for f is not bounded away from 0. Such bounds are used for example in Section 3 on transversality. Also the mere $\mathscr{C}^1$ regularity of the fibre map τ prevents one from using the type of oscillatory cancellation lemma that is used in this article (Lemma 15). The approach of [12] is based on the introduction of *complex standard pairs* and allows the author to obtain a stretched-exponential version of Proposition 2 without functional analytic estimates such as Lemma 1.

Section 2 concerns the estimate of the norm of twisted transfer operators reducing the problem to a single key estimate (Proposition 2). In Section 3 the key notion of transversality is discussed and a certain estimate (Proposition 3) is shown to hold in the case when τ is not cohomologous to a piecewise constant. The key oscillatory cancellation estimate (Proposition 2) is proven in Section 4, crucially using the transversality estimate from Section 3. Finally, in Section 5, the estimate on the twisted transfer operators is used to produce an estimate of exponential mixing.

2. Preparation for the main estimate

From this point onwards we will assume that $\tilde{\lambda} := \inf f' > 2$. In general it would suffice to assume that there exists $n \in \mathbb{N}$ such that $\inf (f^n)' > 1$. In that case we would simply consider a sufficiently large iterate m such that $\inf (f^m)' > 2$ and proceed as now. Let $\Lambda := \sup |f'| \geq \tilde{\lambda}$. For future convenience let $J_n := |(f^n)'|^{-1}$. The twisted transfer operator, for all $b \in \mathbb{R}$, $n \in \mathbb{N}$, is given by the formula

$$\mathscr{L}_b^n h(y) = \sum_{x \in f^{-n}(y)} J_n(x) \cdot h(x) \cdot e^{ib\tau_n(x)}. \tag{2.1}$$

A simple estimate shows that $\|\mathscr{L}_b^n h\|_{\mathbf{L^1}(\mathbb{T}^1)} \leq \|h\|_{\mathbf{L^1}(\mathbb{T}^1)}$. We will work extensively with functions of *bounded variation* due to the suitability of this function space for discontinuities. The Banach space is denoted $(\mathbf{BV}, \|\cdot\|_{\mathbf{BV}})$, variation is defined by $\operatorname{Var}(h) := \sup \left\{ \int_{\mathbb{T}^1} h \cdot \eta' : \eta \in \mathscr{C}^1(\mathbb{T}^1, \mathbb{C}), |\eta| \leq 1 \right\}$, and $\|\cdot\|_{\mathbf{BV}} := \operatorname{Var}(\cdot) + \|\cdot\|_{\mathbf{L^1}(\mathbb{T}^1)}$ as usual. We have the following *Lasota-Yorke* inequality.

Lemma 1. *There exists $\lambda > 1$, $C_\lambda > 0$ such that, for all $n \in \mathbb{N}$, $b \in \mathbb{R}$, $h \in \mathbf{BV}$,*

$$\|\mathscr{L}_b^n h\|_{\mathbf{BV}} \leq C_\lambda \lambda^{-n} \|h\|_{\mathbf{BV}} + C_\lambda (1 + |b|) \|h\|_{\mathbf{L^1}(\mathbb{T}^1)}.$$

Proof. The proof is essentially standard (see for example [19]), but it is important to note the factor of $|b|$ which appears in front of the $\mathbf{L^1}$ norm.

We already know that $\|\mathscr{L}_b h\|_{\mathbf{L^1}} \leq \|h\|_{\mathbf{L^1}}$. Note that

$$\operatorname{Var}(h) = \sup \left\{ \int_{\mathbb{T}^1} h \cdot \eta' : \eta \in \mathscr{C}^1(\mathbb{T}^1, \mathbb{C}), |\eta| \leq 1 \right\}.$$

Consequently we must estimate $\int_{\mathbb{T}^1} \mathscr{L}_b h \cdot \eta' = \int_{\mathbb{T}^1} h \cdot (e^{ib\tau} \cdot \eta' \circ f)$. In order to do this note that (for convenience we denote $J := J_1 = 1/|f'|$)

$$\left[J \cdot \eta \circ f \cdot e^{ib\tau} \right]' = J' \cdot \eta \circ f \cdot e^{ib\tau} + ib\tau' \cdot J \cdot \eta \circ f \cdot e^{ib\tau} + (e^{ib\tau} \cdot \eta' \circ f).$$

This means that

$$\left| \int_{\mathbb{T}^1} \mathscr{L}_b h \cdot \eta' \right| \leq \|J'\|_{\mathbf{L^\infty}} \|h\|_{\mathbf{L^1}} + |b| \|\tau' \cdot J\|_{\mathbf{L^\infty}} \|h\|_{\mathbf{L^1}} + \left| \int_{\mathbb{T}^1} h \cdot \left[J \cdot \eta \circ f \cdot e^{ib\tau} \right]' \right|.$$

The remaining problem is that $[J \cdot \eta \circ f \cdot e^{ib\tau}]$ could be discontinuous. Therefore we introduce the quantity $\phi : \mathbb{T}^1 \to \mathbb{R}$ which is piecewise affine (discontinuous only where $[J \cdot \eta \circ f \cdot e^{ib\tau}]$ is discontinuous) and such that $([J \cdot \eta \circ f \cdot e^{ib\tau}] - \phi)(x)$ tends to 0 as x approaches any discontinuity point. This means that $[J \cdot \eta \circ f \cdot e^{ib\tau} - \phi]$ is continuous and piecewise[4] $\mathscr{C}^1$. Note that $\|\phi\|_{\mathbf{L^\infty}} \leq \|J\|_{\mathbf{L^\infty}}$ and so $\left\| [J \cdot \eta \circ f \cdot e^{ib\tau}] - \phi \right\|_{\mathbf{L^\infty}} \leq 2 \|J\|_{\mathbf{L^\infty}}$. On the other hand, taking advantage of the finite number of discontinuities in this setting, we know that $\|\phi'\|_{\mathbf{L^\infty}}$ is bounded by

[4]That $[J \cdot \eta \circ f \cdot e^{ib\tau} - \phi]$ is continuous and piecewise $\mathscr{C}^1$ means that it may be approximated by a $\mathscr{C}^1$ function with error small in the appropriate sense that makes no difference to the final estimate.

some constant which depends on the size of the smallest image of an element of the partition of smoothness. We have shown that

$$\operatorname{Var}(\mathscr{L}_b h) \leq 2\,\|J\|_{\mathbf{L}^\infty} \operatorname{Var}(h) + (\|J'\|_{\mathbf{L}^\infty} + |b|\,\|\tau' \cdot J\|_{\mathbf{L}^\infty} + \|\phi'\|_{\mathbf{L}^\infty})\,\|h\|_{\mathbf{L}^1}.$$

This suffices[5] since we assumed that $\inf|f'| > 2$ and so $2\,\|J\|_{\mathbf{L}^\infty} < 1$. Consequently the above estimate may be iterated to produce an estimate for all $n \in \mathbb{N}$. □

These estimates and the compactness of the embedding $\mathbf{BV} \hookrightarrow \mathbf{L^1}(\mathbb{T}^1)$, by the usual arguments (see, for example, [21]), imply that the operator $\mathscr{L}_b : \mathbf{BV} \to \mathbf{BV}$ has spectral radius not greater than 1 and essential spectral radius not greater than $\lambda^{-1} \in (0,1)$. The spectral radius of $\mathscr{L}_0 : \mathbf{BV} \to \mathbf{BV}$ is equal to 1.

It is convenient to introduce the equivalent norm

$$\|h\|_{(b)} := (1+|b|)^{-1}\,\|h\|_{\mathbf{BV}} + \|h\|_{\mathbf{L^1}(\mathbb{T}^1)}.$$

The main purpose of this section is to prove the following result.

Proposition 1. *There exists $b_0 > 0$, $\rho > 0$, and $\gamma_2 > 0$ such that*

$$\|\mathscr{L}_b^{n(b)}\|_{(b)} \leq e^{-n(b)\gamma_2}, \qquad \textit{for all } |b| \geq b_0,\ n(b) := \lceil \rho \ln|b| \rceil.$$

The remainder of the section will be devoted to the proof of the above. The proof is self-contained apart from using Proposition 2 (see below), whose proof is postponed to Section 4.

Lemma 2. *For all $n \in \mathbb{N}$, $h \in \mathbf{BV}$, $b \in \mathbb{R}$,*

$$\|\mathscr{L}_b^n h\|_{(b)} \leq C_\lambda \lambda^{-n}\,\|h\|_{(b)} + (C_\lambda + 1)\,\|h\|_{\mathbf{L^1}(\mathbb{T}^1)}.$$

Proof. This is a direct result of the definition of the norm and the Lasota-Yorke estimate (Lemma 1). □

First we deal with the easy case when $\|h\|_{\mathbf{BV}}$ is large in comparison to $\|h\|_{\mathbf{L^1}(\mathbb{T}^1)}$. Let $n_0 := \lceil \ln(4C_\lambda)/\ln\lambda \rceil$.

Lemma 3. *Suppose that $h \in \mathbf{BV}$, satisfying $2(C_\lambda+1)(1+|b|)\,\|h\|_{\mathbf{L^1}(\mathbb{T}^1)} \leq \|h\|_{\mathbf{BV}}$. Then $\|\mathscr{L}_b^{n_0} h\|_{(b)} \leq \frac{3}{4}\,\|h\|_{(b)}$.*

Proof. The definition of $n_0 \in \mathbb{N}$ is such that $C_\lambda \lambda^{-n_0} + \frac{1}{2} \leq \frac{3}{4}$. The conclusion then follows from Lemma 2. □

This means that we only need to worry about estimating in the case where $2(C_\lambda + 1)(1+|b|)\,\|h\|_{\mathbf{L^1}(\mathbb{T}^1)} > \|h\|_{\mathbf{BV}}$. This is the case where the density can be considered to be "almost constant" as long as we look at the scale of $|b|^{-1}$. Furthermore it will suffice to estimate the $\mathbf{L^1}$ norm and not the $\mathbf{BV}$ norm as demonstrated by the following calculation. Using Lemma 2, for any $n \in \mathbb{N}$,

(2.2)
$$\begin{aligned}
\left\|\mathscr{L}_b^{2n} h\right\|_{(b)} &\leq C_\lambda \lambda^{-n}\,\|\mathscr{L}_b^n h\|_{(b)} + (C_\lambda+1)\,\|\mathscr{L}_b^n h\|_{\mathbf{L^1}(\mathbb{T}^1)} \\
&\leq C_\lambda^2 \lambda^{-2n}\,\|h\|_{(b)} + C_\lambda(C_\lambda+1)\lambda^{-n}\,\|h\|_{\mathbf{L^1}(\mathbb{T}^1)} + (C_\lambda+1)\,\|\mathscr{L}_b^n h\|_{\mathbf{L^1}(\mathbb{T}^1)} \\
&\leq 2C_\lambda(C_\lambda+1)\lambda^{-n}\,\|h\|_{(b)} + (C_\lambda+1)\,\|\mathscr{L}_b^n h\|_{\mathbf{L^1}(\mathbb{T}^1)}.
\end{aligned}$$

[5]By considering higher iterates of the same argument, if one were interested in optimal estimates, λ can be chosen arbitrarily close to $\limsup_{n\to\infty} |J_n|^{\frac{1}{n}}$.

It therefore remains to obtain exponential contraction of $\|\mathscr{L}_b^n h\|_{\mathbf{L^1}(\mathbb{T}^1)}$ in terms of $\|h\|_{(b)}$ in the case when $2(C_\lambda + 1)(1+|b|)\,\|h\|_{\mathbf{L^1}(\mathbb{T}^1)} > \|h\|_{\mathbf{BV}}$.

In order to later deal with discontinuities we now introduce a "growth lemma" suitable for this setting. Fix $\delta > 0$ such that, for any interval $\omega \subset \mathbb{T}^1$ of size $|\omega| \leq \delta$, the image $f\omega$ consists of at most two connected components. We will define unions of open intervals Ω_n for all $n \in \mathbb{N}$ iteratively. Let $\Omega_0 \subset \mathbb{T}^1$ be an interval, $|\Omega_0| \leq \delta$. Suppose that Ω_n is already defined. Let ω be one of the connected components of Ω_n. The image $f\omega$ is the union of intervals; some could be large, some could be small. It is convenient to maintain all intervals of size less than δ, and so we artificially chop long intervals so that they are always of size greater than $\delta/2$ and less than δ. In this fashion let the set $\{\omega_k\}_k$ be a set of open intervals which exhausts ω except for a zero measure set and such that each $f\omega_k$ is a single interval of length not greater than δ. The set Ω_{n+1} is defined to be a partition of Ω_n produced by following the same procedure for each connected component of Ω_n.

We must control the measure of points close to the boundaries of $\Omega_n = \{\omega_j\}_j$. For any $n \in \mathbb{N}$, $x \in \Omega_n$, let $r_n(x) := d(f^n(x), f^n(\partial\Omega_n))$, and hence let ($\mathbf{m}$ denotes Lebesgue measure)

$$\mathrm{Z}_\epsilon \Omega_n := \mathbf{m}(\{x \in \Omega_n : r_n(x) \leq \epsilon\}).$$

Let $\beta := \tilde{\lambda}/2$ and[6] let $C_\beta := 4\Lambda\beta\delta^{-1}\tilde{\lambda}^{-1}(\beta-1)^{-1}$.

Lemma 4. *For all $n \in \mathbb{N}$, $\epsilon > 0$,*

$$\mathrm{Z}_\epsilon \Omega_n \leq \beta^{-n}\tilde{\lambda}^n \mathrm{Z}_{\epsilon/\lambda^n}\Omega_0 + \epsilon C_\beta \left|\Omega_0\right|.$$

Proof. Suppose for the moment that Ω_n consists of just one element, i.e., $\Omega_n = \{\omega\}$. We will estimate $\mathrm{Z}_\epsilon\Omega_{n+1}$. The image $f\omega$ consists of at most two connected components. But some of these connected components could be large, in which case they will be cut into smaller pieces of size between $\delta/2$ and δ. The set $\partial\Omega_{n+1}$ consists of points which come from one of three different origins: from $\partial\Omega_n$; from a cut due to the discontinuities of the map; or from the artificial cuts. The first two possibilities are bounded by $2\mathrm{Z}_{\epsilon/\tilde{\lambda}}\Omega_n$. The total length of $f\omega$ is not greater than $\Lambda\mathbf{m}(\omega)$, and so the total number of artificial cuts is not greater than $2\delta^{-1}\Lambda\mathbf{m}(\omega)$. Summing these terms we obtain the estimate

$$\mathrm{Z}_\epsilon\Omega_{n+1} \leq 2\mathrm{Z}_{\epsilon/\tilde{\lambda}}\Omega_n + 4\epsilon\frac{\Lambda}{\delta\tilde{\lambda}}\mathbf{m}(\omega).$$

The equivalent estimate holds, even when Ω_n consists of more than one element. Recall that $2 = \beta^{-1}\tilde{\lambda}$, and so the above estimate reads as

$$\mathrm{Z}_\epsilon\Omega_{n+1} \leq \beta^{-1}\tilde{\lambda}\mathrm{Z}_{\epsilon/\tilde{\lambda}}\Omega_n + 4\epsilon\frac{\Lambda}{\delta\tilde{\lambda}}\mathbf{m}(\Omega_n),$$

and iteration produces the estimate (since $\sum_{j=0}^{\infty}\beta^{-n} = \frac{\beta}{\beta-1}$)

$$\mathrm{Z}_\epsilon\Omega_n \leq \beta^{-n}\tilde{\lambda}^n\mathrm{Z}_{\epsilon/\tilde{\lambda}^n}\Omega_0 + \epsilon\frac{4\Lambda\beta}{\delta\tilde{\lambda}(\beta-1)}\mathbf{m}(\Omega_0). \qquad \square$$

[6] The quantity β is the one which appears in the growth lemma and represents how expansion dominates over chopping (due to the discontinuities). In the present setting, because there are merely a finite number of discontinuities it suffices to define $\beta := \tilde{\lambda}/2$. Even if there were a countable number of discontinuities it is still possible to have a growth lemma of a similar form, but then β would need to be chosen more carefully.

The argument will depend crucially on the three quantities $\rho_1, \xi, \rho_2 > 0$. Let

$$\rho_1 := \frac{2}{\ln \tilde{\lambda}}, \qquad \xi := \frac{\ln \beta}{2 \ln \tilde{\lambda}}, \qquad \rho_2 := \frac{\xi}{2 \ln \Lambda}, \tag{2.3}$$

and hence let $n_1(b) := \lceil \rho_1 \ln |b| \rceil$, $n_2(b) := \lceil \rho_2 \ln |b| \rceil$. Let $n(b) := n_1(b) + n_2(b)$. For notational simplicity we will often suppress the dependence on b of n, n_1, n_2. Note that $\tilde{\lambda}\beta^{-1} = 2$ and so $\ln \beta < \ln \tilde{\lambda}$ and hence $\xi < \frac{1}{2}$. We use two time scales: The first n_1 iterates are for a small interval of length $|b|^{-(1+\xi)}$ to expand significantly to such an extent that the total size of the image is many times unit size and has been cut into pieces by the dynamics repeatedly. Then we take n_2 iterates to see oscillatory cancellations. The argument will also depend on the choice of $b_0 > 0$. At several points during the argument this quantity will be chosen sufficiently large.

Denote by $\{H_\ell\}_\ell$ the partition of $\mathbb{T}^1$ into subintervals of equal length such that

$$|b|^{-(1+\xi)} \leq |H_\ell| \leq 2\,|b|^{-(1+\xi)}. \tag{2.4}$$

We use this partition to approximate the density h. Denote by h_b the density which is constant on each H_ℓ and equal to the average value of h on H_ℓ. Note that $\|h\|_{\mathbf{L^1}(\mathbb{T}^1)} = \|h_b\|_{\mathbf{L^1}(\mathbb{T}^1)}$.

Lemma 5. *Let $h \in \mathbf{BV}$, $b \in \mathbb{R}$, and $|b| \geq b_0$ such that $2(C_\lambda+1)(1+|b|)\,\|h\|_{\mathbf{L^1}(\mathbb{T}^1)} > \|h\|_{\mathbf{BV}}$ and let h_b be the piecewise constant function as defined in the above paragraph. Then*

$$\|h - h_b\|_{\mathbf{L^1}(\mathbb{T}^1)} \leq 8C_\lambda e^{-n\frac{\xi}{2(\rho_1+\rho_2)}}\,\|h\|_{\mathbf{L^1}(\mathbb{T}^1)}.$$

Proof. Standard[7] approximation results for $\mathbf{BV}$ functions imply that $\|h - h_b\|_{\mathbf{L^1}(\mathbb{T}^1)} \leq 2\,|b|^{-(1+\xi)}\,\|h\|_{\mathbf{BV}}$ since $|H_\ell| \leq 2\,|b|^{-(1+\xi)}$. Substituting the control on $\|h\|_{\mathbf{BV}}$ which is assumed we have

$$\|h - h_b\|_{\mathbf{L^1}(\mathbb{T}^1)} \leq 4\,|b|^{-(1+\xi)}\,(C_\lambda+1)(1+|b|)\,\|h\|_{\mathbf{L^1}(\mathbb{T}^1)}.$$

Ensuring that $b_0 > 1$ we obtain $(1+|b|) \leq 2\,|b|$. Increasing b_0 more if required we may assume that $n(b) \leq 2(\rho_1+\rho_2)\ln|b|$. This means that $|b|^{-\xi} \leq e^{-n(b)\frac{\xi}{2(\rho_1+\rho_2)}}$. Consequently $\|h - h_b\|_{\mathbf{L^1}(\mathbb{T}^1)} \leq 8(C_\lambda+1)e^{-n(b)\frac{\xi}{2(\rho_1+\rho_2)}}\,\|h\|_{\mathbf{L^1}(\mathbb{T}^1)}$. ∎

Using Lemma 5 we know that in the case $2(C_\lambda+1)(1+|b|)\,\|h\|_{\mathbf{L^1}(\mathbb{T}^1)} > \|h\|_{\mathbf{BV}}$,

$$\|\mathscr{L}_b^n h\|_{\mathbf{L^1}(\mathbb{T}^1)} \leq \|\mathscr{L}_b^n h_b\|_{\mathbf{L^1}(\mathbb{T}^1)} + 8(C_\lambda+1)e^{-n\frac{\xi}{2(\rho_1+\rho_2)}}\,\|h\|_{\mathbf{L^1}(\mathbb{T}^1)},$$

since $h_b = \sum_\ell h_b \mathbf{1}_\ell$ where h_b is constant on each interval H_ℓ, and that $\|h\|_{\mathbf{L^1}(\mathbb{T}^1)} = \|h_b\|_{\mathbf{L^1}(\mathbb{T}^1)}$. We now take advantage of the following result. This is the main estimate which takes advantage of the oscillatory cancellation mechanism which is present in this setting.

Proposition 2. *There exists $C_3 > 0$, $\gamma_3 > 0$ such that, for all $|b| \geq b_0$ and ℓ,*

$$\|\mathscr{L}_b^{n(b)} \mathbf{1}_{H_\ell}\|_{\mathbf{L^1}(\mathbb{T}^1)} \leq C_3 e^{-n(b)\gamma_3}\,\|\mathbf{1}_{H_\ell}\|_{\mathbf{L^1}(\mathbb{T}^1)}.$$

---

[7] By [13, §5.2.2], since $h \in \mathbf{BV}$ then for any $\epsilon > 0$ there exists $\tilde{h} \in \mathscr{C}^1$ be such that $\|h - \tilde{h}\|_{\mathbf{L^1}(\mathbb{T}^1)} \leq \epsilon$ and $|\|\tilde{h}'\|_{\mathbf{L^1}(\mathbb{T}^1)} - \|h\|_{\mathbf{BV}}| \leq \epsilon$. For each ℓ let $x_\ell \in H_\ell$ be such that $\tilde{h}(x_\ell) = \int_{H_\ell} \tilde{h}(y)\,dy$. For every $x \in H_\ell$ we have the estimate $|\tilde{h}(x) - \tilde{h}(x_\ell)| \leq \int_{H_\ell} |\tilde{h}'|(y)\,dy$. This means that $\|\tilde{h} - \tilde{h}_b\|_{\mathbf{L^1}(\mathbb{T}^1)} = \sum_\ell \int_{H_\ell} |\tilde{h}(x) - \tilde{h}(x_\ell)|\;dx \leq \sup_\ell |H_\ell| \int_{\mathbb{T}^1} |\tilde{h}'|(y)\,dy$. Using the above approximation argument (to replace h with $\tilde{h}$) this implies the required estimate.

The proof of the above is postponed to Section 4.

Combining Lemma 5 and Proposition 2 we obtain the estimate

$$\begin{aligned}\|\mathscr{L}_b^{n(b)}h\|_{\mathbf{L^1}(\mathbb{T}^1)} &\leq \left(8(C_\lambda+1)e^{-n(b)\frac{\xi}{2(\rho_1+\rho_2)}} + C_3 e^{-n(b)\gamma_3}\right)\|h\|_{\mathbf{L^1}(\mathbb{T}^1)}\\ &\leq C_4 e^{-n(b)\gamma_4}\|h\|_{\mathbf{L^1}(\mathbb{T}^1)}\end{aligned}$$

where $\gamma_4 := \min(\frac{\xi}{2(\rho_1+\rho_2)}, \gamma_3)$ and $C_4 := 8(C_\lambda+1)+C_3$. We now substitute these estimates into (2.2):

$$\begin{aligned}\|\mathscr{L}_b^{2n(b)}h\|_{(b)} &\leq \left(2C_\lambda(C_\lambda+1)\lambda^{-n(b)} + (C_\lambda+1)C_4e^{-n(b)\gamma_4}\right)\|h\|_{(b)}\\ &\leq (C_\lambda+1)(2C_\lambda+C_4)e^{-n(b)\gamma_5}\|h\|_{(b)},\end{aligned}$$

where $\gamma_5 := \min(\ln\lambda, \gamma_4)$. To complete the proof of Proposition 1 we must combine the above estimate with Lemma 3. We choose $b_0 > 0$ sufficiently large such that

$$\|\mathscr{L}_b^{2n(b)}h\|_{(b)} \leq e^{-n(b)\frac{\gamma_5}{2}}\|h\|_{(b)}$$

for all $|b| \geq b_0$. Note that the estimate of Lemma 3 cannot be simply iterated since the assumption of the estimate is not invariant. However we can argue as follows: Either the estimate can be interated or the above estimate applies. Consequently we obtain the exponential rate as required and complete the proof of Proposition 1.

3. Transversality

The purpose of this section is to prove Proposition 3 (see below). This is a crucial estimate which will be required in Section 4. We may assume that $\sup|\tau'| > 0$ since if this does not hold, then τ is actually equal to a piecewise constant function and, in particular, is cohomologous to a piecewise constant function. The first step is to define a forward invariant unstable conefield. Let

$$C_1 := \frac{2\sup|\tau'|}{\tilde{\lambda}-1} > 0.$$

Define the constant conefield with the cones $\mathcal{K} = \{\left(\begin{smallmatrix}\alpha\\ \beta\end{smallmatrix}\right) : |\frac{\beta}{\alpha}| \leq C_1\}$. This conefield is strictly invariant under

$$DF(x) = \left(\begin{smallmatrix}f'(x) & 0\\ \tau'(x) & 1\end{smallmatrix}\right).$$

To see the invariance note that $DF(x) : \left(\begin{smallmatrix}\alpha\\ \beta\end{smallmatrix}\right) \mapsto \left(\begin{smallmatrix}\alpha'\\ \beta'\end{smallmatrix}\right)$ where $\frac{\beta'}{\alpha'} = (\tau'(x) + \frac{\beta}{\alpha})/f'(x)$.

Let $x_1, x_2 \in \mathbb{T}^1$ be two preimages[8] of some $y \in \mathbb{T}^1$, i.e., $f^n(x_1) = f^n(x_2) = y$. We write $x_1 \pitchfork x_2$ (meaning *transversal*) if $DF^n_{x_1}\mathcal{K} \cap DF^n_{x_2}\mathcal{K} = \{0\}$. Note that this transversality depends on n even though the dependence is suppressed in the notation. Define the quantity[9]

$$\varphi(n) := \sup_{y\in\mathbb{T}^1}\ \sup_{x_1\in f^{-n}(y)} \sum_{\substack{x_2\in f^{-n}(y)\\ x_1 \not\pitchfork x_2}} J_n(x_2).$$

This crucial quantity gives control on the fraction of preimages which are not transversal. In this section we prove the following, which is an extension of Tsujii [31, Theorem 1.4] to the present situation where discontinuities are permitted.

[8]Recall that f is defined on the open subset $\bigcup_k \mathbb{I}_k \subset \mathbb{T}^1$. As such if $x \in f^{-n}(y)$, then DF^n_x is well defined.

[9]It is useful to compare the definition of $\varphi(n)$ with $\mathscr{L}_0^n 1(y) = \sum_{x\in f^{-n}(y)} J_n(x)$ (the transfer operator $\mathscr{L}_b$ is defined in (2.1)).

Much of the argument follows the reasoning of the above mentioned reference with some changes due to the more general setting.

Proposition 3. *Let $F : (x,u) \mapsto (f(x), u + \tau(x))$ be a piecewise-$\mathscr{C}^2$ skew product over an expanding base map as above.* ***Either:***

$$\limsup_{n\to\infty} \varphi(n)^{\frac{1}{n}} < 1 \tag{3.1}$$

Or: *There exists Lipschitz $\theta : \mathbb{T}^1 \to \mathbb{R}$ and piecewise constant $\chi : \mathbb{T}^1 \to \mathbb{R}$ such that $\tau = \theta \circ f - \theta + \chi$. Moreover the discontinuities of χ occur only at points where either f or τ is discontinuous.*

Before proving the above, let us record a consequence of the transversality. Let $\tau_n := \sum_{j=0}^{n-1} \tau \circ f^j$.

Lemma 6. *If $f^n(x_1) = f^n(x_2)$ and $x_1 \pitchfork x_2$, then*

$$|(\tau_n' \cdot J_n)(x_1) - (\tau_n' \cdot J_n)(x_2)| > C_1(J_n(x_1) + J_n(x_2)).$$

Proof. Assume that $\frac{\tau_n'}{(f^n)'}(x_1) \geq \frac{\tau_n'}{(f^n)'}(x_2)$, the other case being analogous. Note that

$$DF^n(x_1)\left(\begin{smallmatrix} 1 \\ -C_1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} (f^n)'(x_1) \\ \tau_n'(x_1) - C_1 \end{smallmatrix}\right), \quad DF^n(x_2)\left(\begin{smallmatrix} 1 \\ C_1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} (f^n)'(x_2) \\ \tau_n'(x_2) + C_1 \end{smallmatrix}\right).$$

Transversality implies that $(\tau_n'(x_1) - C_1)/(f^n)'(x_1) > (\tau_n'(x_2) + C_1)/(f^n)'(x_2)$. $\square$

The remainder of this section is devoted to the proof of Proposition 3. As mentioned in the introduction it is known that there exists an f-invariant probability measure ν which is equivalent to Lebesgue; i.e., the invariant density is bounded and bounded away from zero. Let h_ν denote the density of ν. It is convenient to introduce the quantity

$$\tilde{\varphi}(n, L, y) := \sum_{\substack{x \in f^{-n}(y) \\ DF^n(x)\mathcal{K} \supset L}} J_n(x) \cdot \frac{h_\nu(x)}{h_\nu(y)} \tag{3.2}$$

where $L \in \mathbb{RP}^1$ (an element of real projective space, i.e., a line in $\mathbb{R}^2$ which passes through the origin). Let $\tilde{\varphi}(n) := \sup_y \sup_L \tilde{\varphi}(n, L, y)$. The benefit of this definition is that $\tilde{\varphi}(n)$ is submultiplicative, i.e., $\tilde{\varphi}(n+m) \leq \tilde{\varphi}(n)\tilde{\varphi}(m)$ for all $n, m \in \mathbb{N}$, and $\tilde{\varphi}(n) \leq 1$ for all $n \in \mathbb{N}$.

Lemma 7. *The following statements are equivalent.*

(i) $\limsup_{n\to\infty} \varphi(n)^{\frac{1}{n}} = 1$.

(ii) $\lim_{n\to\infty} \tilde{\varphi}(n)^{\frac{1}{n}} = 1$.

(iii) *For all $n \in \mathbb{N}$, $y \in \mathbb{T}^1$ there exists $L_n(y) \in \mathbb{RP}^1$ such that, for every $x \in f^{-n}(y)$, $DF^n(x)\mathcal{K} \supset L_n(y)$.*

(iv) *There exists a measurable F-invariant unstable direction; i.e., there exists $\ell : \mathbb{T}^1 \to \mathbb{R}$ such that $\tau' = f' \cdot \ell \circ f - \ell$ and so*

$$DF(x)\left(\begin{smallmatrix} 1 \\ \ell(x) \end{smallmatrix}\right) = f'(x)\left(\begin{smallmatrix} 1 \\ \ell \circ f(x) \end{smallmatrix}\right).$$

(v) *Statement* (iv) *holds with ℓ of bounded variation.*

(vi) *There exists $\theta : \mathbb{T}^1 \to \mathbb{T}^1$ such that $\tau - \theta \circ f + \theta$ is piecewise constant (discontinuities only where either f or τ is discontinuous). Moreover θ is differentiable with derivative of bounded variation.*

Note that $\varphi(n)$ is uniformly bounded; i.e., there exists some $C > 0$ such that $\varphi(n) \leq C$ for all $n \in \mathbb{N}$. This can be seen by observing that $\varphi(n) \leq \|\mathscr{L}_0^n 1\|_{\mathbf{L}^\infty} \leq \|\mathscr{L}_0^n 1\|_{\mathbf{BV}}$ and using Lemma 1. This means that $\limsup_{n\to\infty} \varphi(n)^{\frac{1}{n}} \leq 1$, and hence the above lemma immediately implies Proposition 3. In the remainder of this section we prove the above lemma. First a simple fact that we will use repeatedly.

Lemma 8. $|J_n \cdot \tau_n'| \leq \frac{1}{2}C_1$.

Proof. First observe that $\tau_n' = \sum_{i=0}^{n-1} \tau' \circ f^i \cdot (f^i)'$. Consequently $|J_n \cdot \tau_n'| \leq |\tau'| \sum_{i=1}^{n} \tilde{\lambda}^{-i}$. For all $n \in \mathbb{N}$ the sum $\sum_{i=1}^{n} \tilde{\lambda}^{-i}$ is bounded from above by $(\tilde{\lambda}-1)^{-1}$. And so, using also the definition of C_1, we know that $|J_n \cdot \tau_n'| \leq \sup|\tau'|/(\tilde{\lambda}-1) = \frac{1}{2}C_1$. □

Proof of (i) $\Longrightarrow$ (ii). Suppose that $m \in \mathbb{N}$, $n = n(m) = \lceil 2\frac{\ln \Lambda}{\ln \tilde{\lambda}} m \rceil$, $y \in \mathbb{T}^1$ and $x_1, x_2 \in f^{-n}(y)$. Note that $n > m$ since $\Lambda \geq \tilde{\lambda}$. Let $p = n - m$. Further suppose that

$$DF^n(x_1)\mathcal{K} \cap DF^n(x_2)\mathcal{K} \neq \{0\}.$$

The slopes of the edges of $DF^n(x_1)\mathcal{K}$ are $\frac{\tau_n'}{(f^n)'}(x_1) \pm C_1 J_n(x_1)$. Let

$$L(x_1) := DF^n(x_1)(\mathbb{R} \times \{0\}).$$

The slope of L is $\frac{\tau_n'}{(f^n)'}(x_1)$. Since we assume the cones $DF^n(x_1)\mathcal{K}$ and $DF^n(x_2)\mathcal{K}$ are not transversal this implies that the difference in slope between one of the edges of $DF^n(x_2)\mathcal{K}$ and L is not greater than

$$C_1 J_n(x_1) \leq C_1 \tilde{\lambda}^{-n}. \tag{3.3}$$

Now consider the cone $DF^n(x_2)\mathcal{K}$ and the cone $DF^m(f^p x_2)\mathcal{K} \supset DF^n(x_2)\mathcal{K}$. The slopes of the edges of the first are

$$\frac{\tau_n'}{(f^n)'}(x_2) \pm C_1 J_n(x_2) = \frac{\tau_m'}{(f^m)'} \circ f^p(x_2) + \frac{\tau_p'}{(f^m)' \circ f^p \cdot (f^p)'}(x_2) \pm C_1 J_n(x_2),$$

whilst the slopes of the edges of the second are

$$\frac{\tau_m'}{(f^m)'} \circ f^p(x_2) \pm C_1 J_m \circ f^p(x_2).$$

Consequently the slopes of the edges of the two cones are separated by at least

$$J_m \circ f^p(x_2) \left(C_1 - \sup\left|\tau_p' \cdot J_p\right|\right) - C_1 J_n(x_2).$$

By Lemma 8 we know that $\left|\tau_p' \cdot J_p\right| \leq \frac{1}{2}C_1$. This means that the above term is bounded from below by

$$\tfrac{1}{2}C_1 \Lambda^{-m} - C_1 \tilde{\lambda}^{-n} \geq \tfrac{1}{2}C_1 \tilde{\lambda}^{-\frac{n}{2}} - C_1 \tilde{\lambda}^{-n},$$

where we used that the assumed relation between n and m implies that $m \leq \frac{n}{2}\frac{\ln \tilde{\lambda}}{\ln \Lambda}$ and so $\Lambda^{-m} \geq \tilde{\lambda}^{-\frac{n}{2}}$. Recall now (3.3). For all n sufficiently large $\frac{1}{2}C_1 \tilde{\lambda}^{-\frac{n}{2}} - C_1 \tilde{\lambda}^{-n} \geq C_1 \tilde{\lambda}^{-n}$. To conclude, we have shown that $DF^n(x_1)\mathcal{K} \cap DF^n(x_2)\mathcal{K} \neq \{0\}$

implies that $DF^m(f^p x_2)\mathcal{K} \supset L(x_1)$ where $L(x_1)$ is defined as before. This means that (let $z = f^p x_2$):

$$\sum_{\substack{x_2 \in f^{-n}(y) \\ x_1 \not\pitchfork x_2}} J_n(x_2) \leq \sum_{\substack{x_2 \in f^{-n}(y) \\ DF^m(f^p x_2)\mathcal{K} \supset L(x_1)}} J_m(f^p x_2) \cdot J_p(x_2)$$

$$\leq \sum_{\substack{z \in f^{-m}(y) \\ DF^m(z)\mathcal{K} \supset L(x_1)}} J_m(z) \sum_{x_2 \in f^{-p}(z)} J_p(x_2).$$

Finally this implies that $\varphi(n) \leq C_2 \tilde{\varphi}(m(n))$, where $C_2 := \sup h_\nu / \inf h_\nu > 0$. □

Proof of (ii) $\Longrightarrow$ (iii). By submultiplicativity and the fact that $\tilde{\varphi}(n) \leq 1$ for all $n \in \mathbb{N}$ the assumption $\lim_{n\to\infty} \tilde{\varphi}(n)^{\frac{1}{n}} = 1$ implies that $\tilde{\varphi}(n) = 1$ for all $n \in \mathbb{N}$. Consequently the following statement holds:

For all n there exists $y_n \in \mathbb{T}^1$ and $L_n \subset \mathbb{RP}^1$ such that, for all $x \in f^{-n}(y_n)$, $DF^n(x)\mathcal{K} \supset L_n$.

It remains to prove that the above statement implies statement (iii). We will prove the contrapositive. Suppose the negation of statement (iii); i.e., there exists $n_0 \in \mathbb{N}$, $y \in \mathbb{T}^1$, $x_1, x_2 \in f^{-n_0}(y)$ such that $DF^{n_0}(x_1)\mathcal{K} \cap DF^{n_0}(x_2)\mathcal{K} = \{0\}$. Let g_1, g_2 denote the two inverse maps corresponding to x_1, x_2. These inverses are defined on some interval containing y, and due to the openness of the transversality of cones we can assume that $DF^{n_0}(g_1(y))\mathcal{K} \cap DF^{n_0}(g_2(y))\mathcal{K} = \{0\}$ for all $y \in \omega_*$ where $\omega_* \subset \mathbb{T}^1$ is an open interval. Since f is covering there exists $m_0 \in \mathbb{N}$ such that $f^{m_0}(\omega_*) = \mathbb{T}^1$. Let $m = m_0 + n_0$. For all $y \in \mathbb{T}^1$ there exists $z \in f^{-m_0}(y)$, and there exists $x_1, x_2 \in f^{-n_0}(z)$ with the above transversality property. This means that for all $y \in \mathbb{T}^1$ there exist $x_1, x_2 \in f^{-m}(y)$ such that $DF^m(x_1)\mathcal{K} \cap DF^m(x_2)\mathcal{K} = \{0\}$, since

$$DF^m(x_1)\mathcal{K} \cap DF^m(x_2)\mathcal{K} = DF^{m_0}(y)(DF^{n_0}(x_1)\mathcal{K} \cap DF^{n_0}(x_2)\mathcal{K}).$$

This contradicts the above statement concerning the existence of some L_n such that $DF^n(x)\mathcal{K} \supset L_n$ for all $x \in f^{-n}(y)$. □

Proof of (iii) $\Longrightarrow$ (iv). For all $x \in \mathbb{T}^1$ let $\ell_n(x)$ denote the slope of $L_n(x)$, i.e., $\binom{1}{\ell_n(x)} \in L_n(x)$. The uniform expansion means that the image of unstable cones contracts, and consequently for each x then $\ell_n(x) \to \ell(x)$ as $n \to \infty$. The function $\ell(x)$ enjoys the property that $\tau_n'(x) + \ell(x) = (f^n)'(x) \cdot \ell(f^n x)$. □

Proof of (iv) $\Longleftrightarrow$ (v). The implication (v) $\Longrightarrow$ (iv) is immediate. Assume that statement (iv) holds. Since ℓ is invariant we know that for any $n \in \mathbb{N}$, $x \in f^{-n}(y)$ that

$$\ell(y) = \frac{\tau_n'}{(f^n)'}(x) + \frac{\ell}{(f^n)'}(x).$$

For large n the second term on the right hand side becomes very small. Note that because we assume (iv) holds, if we want to calculate ℓ at y it does not matter which preimage x we consider. Fix some $\omega_0 \subset \mathbb{T}^1$ a disjoint union of intervals and a bijection $g : \mathbb{T}^1 \to \omega_0$ such that $f \circ g$ is the identity. We can do this in such a way that g is $\mathscr{C}^2$ on each component of ω_0.[10] Of course $f^n \circ g^n = \mathbf{id}$ for all $n \in \mathbb{N}$.

[10]Note that if the map f was full branch (or if there was at least one smooth onto component of f) we could choose ω_0 and $g : \mathbb{T}^1 \to \omega_0$ such that g is $\mathscr{C}^2$, but this cannot be expected in general.

Consequently

$$\begin{aligned}\ell &= \frac{\tau_n'}{(f^n)'} \circ g^n + \frac{\ell}{(f^n)'} \circ g^n \\ &= \sum_{j=0}^{n-1} \frac{\tau'}{(f^{n-j})'} \circ g^{n-j} + \frac{\ell}{(f^n)'} \circ g^n.\end{aligned} \tag{3.4}$$

Note that $\|\frac{\ell}{(f^n)'} \circ g^n\|_{\mathbf{L}^\infty(\mathbb{T}^1)} \to 0$ as $n \to \infty$. We will show that $\sum_{j=1}^{\infty} \frac{\tau'}{(f^j)'} \circ g^j$ is of bounded variation. Each term in this infinite sum is piecewise $\mathscr{C}^1$ and has only a finite number of discontinuities. The function g may possess a finite number of discontinuities, but since it is one-to-one there exists some $C > 0$ such that the number of discontinuities of g^j is not greater than Cj for all $j \in \mathbb{N}$. We estimate

$$\|\tfrac{\tau'}{(f^j)'} \circ g^j\|_{\mathbf{BV}} \leq \|\tau'\|_{\mathbf{BV}} \|\tfrac{1}{(f^j)'} \circ g^j\|_{\mathbf{BV}} + Cj \, \|\tau'\|_{\mathbf{L}^\infty} \|\tfrac{1}{(f^j)'}\|_{L^\infty}$$

and observe that $\|\frac{1}{(f^j)'} \circ g^j\|_{\mathbf{BV}} \leq \|\frac{1}{(f^j)'}\|_{\mathbf{L}^\infty} + \operatorname{Var}(\frac{1}{(f^j)'} \circ g^j)$. Since $\|\frac{1}{(f^j)'}\|_{L^\infty}$ decreases exponentially it remains to estimate $\operatorname{Var}(\frac{1}{(f^j)'} \circ g^j)$. Following closely the argument used in the proof of Lemma 1 we observe that

$$\operatorname{Var}(\tfrac{1}{(f^j)'} \circ g^j) \leq \sup_{|\eta| \leq 1} \int_{\mathbb{T}^1} \tfrac{1}{(f^j)'} \circ g^j \cdot \eta' \leq \sup_{|\eta| \leq 1} \int_{g^j(\mathbb{T}^1)} \eta' \circ f^j.$$

Again, as in the proof of Lemma 1, we note that $\eta' \circ f^k = (\eta \circ f^k / f^k)' + \eta \circ f^k \cdot (f^k)'/(f^k)^2$. The difference now is that the part of the estimate corresponding to $\|\eta \circ f^k \cdot (f^k)'/(f^k)^2\|_{\mathbf{L}^1(g^j(\mathbb{T}^1))}$ can be bounded by the measure of the set $g^j(\mathbb{T}^1)$ and this decreases exponentially. This all means that the $\mathbf{BV}$ norm of the term $\frac{\tau'}{(f^j)'} \circ g^j$ is exponentially decreasing with j and so the sum converges in $\mathbf{BV}$. Consequently ℓ must be of bounded variation. □

Proof of (v) $\iff$ (vi). First we prove (v) $\implies$ (vi). For all $y \in \mathbb{T}^1$ let

$$\theta(y) := \int_0^y \ell(x) \, dx.$$

This defines a Lipschitz function on $\mathbb{T}^1$, differentiable in the sense that the derivative is of bounded variation. There exists a partition $\{\omega_m\}_m$ such that τ and f are $\mathscr{C}^2$ when restricted to each element of the partition. Write $\omega_m = (a_m, b_m)$. If $y \in \omega_m$, then $\tau(y) = \tau(a_m) + \int_{a_m}^y \tau'(x) \, dx$. Substituting the equation $\tau' = f' \cdot \ell \circ f - \ell$ we obtain

$$\begin{aligned}\tau(y) &= \tau(a_m) + \int_{a_m}^y f' \cdot \ell \circ f(x) \, dx - \int_{a_m}^y \ell(x) \, dx \\ &= \tau(a_m) + \int_{f(a_m)}^{f(y)} \ell(x) \, dx - \int_{a_m}^y \ell(x) \, dx \\ &= \theta \circ f(y) - \theta(y) + (\theta(a_m) - \theta \circ f(a_m) + \tau(a_m)) \\ &= \theta \circ f(y) - \theta(y) + \chi_m.\end{aligned}$$

Let χ denote the piecewise constant function equal to χ_m on each ω_m. The implication (vi) $\implies$ (v) follows by differentiating $\tau - \theta \circ f + \theta = \chi$. □

Proof of (iv) $\implies$ (i). Suppose that (iv) holds. This means that there exists $\ell : \mathbb{T}^1 \to \mathbb{R}$ such that $\tau' = f' \cdot \ell \circ f - \ell$. Using the formula (3.4) for ℓ we know that $\sup|\ell| \leq \frac{3}{2} \sum_{k=1}^{\infty} \frac{\sup|\tau|}{\tilde{\lambda}^k} = \frac{3}{2} \frac{\sup|\tau|}{\tilde{\lambda}-1} \leq C_1$. This means that the vector $\left(\begin{smallmatrix}1 \\ \ell(x)\end{smallmatrix}\right) \in \mathcal{K}$

for each x. Consequently the vector $\binom{1}{\ell(y)}$ is contained within $DF^n(x)\mathcal{K}$ for all $x \in f^{-n}(y)$ since $DF^n(x)\binom{1}{\ell(x)} = (f^n)'(x)\binom{1}{\ell(y)}$. The important point is that this is the same vector $\binom{1}{\ell(y)}$, irrespective of which preimage $x \in f^{-n}(y)$ is being considered. This has the implication that $x_1 \not\pitchfork x_2$, i.e., $DF^n(x_1)\mathcal{K} \cap DF^n(x_2)\mathcal{K} \neq \{0\}$, for every $x_1, x_2 \in f^{-n}(y)$. In other words, none of the preimages are transversal. This in turn implies that, in this special case, $\varphi(n) = \sup_{y\in\mathbb{T}^1}\sup_{x_1\in f^{-n}(y)}\sum_{x_2\in f^{-n}(y)} J_n(x_2)$. Here we use again the connection to $\mathscr{L}_0^n 1$ and that this quantity is uniformly bounded away from zero for all large $n \in \mathbb{N}$ since $\mathscr{L}_0^n 1$ converges in the $\mathbf{BV}$ norm to the invariant density which is bounded away from zero and the $\mathbf{BV}$ norm dominates $\mathbf{L}^\infty$. □

4. The main estimate

This section is devoted to the proof of Proposition 2, which was stated in Section 2. Throughout this section we assume that the first alternative of Proposition 3 holds. Let $\gamma_1 := \limsup_{n\to\infty} -\frac{1}{n}\log\varphi(n) > 0$, and fix $\gamma \in (0, \gamma_1)$. There exists $C_\gamma > 0$ such that

$$\varphi(n) \leq C_\gamma e^{-n\gamma} \quad \text{for all } n \in \mathbb{N}. \tag{4.1}$$

In order to prove Proposition 2 we must estimate $\|\mathscr{L}_b^{n(b)}\mathbf{1}_\Omega\|_{\mathbf{L}^1(\mathbb{T}^1)}$ where Ω is an interval such that $|b|^{-(1+\xi)} \leq |\Omega| \leq 2\,|b|^{-(1+\xi)}$. Let $\Omega_0 = \Omega$ and, using the notation of Lemma 4, denote by $\{\omega_j\}_j$ the connected components of Ω_n. Let $h_j := f^{n(b)}\big|_{\omega_j}^{-1}$. Note that $\|\mathbf{1}_\Omega\|_{\mathbf{L}^1(\mathbb{T}^1)} = |\Omega|$. We must estimate

$$\|\mathscr{L}_b^n\mathbf{1}_\Omega\|_{\mathbf{L}^1(\mathbb{T}^1)} = \int_{\mathbb{T}^1}\left|\sum_j \left(J_n\cdot e^{ib\tau_n}\right)\circ h_j(z)\cdot\mathbf{1}_{f^n\omega_j}(z)\right|\,dz. \tag{4.2}$$

Introduce a partition of $\mathbb{T}^1$ into equal sized subintervals $\{I_p\}_p$ such that

$$|b|^{-(1-\xi)} \leq |I_p| \leq 2\,|b|^{-(1-\xi)}. \tag{4.3}$$

For each p, fix some $y_p \in I_p$ as a reference. To proceed we would like to ensure that the subintervals $f^n\omega_j$ make full crossings of the intervals I_p. For each p let G_p denote the set of indexes j such that $f^n\omega_j \supset I_p$. Let $G_p^{\complement}$ denote the complement of G_p. The integrals associated to indexes in the set $G_p^{\complement}$ are estimated as follows:

$$\sum_p\int_{I_p}\left|\sum_{j\in G_p^{\complement}}\left(J_n\cdot e^{ib\tau_n}\right)\circ h_j(z)\cdot\mathbf{1}_{f^n\omega_j}(z)\right|\,dz \leq \sum_p\sum_{j\in G_p^{\complement}}\left|\omega_j\cap f^{-n}I_p\right|.$$

That $j \in G_p^{\complement}$ implies that one of the end points of $f^n\omega_j$ is contained within I_p. Consequently $\omega_j \cap f^{-n}I_p$ is contained within the set $\{x \in \Omega_n : r_n(x) < \epsilon\}$ where $\epsilon = |I_p| \leq 2\,|b|^{-(1-\xi)}$. This means that

$$\sum_p\sum_{j\in G_p^{\complement}}\left|\omega_j\cap f^{-n}I_p\right| \leq \mathrm{Z}_\epsilon\Omega_n.$$

Applying the estimate of Lemma 4 gives a bound of

$$\begin{aligned}\mathrm{Z}_\epsilon\Omega_n &\leq \beta^{-n}\lambda^n\frac{2\epsilon}{\lambda^n} + \epsilon C_\beta\,|\Omega| \\ &\leq 8\,|b|^{-(1+\xi)}\left(e^{-n\ln\beta}e^{n\frac{2\xi}{\rho_1+\rho_2}} + 2C_\beta e^{-n\frac{1-\xi}{\rho_1+\rho_2}}\right).\end{aligned}$$

Recalling the definitions of ξ and ρ_1, note that $\frac{2\xi}{\rho_1+\rho_2} < \frac{2\xi}{\rho_1} = \xi \ln \lambda$ and so $\ln \beta - \frac{2\xi}{\rho_1+\rho_2} > \ln \beta - \xi \ln \lambda > 0$. Let $\gamma_6 := \min(\ln \beta - \frac{2\xi}{\rho_1+\rho_2}, \frac{1-\xi}{\rho_1+\rho_2}) > 0$, $C_5 := 8(1+2C_\beta)$. This means that

$$\sum_p \int_{I_p} \Big| \sum_{j \in G_p^{\complement}} \left(J_n \cdot e^{ib\tau_n}\right) \circ h_j(z) \cdot \mathbf{1}_{f^n \omega_j}(z) \Big| \ dz \leq |\Omega| \, C_5 e^{-\gamma_6 n}. \tag{4.4}$$

Now we may proceed to estimate (4.2) summing only over the indexes $j \in G_p$. Since $|\sum_k a_k|^2 = \sum_{jk} a_j \overline{a_k}$, using also Jensen's inequality, we have

(4.5)

$$\begin{aligned} \sum_p \int_{I_p} \Big| \sum_{j \in G_p} \left(J_n \cdot e^{ib\tau_n}\right) \circ h_j(z) \Big| \ dz &= \sum_p \int_{I_p} \left(\sum_{j,k \in G_p} (K_{j,k} \cdot e^{ib\theta_{j,k}})(z) \right)^{\frac{1}{2}} dz \\ &\leq \left(\sum_p \sum_{j,k \in G_p} \left| \int_{I_p} (K_{j,k} \cdot e^{ib\theta_{j,k}})(z) \ dz \right| \right)^{\frac{1}{2}} \end{aligned}$$

where $K_{j,k} := J_n \circ h_j \cdot J_n \circ h_k$ and we define the following crucial quantity related to the phase difference between different preimages of the same point:

$$\theta_{j,k}(x) := \left(\tau_n \circ h_j - \tau_n \circ h_k\right)(x).$$

Lemma 9. *There exists $C_6 > 0$ such that $J_n'(x) \leq C_6$ for all $x \in \mathbb{T}^1$, $n \in \mathbb{N}$.*

Proof. Note that $J_n = \prod_{j=0}^{n-1} \frac{1}{f'} \circ f^j$. Consequently $J_n' = \sum_{j=0}^{n-1} \frac{f''}{f'} \circ f^j \cdot J_{n-j} \circ f^j$. And so $|J_n'| \leq \sup |f''| / (\lambda - 1)$ for any $n \in \mathbb{N}$. □

Lemma 10. *There exists $C_7 > 0$, independent of $n \in \mathbb{N}$, such that $|\theta_{j,k}''| \leq C_7$.*

Proof. Suppose that $g : \mathbb{T}^1 \to \omega$ such that $g \circ f^n = \mathbf{id}$. Let $g^{(j)} := f^{n-j} \circ g$. Note that

$$(\tau_n \circ g)' = \sum_{j=0}^{n-1} (\tau' \cdot J_j) \circ g^{(j)}.$$

Consequently

$$(\tau_n \circ g)'' = \sum_{j=0}^{n-1} \left(J_j^2 \cdot \tau'' + \tau' \cdot J_j' \cdot J_j\right) \circ g^{(j)}.$$

By Lemma 9 we know that $J_n' \leq C_6$. Since τ is $\mathscr{C}^2$ and $J_n \leq \lambda^n$ the above term is uniformly bounded for any $n \in \mathbb{N}$. □

Let $g_j := f^{n_1} \circ h_j$. For each p let A_p denote the set of pairs $(j,k) \in G_p \times G_p$ such that[11] $g_j(y_p) \pitchfork g_k(y_p)$ (this is the case where we see oscillatory cancellations since the two preimages are transversal at iterate n_2). Let $A_p^{\complement}$ denote the complement set, i.e., the set of pairs (j,k) such that $g_j(y_p) \not\pitchfork g_k(y_p)$:

$$\begin{aligned} \sum_{j,k \in G_p} \left| \int_{I_p} (K_{j,k} \cdot e^{ib\theta_{j,k}})(z) \ dz \right| &\leq \sum_j \sum_{k:(j,k) \in A_p} \left| \int_{I_p} (K_{j,k} \cdot e^{ib\theta_{j,k}})(z) \ dz \right| \\ &\quad + \sum_j \sum_{k:(j,k) \in A_p^{\complement}} \int_{I_p} K_{j,k}(z) \ dz. \end{aligned} \tag{4.6}$$

[11] Recall that at the beginning of Section 4, for each p we fixed some point $y_p \in I_p$ for reference.

Before estimating the above it is convenient to give the following distortion estimates.

Lemma 11. $K'_{j,k} \leq 2C_6 K_{j,k}$.

Proof. Differentiating we obtain $K'_{j,k} = K_{j,k}\,(J'_m \circ h_j + J'_m \circ h_k)$. By Lemma 9 we know that $J'_n \leq C_6$. $\square$

Recall that $|b|^{-(1+\xi)} \leq |\Omega| \leq 2\,|b|^{-(1+\xi)}$ and $n_1 = \lceil \rho_1 \ln |b| \rceil$.

Lemma 12. *For all* $y \in \mathbb{T}^1$,

$$\sum_{x \in f^{-n_1}(y) \cap \Omega} J_{n_1}(x) \leq 6C_\lambda\,|\Omega|\,.$$

Proof. First note that

$$\sum_{x \in f^{-n_1}(y) \cap \Omega} J_{n_1}(x) = \sum_{x \in f^{-n_1}(y)} J_{n_1}(x) \cdot \mathbf{1}_\Omega = (\mathscr{L}_0^{n_1}\mathbf{1}_\Omega)(y),$$

and so it suffices to estimate $\|\mathscr{L}_0^{n_1}\mathbf{1}_\Omega\|_{\mathbf{L}^\infty(\mathbb{T}^1)}$. In one dimension $\|\cdot\|_{\mathbf{L}^\infty(\mathbb{T}^1)} \leq 2\,\|\cdot\|_{\mathbf{BV}}$. Moreover $\|\mathbf{1}_\Omega\|_{\mathbf{L}^1(\mathbb{T}^1)} = |\Omega|$ and $\|\mathbf{1}_\Omega\|_{\mathbf{BV}} = 2$. So, with the help of the estimate from Lemma 1,

$$\|\mathscr{L}_0^{n_1}\mathbf{1}_\Omega\|_{\mathbf{L}^\infty(\mathbb{T}^1)} \leq 2(2C_\lambda \lambda^{-n_1} + C_\lambda\,|\Omega|).$$

Note that $\lambda^{-n_1} = |b|^{-2}$ since $n_1 \geq \rho_1 \ln |b|$ and $\rho_1 = 2/\ln \lambda$. Consequently ($\xi \leq \frac{1}{2}$) the above quantity is bounded by $6C_\lambda\,|b|^{-(1+\xi)}$. $\square$

In a similar way to the above $\sum_{x \in f^{-n}(y)} J_n(x) = (\mathscr{L}_0^n \mathbf{1})(y)$ for all $y \in \mathbb{T}^1$, $n \in \mathbb{N}$. We may again apply the estimate from Lemma 1 and so

$$\sum_{x \in f^{-n}(y)} J_n(x) \leq 3C_\lambda. \tag{4.7}$$

By Lemma 11 we know that $K'_{j,k} \leq 2C_6 K_{j,k}$. Hence, by Gronwall's inequality,

$$K_{j,k}(z) \leq e^{2C_6|I_p|} K_{j,k}(x_p) \tag{4.8}$$

for all $z \in I_p$. Choosing b_0 large insures that $|I_p|$ is small and so

$$\int_{I_p} K_{j,k}(z)\,dz \leq 2\,|I_p|\,K_{j,k}(x_p).$$

Now we consider the sum in (4.6) corresponding to the noncancelling pairs (this is the second of the two terms on the right hand side). Using the above estimates

$$\sum_p \sum_j \sum_{k:(j,k)\in A_p^{\complement}} \int_{I_p} K_{j,k}(z)\,dz \leq 2 \sum_j \sum_{k:(j,k)\in A_p^{\complement}} K_{j,k}(x_p).$$

Note that

$$\sum_j J_n \circ h_j(x_p) \leq \sum_{y \in f^{-n}(x_p)\cap\Omega} J_n(y) \leq \sum_{z \in f^{-n_2}(x_p)} J_{n_2}(z) \sum_{y \in f^{-n_1}(z)\cap\Omega} J_{n_1}(y)$$

where $n_1 = \rho_1 \ln |b|$, $n_2 = \rho_2 \ln |b|$. Using also the estimate of Lemma 12

$$\sum_j J_n \circ h_j(x_p) \leq 6C_\lambda\,|\Omega| \sum_{z \in f^{-n_2}(x_p)} J_{n_2}(z). \tag{4.9}$$

Using the above estimates, together with (4.7) and (4.1), we obtain

$$\sum_p \sum_j \sum_{k:(j,k)\in A_p^{\complement}} \int_{I_p} K_{j,k}(z)\ dz \le (6C_\lambda)^2 |\Omega|^2 C_\lambda \sum_{\substack{z_2\in f^{-n_2}(x_p)\\ z_1 \not\sim z_2}} J_{n_2}(z_2) \le (6C_\lambda)^2 |\Omega|^2 C_\lambda C_\gamma e^{-n_2\gamma}. \tag{4.10}$$

Let us now consider the case where $(j,k) \in A_p$ and so estimate the remaining term of (4.6).

Lemma 13. *Suppose that $(j,k) \in A_p$. Then*

$$|\theta'_{j,k}(x_p)| > \tfrac{1}{2}C_1(J_{n_2}\circ g_j + J_{n_2}\circ g_k)(x_p).$$

Proof. Differentiating, since $\tau_n \circ h_j = \tau_{n_1}\circ h_j + \tau_{n_2}\circ f^{n_1}\circ h_j$, we obtain

$$\theta'_{j,k} = (\tau'_{n_1}\cdot J_n)\circ h_j - (\tau'_{n_1}\cdot J_n)\circ h_k + (\tau'_{n_2}\cdot J_{n_2})\circ g_j - (\tau'_{n_2}\cdot J_{n_2})\circ g_k.$$

Applying the estimate of Lemma 8 means that the first two terms can be estimated as

$$\left|(\tau'_{n_1}\cdot J_n)\circ h_j - (\tau'_{n_1}\cdot J_n)\circ h_k)\right| \le \tfrac{1}{2}C_1(J_{n_2}\circ g_j + J_{n_2}\circ g_k).$$

Using the estimate of Lemma 6 we have that

$$\left|(\tau'_{n_2}\cdot J_{n_2})\circ g_j - (\tau'_{n_2}\cdot J_{n_2})\circ g_k\right| > C_1(J_{n_2}\circ g_j + J_{n_2}\circ g_k). \qquad \square$$

The above lemma says that we have the required transversality at the point x_p. The following lemma says that the interval I_p has been chosen sufficiently small such that this same transversality holds for the entire interval I_p.

Lemma 14. *Suppose that $(j,k) \in A_p$. Then $|\theta'_{j,k}(y)| > \frac{1}{2}C_1\Lambda^{-n_2}$ for all $y \in I_p$.*

Proof. By Lemma 10 and Lemma 13 we know that $|\theta'_{j,k}(y)| > C_1\Lambda^{-n_2} - |I_p|\, C_7$. To complete the proof it remains to show that

$$|I_p| \le \frac{C_1}{2C_7}\Lambda^{-n_2}. \tag{4.11}$$

Recall that, by choice of the partition, $|I_p| \le 2\,|b|^{-(1-\xi)}$ and note that $\Lambda^{n_2} \le \Lambda\,|b|^{\rho_2 \ln\Lambda}$. This means that $|I_p| \le \Lambda^{-n_2} 2\Lambda\,|b|^{-(1-\xi-\rho_2\ln\Lambda)}$. Furthermore, by choice of ξ and ρ_2, we have $\rho_2 = \frac{\xi}{2\ln\Lambda}$ and $\xi \le \frac{1}{2}$, which means that $(1-\xi-\rho_2\ln\Lambda) \ge \frac{1}{4}$. Consequently $|I_p| \le \Lambda^{-n_2}2\Lambda\,|b|^{-\frac{1}{4}}$ and so, again increasing b_0 if required, $|I_p| \le \frac{C_1}{2C_7}\Lambda^{-n_2}$ for all $|b| \ge b_0$. $\square$

The key part of the argument is the following lemma concerning oscillatory integrals.

Lemma 15. *Suppose J is an interval, $\theta \in \mathscr{C}^2(J,\mathbb{R})$, $K \in \mathscr{C}^1(J,\mathbb{C})$, $b \in \mathbb{R}\setminus\{0\}$ and there exists $\kappa > 0$ such that* $\inf|\theta'| \ge \kappa$. *Then*

$$\left|\int_J K\cdot e^{ib\theta(x)}\ dx\right| \le \frac{1}{|b|}\left(\tfrac{2}{\kappa}\sup|K| + \tfrac{1}{\kappa^2}\sup|K|\sup|\theta''|\,|J| + \tfrac{1}{\kappa}\sup|K'|\,|J|\right).$$

Proof of Lemma 15. First change the variables, $y = \theta(x)$, then integrate by parts:

$$\begin{aligned}\int_J K \cdot e^{ib\theta(x)}\ dx &= \int_{\theta(J)} \frac{K}{\theta'} \circ \theta^{-1}(y) e^{iby}\ dy \\ &= -\frac{i}{b}\left[\frac{K}{\theta'} \circ \theta^{-1}(y) e^{iby}\right]_{\theta(J)} \\ &\quad + \frac{i}{b}\int_{\theta(J)} \left(\frac{K\theta''}{(\theta')^2 \cdot \theta'} + \frac{K'}{(\theta')^2}\right) \circ \theta^{-1}(y) e^{iby}\ dy.\end{aligned}$$

Changing variables again, we obtain

$$\int_J K \cdot e^{ib\theta(x)}\ dx = -\frac{i}{b}\left[\frac{K}{\theta'}e^{ib\theta}\right]_J + \frac{i}{b}\int_J \left(\frac{K\theta''}{(\theta')^2} + \frac{K'}{\theta'}\right)(x) e^{ib\theta(x)}\ dx.$$

The required estimate follows immediately. □

In preparation to applying the above lemma, note that (4.8) implies $\sup_{I_p} K_{j,k} \leq 2K_{j,k}(x_p)$ and similarly $\sup_{I_p} K'_{j,k} \leq 4C_6 K_{j,k}(x_p)$. By Lemma 14 we know that $|\theta'_{j,k}| > \frac{1}{2}C_1\Lambda^{-n_2}$. By Lemma 10 we know that $|\theta''_{jk}| \leq C_7$. We also know that $|I_p| \leq \frac{C_1}{2C_7}\Lambda^{-n_2}$ by (4.11). Using these estimates with Lemma 15 we obtain

$$\left|\int_{I_p} (K_{j,k} \cdot e^{ib\theta_{j,k}})(z)\ dz\right| \leq \frac{1}{|b|} K_{j,k}(x_p) C_8 \Lambda^{n_2}$$

where $C_8 := 8C_6(1 + C_7 + 2C_6)C_1^{-2}$. This means that for the first sum in (4.6) we obtain, using (4.7), Lemma 12 and decomposing $n = n_1 + n_2$ as we did before (4.10), the estimate

$$\sum_p \sum_{\substack{j,k \\ (j,k)\in A_p}} \left|\int_{I_p} (K_{j,k} \cdot e^{ib\theta_{j,k}})(z)\ dz\right| \leq \frac{1}{|b|} |b|^{1-\xi} C_8 \Lambda^{n_2} C_\lambda^2 (6C_\lambda)^2 |\Omega|^2$$

(the term $|b|^{1-\xi}$ comes from the sum over p). Since $|b| \geq e^{n_2 \frac{2\xi}{3\rho_1}}$ (increasing b_0 again if required),

$$\frac{1}{|b|} |b|^{1-\xi} \Lambda^{n_2} = \frac{1}{|b|^{\xi}} \Lambda^{n_2} \leq e^{-n_2(\frac{2\xi}{3\rho_2} - \ln\Lambda)}.$$

Let $\gamma_7 := \frac{\xi}{6\rho_2} > 0$. This means that (since $\rho_2 = \frac{\xi}{2\ln\Lambda}$ by (2.3))

$$\sum_p \sum_{\substack{j,k \\ (j,k)\in A_p}} \left|\int_{I_p} (K_{j,k} \cdot e^{ib\theta_{j,k}})(z)\ dz\right| \leq C_8 C_\lambda^2 (6C_\lambda)^2 |\Omega|^2 e^{-n_2\gamma_7}. \tag{4.12}$$

In order to estimate the final term in (4.5) we use (4.6) and sum the estimates (4.10) and (4.12) to obtain

$$\sum_p \sum_{j,k\in G_p} \left|\int_{I_p} (K_{j,k} \cdot e^{ib\theta_{j,k}})(z)\ dz\right| \leq |\Omega|^2 e^{-n\gamma_8} \left((6C_\lambda)^2 C_\lambda C_\gamma + C_8 C_\lambda^2 (6C_\lambda)^2\right),$$

where $\gamma_8 := \frac{\rho_1}{\rho_1+\rho_2}\min(\gamma, \gamma_7) > 0$. Let $C_3 := 6C_\lambda \left(C_\lambda C_\gamma + C_8 C_\lambda^2\right)^{\frac{1}{2}}$. Taking the square root of the above, we obtain

$$\sum_p \int_{I_p} \left|\sum_{j\in G_p} \left(J_n \cdot e^{ib\tau_m}\right) \circ h_j(z)\right|\ dz \leq |\Omega|\, e^{-n\frac{\gamma_8}{2}} C_3.$$

Including also the estimate (4.4) we have shown that (let $\gamma_3 := \min(\gamma_6, \frac{\gamma_8}{2}) > 0$)

$$\int_{\mathbb{T}^1} \left| \sum_j \left(J_n \cdot e^{ib\tau_n}\right) \circ h_j(z) \cdot \mathbf{1}_{f^n\omega_j}(z) \right| \, dz \leq (C_3 + C_5)\,|\Omega|\, e^{-n(b)\gamma_3}.$$

This completes the proof of Proposition 2.

5. Rate of mixing

Here we use the estimates of Proposition 1 concerning the twisted transfer operators in order to estimate the rate of mixing. Let $g, h \in \mathscr{C}^1(\mathbb{T}^2, \mathbb{C})$ be two observables. We assume, without loss of generality, that g is mean zero with respect to the invariant measure, i.e., $\int g(x,u) \cdot h_\nu(x)\, dxdu = 0$ (recall that h_ν denotes the density of the invariant measure). Denote by $\hat{g}_b$ and $\hat{h}_b$ their Fourier components (in the fibre coordinate), i.e.,

$$g(x,u) = \sum_{b\in\mathbb{Z}} \hat{g}_b(x) e^{ibu},$$

and similarly for $h(x,u)$. Using the regularity of the observables (in particular the smoothness in the fibre direction) we have[12] that $\|\hat{g}_b\|_{\mathbf{BV}} \leq \|g\|_{\mathscr{C}^1}$ and $\|\hat{g}_b\|_{\mathbf{L}^\infty} \leq |b|^{-1}\|g\|_{\mathscr{C}^1}$. The analogous results hold for h. That g is mean zero implies that $\int_{\mathbb{T}^1} \hat{g}_0(x) \cdot h_\nu(x)\, dx = 0$. By substituting the Fourier series and simple manipulations we obtain the formula

$$\int_{\mathbb{T}^2} (g \cdot h \circ F^n)(x,u) \cdot h_\nu(x)\, dx\, du = \sum_{b\in\mathbb{Z}} \int_{\mathbb{T}^1} \mathscr{L}_b^n(\hat{g}_b \cdot h_\nu)(x) \cdot \hat{h}_{-b}(x)\, dx.$$

We separate the sum into three pieces: the term where $b = 0$, a finite number of terms where $b \neq 0$ and $|b| \leq b_0$, and the infinite sum of the remaining terms. When $b = 0$ we take advantage of the exponential mixing of the base map f (a standard consequence of the covering property combined with Lemma 1) and so

$$\left| \int_{\mathbb{T}^1} \mathscr{L}_0^n(\hat{g}_b \cdot h_\nu)(x) \cdot \hat{h}_{-b}(x)\, dx \right| \leq C_\lambda \lambda^{-n} \|\hat{g}_0\|_{\mathbf{BV}} \|\hat{h}_0\|_{\mathbf{L}^\infty},$$

using also that $\hat{g}_0$ has zero mean. For the terms where $b \neq 0$, $|b| \leq b_0$ we use the following lemma. The proof follows exactly the proof of [2, Lemma 7.21].

Lemma 16. *Suppose that τ is not Lipschitz-cohomologous to a piecewise constant*[13] *and that $b \neq 0$. Then the spectral radius of $\mathscr{L}_b : \mathbf{BV} \to \mathbf{BV}$ is strictly less than* 1.

Proof. We already know (Lemma 1) that $\mathscr{L}_b : \mathbf{BV} \to \mathbf{BV}$ is quasi compact. Suppose that there exists $\sigma \in \mathbb{C}$, $|\sigma| = 1$, and nonzero $h \in \mathbf{BV}$ such that

$$\mathscr{L}_b h = \sigma h.$$

Note that $|h| = |\mathscr{L}_b h| \leq \mathscr{L}_0 |h|$. Combined with the fact that $\int |h|(x)\, dx = \int \mathscr{L}_0 |h|(x)\, dx$ this means that $|h| = \mathscr{L}_0 |h|$. Consequently $\mathscr{L}_0 |h| = |\mathscr{L}_b h|$ and

$$\mathscr{L}_0^n |h| = |\mathscr{L}_b^n h|$$

for any $n \in \mathbb{N}$. This means that, for each y, the quantity $\arg(e^{ib\tau_n(x)} \cdot h(x))$ takes the same value for all $x \in f^{-n}(y)$. Choose $k \in \mathbb{N}$ such that $kb > b_0$ and consider

[12] We use the definition $\|g\|_{\mathscr{C}^1} := \sup|g| + \sup|g'|$.

[13] I.e., assume that the first alternative of Proposition 3 holds.

$h^k \in \mathbf{BV}$. This means that $\arg(e^{ibk\tau_n(x)} \cdot h^k(x))$ takes the same value for all $x \in f^{-n}(y)$. Consequently $\mathscr{L}_0^n \left|h^k\right| = \left|\mathscr{L}_{kb}^n h^k\right|$ and hence

$$\left\|\mathscr{L}_0^n \left|h^k\right|\right\|_{\mathbf{L^1}} = \left\|\mathscr{L}_{kb}^n h^k\right\|_{\mathbf{L^1}}.$$

This is a contradiction since $\left\|\mathscr{L}_0^n \left|h^k\right|\right\|_{\mathbf{L^1}} = \left\|h^k\right\|_{\mathbf{L^1}} > 0$, but, by Proposition 1, $\left\|\mathscr{L}_{kb}^n h^k\right\|_{\mathbf{L^1}} \to 0$ as $n \to \infty$. □

Now we deal with the terms where $|b| \geq b_0$. Recall that in Proposition 1 we obtained the estimate $\|\mathscr{L}_b^{n(b)}\|_{(b)} \leq e^{-n(b)\gamma_2}$ where $\rho \ln|b| \leq n(b) \leq \rho \ln|b| + 2$. We may assume that $\gamma_2 > 0$ is sufficiently small such that $\rho\gamma_2 < 1$. Consequently the above estimate implies that there exists $\alpha \in (0,1)$ such that $\|\mathscr{L}_b^n\|_{(b)} \leq |b|^\alpha e^{-n\gamma_2}$ for all $n \in \mathbb{N}$, $|b| \geq b_0$ (the above estimate holds only for multiples of $n(b)$; for intermediate values of n we use again Lemma 2 and increase b_0 if required). Note that

$$\begin{aligned}\left|\int_{\mathbb{T}^1} \mathscr{L}_b^n(\hat{g}_b \cdot h_\nu)(x) \cdot \hat{h}_{-b}(x)\, dx\right| &\leq \|\mathscr{L}_b^n(\hat{g}_b \cdot h_\nu)\|_{\mathbf{L^1}} \|\hat{h}_{-b}\|_{\mathbf{L^\infty}} \\ &\leq \|\mathscr{L}_b^n\|_{(b)} \|\hat{g}_b \cdot h_\nu\|_{(b)} \|\hat{h}_{-b}\|_{\mathbf{L^\infty}} \\ &\leq \|\mathscr{L}_b^n\|_{(b)} \left((1+|b|)^{-1} \|\hat{g}_b\|_{\mathbf{BV}} + \|\hat{g}_b\|_{\mathbf{L^1}}\right) \|h_\nu\|_{\mathbf{BV}} \|\hat{h}_{-b}\|_{\mathbf{L^\infty}} \\ &\leq \|\mathscr{L}_b^n\|_{(b)} |b|^{-2} \|h_\nu\|_{\mathbf{BV}} \|g\|_{\mathscr{C}^1} \|h\|_{\mathscr{C}^1}.\end{aligned}$$

It remains to observe that

$$\sum_{|b|\geq b_0} \|\mathscr{L}_b^n\|_{(b)} |b|^{-2} \leq \sum_{|b|\geq b_0} |b|^{-(2-\alpha)} e^{-n\gamma_2}.$$

Crucially $(2-\alpha) > 1$ and so this is summable. This proves exponential mixing for $\mathscr{C}^1$ observables, which, by an approximation argument [10, Proof of Corollary 1], implies exponential mixing for Hölder observables.

Acknowledgements

The authors are grateful for the hospitality of Carlangelo Liverani and for ERC Advanced Grant MALADY (246953). They are also grateful for advice and support from all at the Budapest and Vienna dynamical systems groups. It is a pleasure to thank Viviane Baladi for a careful reading of a preliminary version and helpful comments. The authors are grateful to the anonymous referee for valuable comments.

References

[1] José F. Alves, Maria Carvalho, and Jorge Milhazes Freitas, *Statistical stability for Hénon maps of the Benedicks-Carleson type*, Ann. Inst. H. Poincaré Anal. Non Linéaire **27** (2010), no. 2, 595–637, DOI 10.1016/j.anihpc.2009.09.009. MR2595193 (2011g:37052)

[2] Artur Avila, Sébastien Gouëzel, and Jean-Christophe Yoccoz, *Exponential mixing for the Teichmüller flow*, Publ. Math. Inst. Hautes Études Sci. **104** (2006), 143–211, DOI 10.1007/s10240-006-0001-5. MR2264836

[3] Viviane Baladi, Michael Benedicks, and Véronique Maume-Deschamps, *Almost sure rates of mixing for i.i.d. unimodal maps* (English, with English and French summaries), Ann. Sci. École Norm. Sup. (4) **35** (2002), no. 1, 77–126, DOI 10.1016/S0012-9593(01)01083-7. MR1886006 (2003d:37027)

[4] Viviane Baladi and Carlangelo Liverani, *Exponential decay of correlations for piecewise cone hyperbolic contact flows*, Comm. Math. Phys. **314** (2012), no. 3, 689–773, DOI 10.1007/s00220-012-1538-4. MR2964773

[5] Viviane Baladi and Brigitte Vallée, *Exponential decay of correlations for surface semi-flows without finite Markov partitions*, Proc. Amer. Math. Soc. **133** (2005), no. 3, 865–874 (electronic), DOI 10.1090/S0002-9939-04-07671-3. MR2113938 (2006d:37047)

[6] Oliver Butterley, *Area expanding $\mathcal{C}^{1+\alpha}$ suspension semiflows*, Comm. Math. Phys. **325** (2014), no. 2, 803–820, DOI 10.1007/s00220-013-1835-6. MR3148102

[7] Nikolai Chernov and Roberto Markarian, *Chaotic billiards*, Mathematical Surveys and Monographs, vol. 127, American Mathematical Society, Providence, RI, 2006. MR2229799 (2007f:37050)

[8] N. I. Chernov, *Markov approximations and decay of correlations for Anosov flows*, Ann. of Math. (2) **147** (1998), no. 2, 269–324, DOI 10.2307/121010. MR1626741 (99d:58101)

[9] Mark F. Demers and Carlangelo Liverani, *Stability of statistical properties in two-dimensional piecewise hyperbolic maps*, Trans. Amer. Math. Soc. **360** (2008), no. 9, 4777–4814, DOI 10.1090/S0002-9947-08-04464-4. MR2403704 (2009f:37021)

[10] Dmitry Dolgopyat, *On decay of correlations in Anosov flows*, Ann. of Math. (2) **147** (1998), no. 2, 357–390, DOI 10.2307/121012. MR1626749 (99g:58073)

[11] Dmitry Dolgopyat and Carlangelo Liverani, *Energy transfer in a fast-slow Hamiltonian system*, Comm. Math. Phys. **308** (2011), no. 1, 201–225, DOI 10.1007/s00220-011-1317-7. MR2842975 (2012h:37072)

[12] P. Eslami, *Stretched-exponential mixing for $\mathscr{C}^{1+\alpha}$ skew products with discontinuities*, preprint, arXiv:1405.6981, 2014.

[13] Lawrence C. Evans and Ronald F. Gariepy, *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1992. MR1158660 (93f:28001)

[14] Jorge Milhazes Freitas, *Continuity of SRB measure and entropy for Benedicks-Carleson quadratic maps*, Nonlinearity **18** (2005), no. 2, 831–854, DOI 10.1088/0951-7715/18/2/019. MR2122687 (2005m:37082)

[15] S. Galatolo and I. Nisoli, *Rigorous computation of invariant measures and fractal dimension for piecewise hyperbolic maps: 2D Lorenz like maps*, To appear in Ergod. Th. & Dynam. Sys., 2014.

[16] Stefano Galatolo, Jérôme Rousseau, and Benoit Saussol, *Skew products, quantitative recurrence, shrinking targets and decay of correlations*, Ergodic Theory Dynam. Systems **35** (2015), no. 6, 1814–1845, DOI 10.1017/etds.2014.10. MR3377286

[17] Sébastien Gouëzel, *Local limit theorem for nonuniformly partially hyperbolic skew-products and Farey sequences*, Duke Math. J. **147** (2009), no. 2, 193–284, DOI 10.1215/00127094-2009-011. MR2495076 (2010i:37012)

[18] Gerhard Keller, *Stochastic stability in some chaotic dynamical systems*, Monatsh. Math. **94** (1982), no. 4, 313–333, DOI 10.1007/BF01667385. MR685377 (84k:58130)

[19] G. Keller, *Markov extensions, zeta functions, and Fredholm theory for piecewise invertible dynamical systems*, Trans. Amer. Math. Soc. **314** (1989), no. 2, 433–497, DOI 10.2307/2001395. MR1005524 (91c:58062)

[20] Gerhard Keller and Carlangelo Liverani, *Stability of the spectrum for transfer operators*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **28** (1999), no. 1, 141–152. MR1679080 (2000b:47030)

[21] G. Keller and C. Liverani, *A spectral gap for a one-dimensional lattice of coupled piecewise expanding interval maps*, Dynamics of coupled map lattices and of related spatially extended systems, Lecture Notes in Phys., vol. 671, Springer, Berlin, 2005, pp. 115–151, DOI 10.1007/11360810_6. MR2395729 (2009f:37022)

[22] Gerhard Keller and Carlangelo Liverani, *Map lattices coupled by collisions*, Comm. Math. Phys. **291** (2009), no. 2, 591–597, DOI 10.1007/s00220-009-0835-z. MR2530174 (2010g:37039)

[23] Carlangelo Liverani, *Decay of correlations*, Ann. of Math. (2) **142** (1995), no. 2, 239–301, DOI 10.2307/2118636. MR1343323 (96e:58090)

[24] Carlangelo Liverani, *Decay of correlations for piecewise expanding maps*, J. Statist. Phys. **78** (1995), no. 3-4, 1111–1129, DOI 10.1007/BF02183704. MR1315241 (96d:58077)

[25] Carlangelo Liverani, *Rigorous numerical investigation of the statistical properties of piecewise expanding maps. A feasibility study*, Nonlinearity **14** (2001), no. 3, 463–490, DOI 10.1088/0951-7715/14/3/303. MR1830903 (2002b:37128)

[26] Carlangelo Liverani, *On contact Anosov flows*, Ann. of Math. (2) **159** (2004), no. 3, 1275–1312, DOI 10.4007/annals.2004.159.1275. MR2113022 (2005k:37048)

[27] Ippei Obayashi, *Exponential decay of correlations for surface semiflows with an expanding direction*, J. Math. Kyoto Univ. **49** (2009), no. 2, 427–440. MR2571851 (2011b:37004)
[28] Mark Pollicott, *On the rate of mixing of Axiom A flows*, Invent. Math. **81** (1985), no. 3, 413–426, DOI 10.1007/BF01388579. MR807065 (87i:58148)
[29] David Ruelle, *Flots qui ne mélangent pas exponentiellement* (French, with English summary), C. R. Acad. Sci. Paris Sér. I Math. **296** (1983), no. 4, 191–193, DOI 10.1142/9789812833709_0024. MR692974 (84h:58085)
[30] Omri Sarig, *Subexponential decay of correlations*, Invent. Math. **150** (2002), no. 3, 629–653, DOI 10.1007/s00222-002-0248-5. MR1946554 (2004e:37010)
[31] Masato Tsujii, *Decay of correlations in suspension semi-flows of angle-multiplying maps*, Ergodic Theory Dynam. Systems **28** (2008), no. 1, 291–317, DOI 10.1017/S0143385707000430. MR2380311 (2009b:37051)
[32] Lai-Sang Young, *Statistical properties of dynamical systems with some hyperbolicity*, Ann. of Math. (2) **147** (1998), no. 3, 585–650, DOI 10.2307/120960. MR1637655 (99h:58140)
[33] Lai-Sang Young, *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999), 153–188, DOI 10.1007/BF02808180. MR1750438 (2001j:37062)

Fakultät für Mathematik, Universität Wien, Oskar-Morgenstern-Platz 1, 1090 Wien, Austria

E-mail address: `oliver.butterley@univie.ac.at`

Dipartimento di Matematica, II Università di Roma (Tor Vergata), Via della Ricerca Scientifica, 00133 Roma, Italy

E-mail address: `eslami@mat.uniroma2.it`