A SMOOTH MIXING FLOW ON A SURFACE WITH NON-DEGENERATE FIXED POINTS

JON CHAIKA AND ALEX WRIGHT

ABSTRACT. We construct a smooth, area preserving, mixing flow with finitely many non-degenerate fixed points and no saddle connections on a closed surface of genus 5. This resolves a problem that has been open for four decades.

1. INTRODUCTION

Motivation and main result. Flows on surfaces are a basic example in smooth dynamics, being in a sense the smallest smooth dynamical systems after circle diffeomorphisms, and have been the topic of a vast body of research.

In 1972, the existence of smooth ergodic flows on all closed surfaces except the sphere, projective plane, and Klein bottle was established by Blohin [**Blo72**]. Only a few years later Katok, Sinai, and Stepin indicated the following as an open problem in their 1975 survey paper [**KSS75**, 4.4.1].

"Let T_t be a smooth flow on a surface of genus $p \geq 2$ with smooth positive invariant measure, all of whose fixed points are non-degenerate saddles. Can T_t be mixing? The distinguished results of A. V. Kochergin and A. B. Katok give a negative answer to this question in all probability."

The same question was listed by Katok and Thouvenot in the Handbook of Dynamical Systems [**KT06**, Problem 6.10] and was mentioned by Forni in [**For02**, Page 4]. The purpose of this paper is to provide a positive answer.

Theorem 1.1. *There is a mixing, smooth, area preserving flow on a surface of genus 5 with finitely many fixed points, all non-degenerate, and no saddle connections.*

A saddle connection is a flow trajectory joining two fixed points for the flow. The derivative of a smooth flow is a vector field, which can be written locally on a surface as $A(x,y)\partial_x + B(x,y)\partial_y$. A fixed point is called non-degenerate if at that point the function (A,B) has non-zero Jacobian, i.e. if $A_xB_y - A_yB_x \neq 0$.

Already at the time of Katok, Sinai, and Stepin's question, Kochergin had shown that mixing flows do exist on surfaces if degenerate saddles are allowed [**Koč75**], again on all closed surfaces except the sphere, the projective plane and the Klein bottle. However, the presence of degenerate saddles has such a drastic effect that it is reasonable to believe that the natural class of dynamical systems that should be grouped together is not all smooth area preserving flows on surfaces, but rather those with only finitely many saddles, all of which are non-degenerate.

Kochergin also showed that a class of smooth flows on surfaces with finitely many non-degenerate fixed points arising from Blohin's construction (namely certain suspensions of irrational circle rotations) are never mixing [**Koč76**], [**Koc07a**]. Thus Kochergin's results supports making a large distinction between degenerate and non-degenerate saddles. A consequence of Kochergin's result is that there does not exist a mixing, smooth, area preserving flow on a surface of genus 1 with finitely many fixed points, all non-degenerate, and no saddle connections [**Koc07a**].

The intuition that the types of flows considered in our main theorem are very unlikely to be mixing has proven correct, as Ulcigrai has recently established that such flows are typically not mixing [**Ulc11**]; using different methods Scheglov had established this in genus 2 [**Sch09a**]. It is however known that these flows are generically uniquely ergodic [**Mas82, Vee82**] and weak mixing [**Ulc09**]. (The notion of generic here is measure theoretic.) Many examples are mild mixing [**KKP**].

Kochergin's mechanism for mixing. Consider a small horizontal line segment in $\mathbb{R}^2$. Under the action of $\begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}$ this small horizontal line segment will be sheared until eventually it is close to a long vertical line segment. Similarly, an interval transverse to a flow may eventually get sheared so much that it becomes close to an orbit of the flow. Kochergin's seminal observation is that in this case, if the flow is ergodic, one may expect equidistribution of such flowed transverse intervals, and subsequently hope to conclude that the flow is mixing.

This idea has been used in many subsequent works, and has also been applied to flows on higher dimensional manifolds, such as Fayad's example of a reparameterization of a linear flow on $\mathbb{T}^3$ that is mixing and has singular spectrum [**Fay06**]. (We mention in passing that mixing is easier to obtain in dimension greater two, and indeed Fayad obtains mixing for a flow without fixed points, because there are two dimensions transverse to the flow which may alternately be sheared.)

Kochergin's technique will be the engine of mixing in Theorem 1.1. The shearing effects are most significant near the fixed points of the flow, since trajectories that pass closer to a fixed point will get slowed down more than trajectories that pass farther away. This is why degenerate fixed points can help in establishing mixing; they establish a strong shearing effect. For non-degenerate fixed points, the shearing effect is weaker, and typically many passes near a fixed point are required to accumulate an appreciable amount of shearing. However, passing on different sides of a fixed point produces opposite effects, which are expected to cancel out. This leads to the intuition, which has been made rigorous with the work of Scheglov, Ulcigrai, and others, that area preserving flows should not be mixing.

Suspension flows. Given a flow on a surface with finitely many fixed points and no saddle connections, pick a disjoint union I of intervals transverse to the flow direction. (Typically one picks only one interval, but it will be convenient for us to use four.) Let $T : I \to I$ be the first return map, and let $f : I \to (0, \infty]$ be the first return time function. The flow is isomorphic to the vertical flow F^t on the space

$$Z = \{(x, s) : x \in I, 0 \leq s \leq f(x)\}/((x, f(x)) \sim (T(x), 0)),$$

defined by $F^t(x, s) = (x, s + t)$ for all $0 \leq s + t \leq f(x)$. Note in particular that $F^{f(x)}(x, 0) = (Tx, 0)$.

If the flow is measure preserving, I can be parameterized so that T is a *multi-interval exchange transformation*, i.e., a permutation of a finite number of subintervals that partition I. (If there is only one interval, this is called an *interval exchange transformation.*) Since the flow is smooth, f is smooth away from the discontinuities of T.

The standard model for a non-degenerate fixed point is given by the vector field $x\partial_y + y\partial_x$. This has two incoming trajectories, and two outgoing trajectories. For a flow with only finitely many non-degenerate fixed points, the first return map is, up to a bounded function with bounded derivative, equal to a function of the form $f = 1 - \sum c_i \log |x - x_i|$. Since the roof function is infinite at the x_i, these are the points that orbit into a fixed point. If x_i and x_j orbit into the same discontinuity, then $c_i = c_j$. More precise statements can be found in [**Kat73**] and [**Koč76**, Section 3]; see also [**CF11**].

Moreover, standard arguments show that all T and f satisfying certain technical conditions arise from smooth flows on surfaces with only finitely many non-degenerate fixed points [**CF11**, Section 7]. Thus, to prove our main theorem, we will find an appropriate T and f for which the suspension flow is mixing, and in the last section of this paper will will explain how Theorem 1.1 follows.

Birkhoff sums of non-integrable functions. We will see that the net shearing of an interval transverse to the flow is controlled by Birkhoff sums of f', that is, by sums of the form

$$\sum_{i=0}^{N-1} f'(T^i x).$$

Our roof function f will have f' non-integrable, so the Birkhoff Ergodic Theorem may not be used to understand these sums. Note also that f will not be of bounded variation. Katok has shown that suspension flows over interval exchange transformations with roof functions of bounded variation are never mixing [**Kat80**].

To get enough shearing for mixing, we will require the above Birkhoff sums to grow faster than linearly in N, and we will need fairly precise control.

The problem is that we expect a large amount of cancellation to occur between positive and negative terms in this Birkhoff sum. When $T^i(x)$ is close to and on the right side of a singularity, $f'(T^i x)$ will be very negative. When $T^i(x)$ is close to and on the left side of a singularity, $f'(T^i x)$ will be very positive.

The Katok-Sataev-Veech construction. In turns out that the following result is technically easier to prove than Theorem 1.1.

Theorem 1.2. *There a $\mathbb{Z}_2 = \{0, 1\}$ skew product $T : S^1 \times \mathbb{Z}_2 \to S^1 \times \mathbb{Z}_2$ over a rotation with the following properties: T has four discontinuities, $p_1, p_2, p_3, p_4 \in S^1 \times \mathbb{Z}_2$. There is a fifth point $p_0 \in S^1 \times \mathbb{Z}_2$ that is not a discontinuity, such that the suspension flow over T with roof function*

$$f(p) = 1 - \sum_{i=0}^{4} \log d(p, p_i)$$

is mixing, where d is a distance function on $S^1 \times \mathbb{Z}_2$ that restricts to the standard distance on each of the two copies of S^1.

This choice of T and f do not satisfy the technical conditions to correspond to a smooth flow, because the roof function has one extra non-degenerate singularity that is not at a discontinuity of T. (For a smooth flow with no saddle connections, and all fixed points non-degenerate, each singularity of the roof function must occur at a discontinuity of the multi-interval exchange transformation.) In the final section we choose a closely related but more complicated $\hat{T}$ and $\hat{f}$ to prove Theorem 1.1.

Most of this paper is occupied with the proof of Theorem 1.2. To build T, we modify the Katok-Sataev-Veech construction for producing examples of minimal but non-uniquely ergodic interval exchange transformations [**Vee69, Kat73, Sat75**]. Our T will in fact be uniquely ergodic, but orbits equidistribute very slowly and in a controlled manner.

To obtain mixing, we construct T to be very well approximated by non-minimal $\mathbb{Z}_2$ skew products of rotations, T_k, such that T_k has two minimal components, one of which contains an interval to the left of x_0, and one of which contains an interval to the right of x_0. Quantitative estimates, and highly non-generic choices of parameters such as the continued fraction expansion of the base rotation, allow us to show that this asymmetry in the minimal approximants yields appropriate growth in the Birkhoff sums of f' and hence obtains sufficient shearing for mixing. To get this growth, we must prevent the terms in these sums where $T^i(x)$ is to the right of x_0 from canceling with the terms where $T^i(x)$ is to the left of x_0. This is difficult because T is uniquely ergodic, so all orbits equidistribute. We show that the terms where $T^i(x)$ is in certain decreasing neighborhoods of x_0 dominate these sums, and within these smaller and smaller neighborhoods orbits of certain lengths are not at all equidistributed.

Technically speaking, our analysis makes heavy use of continued fractions and the Denjoy-Koksma inequality. We chose parameters so that the base rotation has orbits which equidistribute very quickly, but the two minimal components in the skew product equidistribute slowly.

Open problems. The proof of Theorem 1.1 is built on the fact that there are non-minimal smooth flows with finitely many non-degenerate fixed points on surfaces of genus 5 such that one minimal component sees only one side of a fixed point. As we observe in Remark 9.9, such flows also exist on surfaces of genus 3 and 4. However, there are no such flows on surfaces of genus 2. Thus, the following seems especially interesting.

Problem 1.3. Is there a mixing, smooth, area preserving flow on a surface of genus 2 with 2 non-degenerate fixed points and no saddle connections?

Some pointers to the literature. We have tabulated over 40 papers that should be mentioned in any complete history of area preserving flows on surfaces. We will omit many of them here. See [**Koc07b**] for a survey.

Some early examples of flows on surfaces that are not mixing include [**Kol53, Kat67, Sch09b, Koč72, Koč76**]. The first result on weak mixing of flows may perhaps be due to von Neuman [**vN32, FL09**].

Novikov has suggested a link between area preserving flows on surfaces, and solid state physics [**Nov82**]. Because of this connection to physics, area preserving flows on surfaces are often called multi-valued Hamiltonian flows.

Arnold pointed out that flows over interval exchange transformations with asymmetric logarithmic singularities arise from non-minimal flows on surfaces, and as a result these flows are now well studied and it is known that mixing in this context is generic [**Ulc07**], see also [**Arn91, Koc03, Koc04c, Koc04a, Koc04d, SK92, Rav17**].

Fraczek-Lemanczyk have provided many examples of smooth area preserving flows on surfaces that are disjoint from all mixing flows, and have proven weak mixing for many suspension flows over rotations [**FL05, FL03**].

Kochergin has given examples of flows over rotations that mix at polynomial speed on rectangles [**Koc04b**]. Fayad-Kanigowski have shown multiple mixing of many suspension flows over rotations with asymmetric logarithmic singularities or degenerate fixed points [**FK**].

Suspension flows over interval exchange transformations with different roof functions are also frequently studied, see for example [**Lem00, Koc02, FL04, FL06**].

While this paper was under revision, Fayad, Forni and Kanigowski showed that for area preserving flows on the torus with one sufficiently strong singularity, the maximal spectral type is typically Lebesgue measure on the line [**FFK**]; Ravotti proved quantitative mixing estimates for minimal components of generic smooth flows on surfaces (suspension flows with asymmetric logarithmic singularites) [**Rav17**]; and in the same same context Kanigowski, Kulaga-Przymus and Ulcigrai proved mixing of all orders [**KKPU**].

The organization of this paper. Most of this paper is occupied with the proof of Theorem 1.2. In the final Section 9 we explain how to modify our construction to yield Theorem 1.1.

Section 2 collects standard results on continued fractions and rotations whose use will be ubiquitous in our analysis. In Section 3 we list the assumptions we must place on the rotation number of the base rotation, and we give an explicit example of a continued fraction expansion satisfying these assumptions. In Section 4 we define the skew product T used to prove Theorem 1.2, and discuss its non-minimal approximants T_k. In Sections 5 and 6 we prove estimates on Birkhoff sums. In Section 7 we prove unique ergodicity of the T used in Theorem 1.2, and in Section 8 we prove Theorem 1.2 from the results in the pervious sections via a standard argument.

Big O and little o notation. Given two sequences of numbers, $\{c_n\}_{n\in\mathbb{N}}$ and $\{d_n\}_{n\in\mathbb{N}}$, we write $c_n = O(d_n)$ if there exists $M \in \mathbb{R}$ so that $-M|d_n| < c_n < M|d_n|$ for all n. We write $c_n = o(d_n)$ if $\lim\limits_{n\to\infty} \frac{c_n}{d_n} = 0$. In Remark 5.3 we specify additional conventions for quantities that depend on more than on parameter.

Warning. Readers should pay careful attention to the typesetting in subscripts, for example to distinguish $q_{n_{k+1}}$ and q_{n_k+1}.

Acknowledgements. Research of JC partially supported by the NSF grants DMS 1004372, 1300550 and a Warnock Chair. Research of AW partially supported by a Clay Research Fellowship. The authors are very grateful to Anatole Katok, Corinna Ulcigrai, and Amie Wilkinson for useful conversations, and to the referees for their careful reading of the paper and helpful suggestions.

2. Continued fractions and rotations

Continued fractions. Fix a positive irrational real number $\alpha \in \mathbb{R}$. Let a_n denote the n-th term in the continued fraction expansion of α, and let p_n/q_n denote the n-th best approximant of α.

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}} \qquad \frac{p_n}{q_n} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_n}}}}$$

A best approximant of α is defined to be any rational number p/q such that if p', q' are integers with $0 < q' \leq q$, then

$$\left|\alpha - \frac{p}{q}\right| \leq \left|\alpha - \frac{p'}{q'}\right|.$$

Theorem 2.1. *The following hold.*

(1) Recursive formulas:

$$\begin{aligned} p_{n+1} &= a_{n+1}p_n + p_{n-1}, \\ q_{n+1} &= a_{n+1}q_n + q_{n-1}, \\ \frac{p_n}{q_n} &= \frac{p_{n-1}}{q_{n-1}} + \frac{(-1)^{n+1}}{q_{n-1}q_n}. \end{aligned}$$

(2) *Alternating property:*

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \cdots < \alpha < \cdots < \frac{p_3}{q_3} < \frac{p_1}{q_1}.$$

(3) *Upper and lower bounds:*

$$\frac{1}{q_k(q_k+q_{k+1})} < |\alpha - \frac{p_k}{q_k}| < \frac{1}{q_k q_{k+1}}.$$

(4) *Best approximants property: The set of best approximants of α is exactly equal to $\{\frac{p_n}{q_n}\}_{n=1}^{\infty}$.*

For proofs of any of these facts, see [**Khi64**]. In particular, see page 36 for the upper and lower bounds.

Rotations. Let $S^1 = \mathbb{R}/\mathbb{Z}$ denote the circle, and let $R : S^1 \to S^1$ denote rotation by α, so $R(x) = x + \alpha$. We will often implicitly identify the circle with the interval $[0, 1)$.

Let d denote the distance on the circle coming from the standard distance on $\mathbb{R}$.

Given $x \in S^1$, define a closest return time as positive integer q such that if $0 < q' < q$, then

$$d(x, R^{q'}(x)) > d(x, R^q(x)).$$

This definition does not depend on x. Define the orbit segment of length q of $x \in S^1$ to be the sequence $\{R^i(x)\}_{i=0}^{q-1}$.

We will say that a subset of S^1 is δ-*separated* if the distance between any two distinct points in the subset is at least δ.

Theorem 2.2. *The following hold.*

(1) *Alternating property: For n odd $R^{q_n}(x) \in x + (-\frac{1}{2}, 0)$, and for n even $R^{q_n}(x) \in x + (0, \frac{1}{2})$.*
(2) *Upper and lower bounds:*

$$\frac{1}{q_n + q_{n+1}} < d(x, R^{q_n}x) < \frac{1}{q_{n+1}}.$$

(3) *Best approximants property: The closest return times are exactly $\{q_n\}_{n=1}^{\infty}$.*
(4) *Separation: Any orbit segment of length at most q_n is at least $d(x, R^{q_{n-1}}x)$ separated.*
(5) *Equidistribution: $\{R^i 0\}_{i=0}^{q_n-1}$ contains exactly one point of each interval $[\frac{i}{q_n}, \frac{i+1}{q_n})$, for $i = 0, \ldots, q_n - 1$.*

(6) *Denjoy-Koksma Inequality: For any function* $g : [0,1) \to \mathbb{R}$ *of bounded variation, and any* $x \in S^1$,

$$\left| \sum_{i=0}^{q_n-1} g(R^i(x)) - q_n \int g \right| \leq \mathrm{Var}(g),$$

where $\mathrm{Var}(g)$ *is the total variation of* g.

The first three statements follow from the corresponding statements in Theorem 2.1. Separation follows from the best approximants property. The equidistribution property follows in an elementary way from the bounds in Theorem 2.1, and the Denjoy-Koksma inequality follows from the equidistribution property. For proofs of the equidistribution property and Denjoy-Koksma, we highly recommend the blog post of Lima [**Lim**] (see Lemma 5 and Theorem 6), which is partially based on the paper by Herman [**Her79**] where Denjoy-Koksma was first proven.

3. Picking α

In this paper, we will require α with very special properties. Precisely, we will require the existence of a subsequence n_k of the positive integers such that the following assumptions on n_k and the continued fraction expansion of α hold. Let $\langle\langle y \rangle\rangle$ denote the fractional part of a real number y, so $\langle\langle y \rangle\rangle \in [0,1)$. Note that by the Alternating Property, $d(x, R^{q_i}(x)) = \langle\langle q_i\alpha \rangle\rangle$ if i is even, and $d(x, R^{q_i}(x)) = 1 - \langle\langle q_i\alpha \rangle\rangle$ if i is odd.

Assumptions.

(1) All n_k are even.
(2) $\sum_{k=1}^{\infty} 2\langle\langle q_{n_k}\alpha \rangle\rangle < \min(\frac{\alpha}{4}, 1 - \frac{\alpha}{4})$.
(3) For all $\ell \geq 1$, we have $1+\sum_{k=1}^{\ell} 2q_{n_k} < q_{n_\ell+1}$ and $1+\sum_{k=1}^{\ell} q_{n_k-1} < q_{n_\ell}$.
(4) $\lim\limits_{\ell\to\infty} a_{n_\ell+1} = \infty$.
(5) $a_n = o(\log\log(q_n))$, and $a_{n+1} = o(\log\log(q_n))$.
(6) $\sum_{k=1}^{\infty} a_{n_k+1}^{-1} = \infty$.
(7) $\sum_{i=1}^{n+1} a_i = O(\log(q_n))$.
(8) $\log(q_{n_{k-1}+1}) = o(\log(q_{n_k}))$.
(9) For all k, $a_{n_k} = 2$ and $a_{n_k-1} = 2$.

Note that Assumptions 5 and 7 are on all n. Our assumptions are in effect for the remainder of this paper. They are stronger than necessary; we saw no benefit in trying to pick the weakest sufficient assumptions.

Remark 3.1. Assumptions 1 and 2 allow us to fix notation. They have no particular dynamical significance for us. Assumption 3 is used in proving unique ergodicity in Section 7, and a few other places, where it guarantees that the sequence q_{n_k} grows very rapidly.

Assumption 4 indicates that at time q_{n_ℓ}, orbits come back very close to themselves, and for a long time after that they almost repeat their paths. This will govern transfer of mass between invariant subsets of certain non-minimal approximants of T. Assumption 6 guarantees that T is uniquely ergodic.

Assumptions 5 and 7 reflect that the continued fraction has mostly small coefficients, and hence α is poorly approximated by rationals, so the rotation by α has especially good equidistribution properties. These assumptions are important in estimating Birkhoff sums.

Assumption 8 indicates that the times n_k are chosen so sparsely that orbits of length q_{n_k} come back vastly closer to themselves than orbits of length $q_{n_{k-1}}$. It will be used in Lemma 5.6.

Assumption 9 is used in the proofs of unique ergodicity of T and $\hat{T}$, the multi-interval exchange that we use to prove Theorem 1.1.

To verify that the assumptions are mutually compatible, we show the following.

Proposition 3.2. *Define a sequence n_k recursively by*

$$n_1 = 10 \qquad \textit{and} \qquad n_k = 10^{k^2} n_{k-1}.$$

Define α by specifying its continued fraction expansion as follows:

$$a_{n_k+1} = k + 8 \textit{ for all } k, \qquad \textit{and} \qquad a_i = 2 \textit{ if } i \notin \{n_k + 1\}_{k \in \mathbb{N}}.$$

Then all the above assumptions are satisfied.

Proof. We will establish the assumptions one at a time.

(1) Obvious.
(2) By the upper bound in Theorem 2.1, $\langle\langle q_{n_k}\alpha\rangle\rangle < \frac{1}{q_{n_k+1}}$. By the second recursive formula, and the fact that all a_i are at least 2, we see that $q_n > 2^n$. Since $n_1 = 10$, $\sum_{k=1}^{\infty} 2\langle\langle q_{n_k}\alpha\rangle\rangle < \frac{4}{2^{10}}$. The alternating property in Theorem 2.1 (using p_1/q_1 and p_2/q_2) gives that $\frac{2}{5} < \alpha < \frac{1}{2}$, so the result follows.
(3) Again by the recursive formula and the fact that all a_i are at least 2, we have $q_{i+k} > 2^k q_i$. Since $a_{n_{k+1}} > 2$ for every k, we have in particular that $q_{n_k+1} > 3q_{n_k}$, whence the inequality follows by induction. The second claim follows because $q_{n_i-1} < \frac{1}{2} q_{n_i}$ and $q_{n_{i-j}-1} < \frac{1}{4}^j q_{n_i}$.
(4) Obvious.
(5) Clear if $i \neq n_k, n_k + 1$, and otherwise it follows because $k^2 = o(\log(n_k))$ and $q_i > 2^i$.

(6) Obvious.
(7) Let i be the greatest number such that $n_i \leq n$. Then there are i numbers k such that $n_k \leq n$. Since $i+8 = a_{n_i+1}$, by Assumption 5, there are $o(\log\log q_n)$ numbers k such that $n_k \leq n$. Again by Assumption 5, each a_{n_k+1} is $o(\log\log q_n)$. So, the sum is at most $2n + o(\log\log q_n)^2$. Since n is $O(\log q_n)$, we have the claim.
(8) Since $q_{i+k} > 2^k q_i$, we have that $q_{n_{k+1}} > 2^{n_{k+1}-n_k} q_{n_k}$, so

$$\log q_{n_{k+1}} > (n_{k+1} - n_k)\log(2) + \log q_{n_k}.$$

On the other hand, $q_{i+1} \leq (a_{i+1}+1)q_i$, so we have that

$$\log(q_{n_k}) \leq \sum_{i=1}^{n_k} \log(a_i + 1) = O(n_k),$$

since there are $O(n_k)$ terms where $\log(a_i+1)$ has size $\log(3)$, and at most $O(\log(n_k))$ terms where $i = n_\ell + 1$ for some ℓ and hence a_{i+1} has size $\log(\ell+8) = O(\log(n_k))$.

Since $(n_{k+1} - n_k)/n_k \to \infty$, the result follows.
(9) Obvious.

□

4. A skew product over a rotation

This section develops the Katok-Sataev-Veech construction [**Vee69, Kat73, Sat75**] in a manner convenient for our purposes.

We now define a $\mathbb{Z}_2 = \{0,1\}$ skew product over the circle rotation R by angle α. (In the notation $\mathbb{Z}_2 = \{0,1\}$, it is implicit that 0 and 1 represent equivalent classes modulo 2.) Set

$$J = \left[0, \sum_{k=1}^{\infty} 2\langle\langle q_{n_k}\alpha\rangle\rangle\right).$$

By Assumption 2, we have that J has length less than 1.

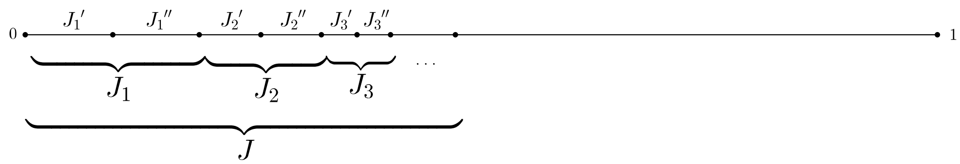
Define a skew product T to be the rotation by α skewed over the interval J, so

$$T : S^1 \times \mathbb{Z}_2 \to S^1 \times \mathbb{Z}_2, \qquad T(x,j) = (R(x), j + \chi_J(R(x))).$$

Here χ_J denotes the characteristic function of J. We will denote the coordinate projections as

$$\pi_{S^1} : S^1 \times \mathbb{Z}_2 \to S^1, \qquad \pi_{\mathbb{Z}_2} : S^1 \times \mathbb{Z}_2 \to \mathbb{Z}_2.$$

We denote by $\iota : S^1 \times \mathbb{Z}_2 \to S^1 \times \mathbb{Z}_2$ the involution $\iota(x,j) = (x, j+1)$.



Notation. For $\ell \geq 1$, set

$$J_\ell = \left[\sum_{k=1}^{\ell-1} 2\langle\langle q_{n_k}\alpha\rangle\rangle, \sum_{k=1}^{\ell} 2\langle\langle q_{n_k}\alpha\rangle\rangle\right),$$

so J is the disjoint union of the J_ℓ. Let

$$\begin{aligned} J'_\ell &= \left[\sum_{k=1}^{\ell-1} 2\langle\langle q_{n_k}\alpha\rangle\rangle, \sum_{k=1}^{\ell-1} 2\langle\langle q_{n_k}\alpha\rangle\rangle + \langle\langle q_{n_\ell}\alpha\rangle\rangle\right), \\ J''_\ell &= \left[\sum_{k=1}^{\ell-1} 2\langle\langle q_{n_k}\alpha\rangle\rangle + \langle\langle q_{n_\ell}\alpha\rangle\rangle, \sum_{k=1}^{\ell} 2\langle\langle q_{n_k}\alpha\rangle\rangle\right) \end{aligned}$$

be the left and right halves of the interval J_ℓ, so $J_\ell = J'_\ell \cup J''_\ell$.

For any subset $Y \subset S^1$, we set

$$\tilde{Y} = Y \times \mathbb{Z}^2 = \pi_{S^1}^{-1}(Y).$$

Moving a bit of one invariant set into another. Here we consider a very general construction. The notation has been chosen to match the situation to which the lemma will be applied.

Lemma 4.1. *Let $T_\ell : X \to X$ be an invertible transformation with two invariant sets U_ℓ and $V_\ell = X \setminus U_\ell$. Suppose there is an involution $\iota : X \to X$ that interchanges U_ℓ and V_ℓ. Let $\tilde{J}'_{\ell+1} \subset X$ be ι invariant, and let $q_{n_{\ell+1}}$ be any integer such that*

$$\tilde{J}'_{\ell+1}, T_\ell(\tilde{J}'_{\ell+1}), \ldots, T_\ell^{q_{n_{\ell+1}}}(\tilde{J}'_{\ell+1})$$

are disjoint. Consider the transformation $T_{\ell+1}$ on X defined by

$$T_{\ell+1}(x) = \begin{cases} \iota(T_\ell(x)) & \text{if } T_\ell(x) \in \tilde{J}'_{\ell+1} \cup T_\ell^{q_{n_{\ell+1}}}(\tilde{J}'_{\ell+1}) \\ T_\ell(x) & \text{otherwise.} \end{cases}$$

Then the set

$$U_{\ell+1} = \left(U_\ell \cup \bigcup_{i=0}^{q_{n_{\ell+1}}-1} T_\ell^i(\tilde{J}'_{\ell+1} \cap V_\ell)\right) \setminus \bigcup_{i=0}^{q_{n_{\ell+1}}-1} T_\ell^i(\tilde{J}'_{\ell+1} \cap U_\ell)$$

is $T_{\ell+1}$ invariant, as is $V_{\ell+1} = X \setminus U_{\ell+1} = \iota(U_{\ell+1})$.

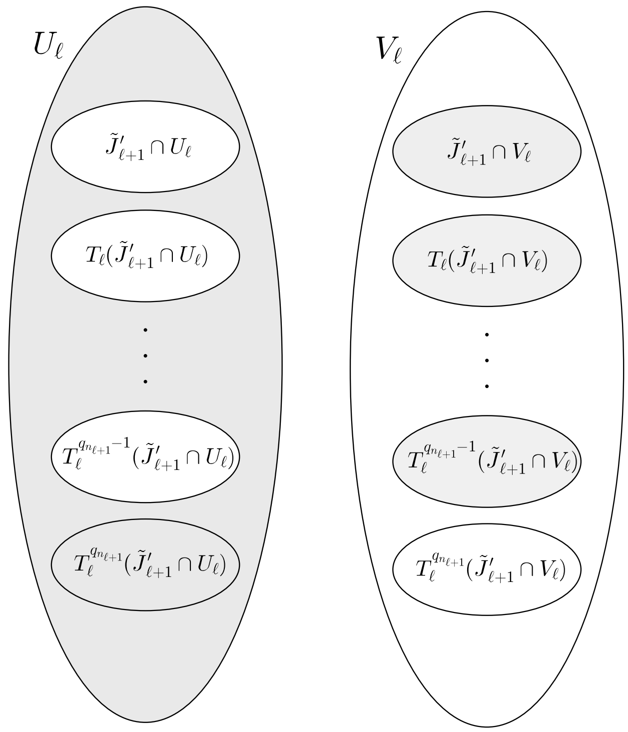


FIGURE 4.1. Schematic picture of Lemma 4.1. The set $U_{\ell+1}$ is shaded.

The proof of Lemma 4.1 is left to the reader, who is invited to convince himself or herself while looking at Figure 4.1.

Non-minimal approximants. Define $T_\ell : [0,1] \times \mathbb{Z}_2 \to [0,1] \times \mathbb{Z}_2$ to be the rotation by α skewed over the interval $[0, \sum_{k=1}^{\ell} 2\langle\langle q_{n_k}\alpha\rangle\rangle)$, so

$$T_\ell(x,j) = (R(x), j + \chi_{[0,\sum_{k=1}^{\ell} 2\langle\langle q_{n_k}\alpha\rangle\rangle)}(R(x))).$$

Set $U_0 = [0,1] \times \{1\}$ and $V_0 = [0,1] \times \{0\}$, and for each $\ell \geq 0$ define

$$U_{\ell+1} = \left(U_\ell \cup \bigcup_{i=0}^{q_{n_{\ell+1}}-1} T_\ell^i(\tilde{J}'_{\ell+1} \cap V_\ell) \right) \setminus \bigcup_{i=0}^{q_{n_{\ell+1}}-1} T_\ell^i(\tilde{J}'_{\ell+1} \cap U_\ell)$$

and $V_{\ell+1} = U_{\ell+1}^c$.

We will show that U_ℓ and V_ℓ are T_ℓ invariant sets, which implies T_ℓ is not minimal. Because T_ℓ is not minimal, and is equal to T on a set of large measure, we refer to T_ℓ as a non-minimal approximant to

T. Our understanding of T will follow from a study of the T_ℓ. Our understanding of T_ℓ will be inductive, using the fact that T_ℓ is almost the same as $T_{\ell-1}$ but with additional skewing.

Lemma 4.2. *For each $\ell > 1$ the intervals*

$$R^i(J'_\ell), \quad i = 0, \dots, q_{n_\ell}$$

are disjoint from each other and $[0, \langle\langle q_{n_\ell}\alpha\rangle\rangle)$ and $[1 - \langle\langle q_{n_\ell}\alpha\rangle\rangle, 1)$.

Recall $J'_\ell = [\sum_{k=1}^{\ell-1} 2\langle\langle q_{n_k}\alpha\rangle\rangle, \sum_{k=1}^{\ell-1} 2\langle\langle q_{n_k}\alpha\rangle\rangle + \langle\langle q_{n_\ell}\alpha\rangle\rangle)$.

Proof. Since $J'_\ell = R^{N_0}([0, \langle\langle q_{n_\ell}\alpha\rangle\rangle))$ when $N_0 = 2\sum_{k=1}^{\ell-1} q_{n_k}$, the intervals in question are

$$R^N([0, \langle\langle q_{n_\ell}\alpha\rangle\rangle)), \quad N = -q_{n_\ell} \text{ and } N = 0 \text{ and } N = N_0, \dots, N_0 + q_{n_\ell}.$$

Thus the intervals in question are contained in the $N_0 + 2q_{n_\ell} + 1$ orbit of an interval of length $\langle\langle q_{n_\ell}\alpha\rangle\rangle$. By Assumption 3, $N_0 + 2q_{n_\ell} + 1 < q_{n_\ell+1}$, so the separation property gives that these intervals are disjoint. ∎

Lemma 4.3. *U_ℓ and V_ℓ are T_ℓ invariant, and $V_\ell = \iota(U_\ell)$.*

Proof. This is proven inductively using Lemma 4.1, where the notation has been chosen to indicate how the result should now be applied.

$T_{\ell+1}$ is obtained from T_ℓ by additionally skewing over $J_{\ell+1} = J'_{\ell+1} \cup R^{q_{n_{\ell+1}}}(J'_{\ell+1})$. This additional skewing amounts to applying ι every time the orbit lands in $\tilde{J}'_{\ell+1} \cup T_\ell^{q_{n_{\ell+1}}}(\tilde{J}'_{\ell+1})$. Thus the definition of $T_{\ell+1}$ as a skew product coincides with the inductive definition of $T_{\ell+1}$ provided in Lemma 4.1.

The disjointness condition in Lemma 4.1 has been verified in the previous lemma. ∎

Theorem 4.4. *For each integer $\ell \geq 1$, there are sets U_ℓ, V_ℓ, defined above, whose disjoint union is $S^1 \times \mathbb{Z}_2$ such that $\iota(U_\ell) = V_\ell$ and the following properties hold.*

(1) U_ℓ contains

$$[0, \langle\langle q_{n_\ell}\alpha\rangle\rangle) \times \{0\} \quad \textit{and} \quad [1 - \langle\langle q_{n_\ell}\alpha\rangle\rangle, 1) \times \{1\}.$$

(2) There is a subinterval J'_ℓ of S^1, defined above, of length $\langle\langle q_{n_\ell}\alpha\rangle\rangle$ such that $\pi_{S_1}(U_\ell \setminus U_{\ell-1}) = \cup_{i=0}^{q_{n_\ell}-1} R^i(J'_\ell)$.

Remark 4.5. In can be shown (using (1) and (2) and Theorem 2.2) that U_ℓ contains most of $[0, \langle\langle q_{n_{\ell-1}}\alpha\rangle\rangle)$, i.e. that $[0, \langle\langle q_{n_{\ell-1}}\alpha\rangle\rangle) \setminus U_\ell$ has small measure. This is important in the proof of Claim (3) of Theorem 6.1 (see Lemma 5.8).

Proof of Theorem 4.4. The first part of the first claim is true by induction. The base case is $\ell = 1$. The inductive step follows from the definition of $U_{\ell+1}$ as well as the disjointness of $\{R^i(J'_\ell)\}_{i=0}^{q_{n_\ell}-1}$ from $[0, \langle\langle q_{n_\ell}\alpha\rangle\rangle)$ and $[1-\langle\langle q_{n_\ell}\alpha\rangle\rangle, 1)$ in Lemma 4.2.

The second claim is by the definition of U_ℓ. □

5. Birkhoff sums for the rotation R

In this section we consider the Birkhoff sums of the function $g(x) = \frac{1}{x}$ over the rotation R. Estimates on these sums are used in the next section to provide the shearing estimates.

Results on Birkhoff sums over rotations have previously been used to prove mixing in a slightly different setting (asymmetric singularities), see for example [**FK, Koc03, SK92**].

Lemma 5.1. *For each positive integer N, there are unique integers b_n such that*

$$N = \sum_{n=1}^{k} b_n q_n,$$

and such that $0 \le b_n \le a_{n+1}$ and $q_n > \sum_{i=0}^{n-1} b_i q_i$ for each n.

Such an expression is called the Ostrowski expansion of N [**Ost22**].

Proof. Pick k such that $q_{k+1} > N \ge q_k$, and let b_k be the unique integer such that $0 \le N - b_k q_k < q_k$. By the second recursive formula in Theorem 2.1,

$$N - a_{k+1}q_k < q_{k+1} - a_{k+1}q_k = q_{k-1}.$$

So we get that $b_k \le a_{k+1}$. Replacing N with $N - b_k q_k$ and iterating this procedure gives the Ostrowski expansion for N. □

Proposition 5.2. *Let $g : (1, 0] \to [0, \infty)$ be a monotone decreasing function. Let*

$$g_N(x) = \max_{i=0,\ldots,N-1} g(R^i(x)).$$

Let $q_{k+1} > N \ge q_k$, and let C_N be any number satisfying $1 \le C_N < 2q_{k+1}$. Then, without any assumptions on α,

$$\begin{aligned}\sum_{i=0}^{N-1} g(R^i x) &= g_N(x) + N\int_{\frac{C_N}{2q_{k+1}}}^{1} g \\ &+ O\left(g\left(\frac{C_N}{2q_{k+1}}\right)\sum_{i=2}^{k+1} a_i + \sum_{i=1}^{\lfloor C_N\rfloor} g\left(\frac{i}{2q_{k+1}}\right)\right).\end{aligned}$$

Remark 5.3. Let Ψ and Q be functions. We say that Q is $O(\Psi)$ if there exists a constant $C > 0$ such that $|Q| \leq C|\Psi|$ for all allowed values of all variables which appear in Ψ and Q. The constant C is independent of everything. If Ψ and Q are functions of N as well as possibly other variables, we say that Q is $o(\Psi)$ if for all $\varepsilon > 0$ there exists a constant $N_0(\varepsilon)$ such that when $N > N_0(\varepsilon)$ we have $|Q| \leq \varepsilon|\Psi|$, for any allowed values of the other variables other than N. The constant $N_0(\varepsilon)$ does not depend on any variable other than ε. For example, Proposition 5.2 asserts that the quantity

$$Q = \sum_{i=0}^{N-1} g(R^i x) - \left(g_N(x) + N \int_{\frac{C_N}{2q_{k+1}}}^{1} g \right)$$

is at most $C|\Psi|$ for some C, where

$$\Psi = g\left(\frac{C_N}{2q_{k+1}}\right) \sum_{i=2}^{k+1} a_i + \sum_{i=1}^{\lfloor C_N \rfloor} g\left(\frac{i}{2q_{k+1}}\right)$$

and the constant C is independent of α, C_N and x.

Proof. Since $N < q_{k+1}$, each orbit segment of length N is $\langle\langle q_k \alpha \rangle\rangle \geq \frac{1}{2q_{k+1}}$ separated. Thus,

$$-g_N(x) + \sum_{i=0}^{N-1} g(R^i x)$$

is within $g\left(\frac{1}{2q_{k+1}}\right)$ of

$$\sum_{i=0}^{N-1} g(R^i x) \chi_{\left[\frac{1}{2q_{k+1}}, 1\right)}(R^i x).$$

Indeed, if $g_N(x) > 2q_{k+1}$ the two expressions above are equal and otherwise the claim is immediate by monotonicity.

Now, note that by separation and monotonicity,

$$\sum_{i=0}^{N-1} g(R^i x) \chi_{\left[\frac{1}{2q_{k+1}}, \frac{C_N}{2q_{k+1}}\right)}(R^i x) \quad \leq \quad \sum_{i=1}^{\lfloor C_N \rfloor} g\left(\frac{i}{2q_{k+1}}\right).$$

Now, consider the Ostrowski expansion $N = \sum_{n=1}^{k} b_n q_n$, and consider an orbit segment of length N to be built from b_n orbit segments of length q_n (for $n = 1, \ldots, k$). Applying Denjoy-Koksma individually to these $\sum_{n=1}^{k} b_n$ orbit segments, we get

$$\sum_{i=0}^{N-1} g(R^i x)\chi_{\left[\frac{C_N}{2q_{k+1}},1\right)}(R^i x) = N\int_{\frac{C_N}{2q_{k+1}}}^{1} g + O\left(g\left(\frac{C_N}{2q_{k+1}}\right)\sum_{i=2}^{k+1} a_i\right).$$

Here we have used the estimate $\sum_{n=1}^{k} b_n \leq \sum_{i=2}^{k+1} a_i$, which is immediate from the definition of the Ostrowski expansion. Denjoy-Koksma gives that the implied constant in the $O(\cdot)$ notation does not depend on anything. □

Corollary 5.4. *With the assumptions in this paper on α, if $g(x) = 1/x$, then for all N*

$$\sum_{i=0}^{N-1} g(R^i x) \quad = \quad g_N(x) + N\log(N) + o\Big(N\big(\log\left(\log(N)\right)\big)^2\Big)$$

and

$$\sum_{i=0}^{N-1} g'(R^i x) \quad = \quad g'_N(x) + o\left((N\log\log(N))^2\right).$$

Remark 5.5. We emphasize that here and for the rest of the paper α is considered to be a constant; we do not assert that the implied estimates in the $O(\cdot)$ and $o(\cdot)$ notation (the choice of C and $N_0(\varepsilon)$ in Remark 5.3) are independent of our specific choice of α.

Proof. Choose k such that $q_k \leq N < q_{k+1}$. Let $1 < C_N \leq q_{k+1}$ be a constant, and let us address the different quantities appearing in the previous proposition.

Recall Assumption 5, which gives that $q_{k+1}/N = o(\log\log(N))$, and note

$$\begin{aligned}\int_{\frac{C_N}{2q_{k+1}}}^{1} g \quad &= \quad \log(N) + \log(q_{k+1}/N) + \log(2) - \log(C_N) \\ &= \quad \log(N) - \log(C_N) + \log(o(\log\log(N))).\end{aligned}$$

Next, note that the same bound for q_{k+1}/N gives

$$\begin{aligned}\sum_{i=1}^{\lfloor C_N \rfloor} g\left(\frac{i}{2q_{k+1}}\right) \quad &= \quad 2q_{k+1}\sum_{i=1}^{\lfloor C_N \rfloor}\frac{1}{i} \\ &= \quad 2q_{k+1}\log(C_N) + O(q_{k+1}) \\ &= \quad o(\log\log(N)N\log C_N).\end{aligned}$$

Finally, note that Assumption 7 gives that $\sum_{i=2}^{k+1} a_i = O(\log(N))$, and hence

$$\begin{aligned} g\left(\frac{C_N}{2q_{k+1}}\right)\sum_{i=2}^{k+1} a_i &= O\left(\frac{2q_{k+1}}{C_N}\log(N)\right) \\ &= o\left(\frac{N\log\log(N)}{C_N}\log(N)\right). \end{aligned}$$

Now setting $C_N = \log(N)$ and using the previous proposition gives the result.

The second bound is similar. ☐

Lemma 5.6. *Set $g(x) = 1/x$. For any k with $q_{n_k} \le N$ and any $x \in [0,1)$ we have*

$$\sum_{i=0}^{N-1} g(R^i x)\chi_{[\langle\langle q_{n_{k-1}}\alpha\rangle\rangle,1)}(R^i x) = o(N\log(N))$$

and

$$\sum_{i=0}^{N-1} g'(R^i x)\chi_{[\langle\langle q_{n_{k-1}}\alpha\rangle\rangle,1)}(R^i x) = O(N^2).$$

Remark 5.7. In Lemma 5.6 we continue to use the notational convention described in Remark 5.3. For example, this lemma asserts that

$$Q = \sum_{i=0}^{N-1} g(R^i x)\chi_{[\langle\langle q_{n_{k-1}}\alpha\rangle\rangle,1)}(R^i x)$$

is $o(\Psi)$ with $\Psi = N\log(N)$, which means that for all $\varepsilon > 0$ there exists a $N_0(\varepsilon)$ such that when $N > N_0(\varepsilon)$ then $|Q| \le \varepsilon|\Psi|$ for all $x \in S^1$ and all k such that $q_{n_k} \le N$. We emphasize that the constant $N_0(\varepsilon)$ does not depend on x or k. We also emphasize that the allowed values for k are k such that $q_{n_k} \le N$; we do not assert that the inequality $|Q| \le \varepsilon|\Psi|$ is true when k does not satisfy this condition.

Proof. Let $N = \sum_{n=1}^{L} b_n q_n$ be the Ostrowski expansion of N. As in the previous proof, using Assumption 7 we get $\sum_{n=1}^{L} b_n = O(\log(N))$. Using Denjoy-Koksma, for each of the b_n orbit segments of length q_n and summing from $n = 1$ to $n = L$ we get

$$\begin{aligned} \sum_{i=0}^{N-1} g(R^i x)\chi_{[\langle\langle q_{n_{k-1}}\alpha\rangle\rangle,1)}(R^i x) &= O\left(-N\log\langle\langle q_{n_{k-1}}\alpha\rangle\rangle + \frac{\log(N)}{\langle\langle q_{n_{k-1}}\alpha\rangle\rangle}\right) \\ &= O\left(N\log(q_{n_{k-1}+1}) + \log(N)q_{n_{k-1}+1}\right). \end{aligned}$$

The implied constant in the $O(\cdot)$ notation does not depend on x or k because the bounds provided by Denjoy-Koksma do not depend on x or k. Assumption 8, which gives $\log(q_{n_{k-1}+1}) = o(\log(N))$, gives the result.

Similarly, to prove the second estimate it suffices to note

$$\begin{aligned}\sum_{i=0}^{N-1} g'(R^i x)\chi_{[\langle\langle q_{n_{k-1}}\alpha\rangle\rangle,1)}(R^i x) &= O\left(\frac{N}{\langle\langle q_{n_{k-1}}\alpha\rangle\rangle} + \frac{\log(N)}{\langle\langle q_{n_{k-1}}\alpha\rangle\rangle^2}\right)\\ &= O\left(Nq_{n_{k-1}+1} + \log(N)q^2_{n_{k-1}+1}\right)\\ &= O(N^2).\end{aligned}$$

The last equality follows by invoking Assumption 8, which states that $\log(q_{n_{k-1}+1}) = o(\log(q_{n_k}))$ and hence implies that $Nq^2_{n_{k-1}+1} = O(N^{1+\epsilon})$ for all $\epsilon > 0$. $\square$

Lemma 5.8. *Set $g(x) = 1/x$. For any large enough ℓ, for any $N > q_{n_\ell}$, let S be an orbit of length q_{n_ℓ} of an interval of length $\langle\langle q_{n_\ell}\alpha\rangle\rangle$. Assume S is disjoint from $[0, \langle\langle q_{n_\ell}\alpha\rangle\rangle)$, and let x be any point disjoint from $\cup_{i=0}^{\sqrt{a_{n_\ell+1}}q_{n_\ell}} R^{-i}(S)$. Then*

$$\sum_{i=0}^{N-1} g(R^i x)\chi_S(R^i x) = o(N\log(N))$$

and

$$\sum_{i=0}^{N-1} g'(R^i x)\chi_S(R^i x) = o(N^2\log(N)^{\frac{1}{3}}).$$

Proof. Denjoy-Koksma gives that the sum of

$$g(x)\chi_{[\langle\langle q_{n_\ell}\alpha\rangle\rangle,1)}(x)$$

over an orbit of length q_{n_ℓ} is at most

$$\begin{aligned}(5.0.1)\quad O(q_{n_\ell}\log q_{n_\ell+1} + q_{n_\ell+1}) = O(q_{n_\ell}\log q_{n_\ell+1}) =\\ O(q_{n_\ell}\log((a_{n_\ell+1}+1)q_{n_\ell})) = O(q_{n_\ell}\log q_{n_\ell}).\end{aligned}$$

The last equality follows because $a_{n_\ell} + 1 = o(q_{n_\ell})$. Note that, by the separation property in Theorem 2.2, each orbit of length $q_{n_\ell+1} > a_{n_\ell+1}q_{n_\ell}$ can hit each interval of S at most once. A point x as in the lemma stays outside of S for time at least $\sqrt{a_{n_\ell+1}}q_{n_\ell}$, then makes a pass through S, then stays outside of S for time at least $q_{n+1} - q_n > \sqrt{a_{n_\ell+1}}q_{n_\ell}$ (Assumption 4), then makes a pass through S, etc. (We refer to a pass through S as an orbit segment that hits each interval of S exactly once.) Therefore, if the orbit makes $m-1$ full passes through

S, plus possibly a final partial pass through, then by (5.0.1), the first sum in the lemma statement is at most

$$mO(q_{n_\ell} \log q_{n_\ell}),$$

while N is at least

$$m\sqrt{a_{n_\ell+1}}q_{n_\ell}.$$

So $N \log N$ is at least

$$m\sqrt{a_{n_\ell+1}}q_{n_\ell} \log(q_{n_\ell}),$$

whence the result follows by Assumption 4.

Similarly, Denjoy-Koksma gives that the sum of

$$g'(x)\chi_{[\langle\langle q_{n_\ell}\alpha\rangle\rangle,1)}(x)$$

over an orbit of length q_{n_ℓ} is at most

$$O(q_{n_\ell}q_{n_\ell+1} + q_{n_\ell+1}^2) = O(a_{n_\ell+1}^2 q_{n_\ell}^2).$$

As before, if the orbit makes $m-1$ full passes through S, plus possibly a final partial pass though, then the sum is at most

$$mO(a_{n_\ell+1}^2 q_{n_\ell}^2),$$

while N is at least

$$m\sqrt{a_{n_\ell+1}}q_{n_\ell}.$$

So $N^2(\log N)^{\frac{1}{3}}$ is at least

$$m^2 a_{n_\ell+1} q_{n_\ell}^2 \log(q_{n_\ell})^{\frac{1}{3}},$$

whence the result follows by Assumption 5. □

6. Birkhoff sums for the skew product T

We now define a function $f : S^1 \times \mathbb{Z}_2 \to \mathbb{R}$, which will serve as the roof function for a suspension flow over T. Recall that d denotes distance on $S^1 \times \mathbb{Z}_2$. Also note that

$$T(x, j) = (R(x), j + \chi_J(R(x)))$$

has discontinuities at $R^{-1}(0) \times \mathbb{Z}_2$ and $R^{-1}(|J|) \times \mathbb{Z}_2$. We will define f to have logarithmic singularities over these discontinuities, as well as an additional logarithmic singularity over $(0, 1) \in S^1 \times \mathbb{Z}_2$.

$$\begin{aligned} f(x,j) = 1 \quad &+ \quad |\log(d(x, R^{-1}(0)))| \\ &+ \quad |\log(d(x, R^{-1}(|J|)))| \\ &+ \quad \chi_{\{1\}}(j) \cdot |\log(d(x, 0))| \end{aligned}$$

Let λ denote Lebesgue probably measure on $S^1 \times \mathbb{Z}_2$. In the next theorem we continue to use the notational convention described in Remark 5.3.

Theorem 6.1. *For all large enough M there exists $G(M) \subset S^1 \times \mathbb{Z}_2$ such that the following hold.*

(1) $G(M)$ is the disjoint union of intervals of length at least $\frac{2}{M\sqrt{\log(M)}}$.

(2) $\lambda(G(M)) \to 1$.

(3) For any N with $N \in [\frac{M}{2}, 2M]$ and for any $p \in G(M)$, either

$$\sum_{i=0}^{N-1} f'(T^i p) - N\log(N) = o(N \log N)$$

or

$$\sum_{i=0}^{N-1} f'(T^i p) + N\log(N) = o(N \log N).$$

(4) For any N with $N \in [\frac{M}{2}, 2M]$ and for any $p \in G(M)$,

$$\sum_{i=0}^{N-1} f''(T^i p) = o(\log(N)^{\frac{1}{3}} N^2).$$

(5) For any $k < 2M$ and $p \in G(M)$,

$$\sum_{i=k}^{k+5\sqrt{\log(M)}} \left|f'(T^i p)\right| = o(M\sqrt{\log(M)}).$$

(6) T^i is continuous on each interval of $G(M)$ for all $0 \leq i \leq 2M + 5\sqrt{\log(M)}$.

The most important conditions are the first three. In the third, it is implicit that the sign is constant on each interval in $G(M)$.

Remark 6.2. In Theorem 6.1 we continue to use the notational convention described in Remark 5.3, now with M as the main variable. For example, part (3) of this theorem asserts that

$$Q = \sum_{i=0}^{N-1} f'(T^i p) \pm N\log(N)$$

is $o(\Psi)$ with $\Psi = N\log N$, which means that for all $\varepsilon > 0$ there exists a $M_0(\varepsilon)$ such that when $M \geq M_0(\varepsilon)$ then $|Q| \leq \varepsilon|\Psi|$ for all $N \in [\frac{M}{2}, 2M]$ and $p \in G(M)$. For clarity, explicitness, if $M > M_0(\epsilon)$, then for any $N \in [\frac{M}{2}, 2M]$ and $p \in G(M)$ we have $|\sum_{i=0}^{N-1} f'(T^i p) - N\log(N)| < \epsilon N\log(N)$ or $|\sum_{i=0}^{N-1} f'(T^i p) + N\log(N)| < \epsilon N\log(N)$.

Proof. Pick ℓ such that $q_{n_\ell} < \frac{M}{2} < q_{n_{\ell+1}}$. Let $Q_M \subset S^1$ be the points within distance $\frac{1}{M(\log M)^{\frac{1}{12}}}$ of the (projections to S^1 of) singularities of f. Define the "bad set" in S^1 to be

$$\begin{aligned} B_{S^1}(M) &= \bigcup_{i=0}^{2M+5\sqrt{\log(M)}} R^{-i}(Q_M) \\ &\cup \bigcup_{i=0}^{2M+5\sqrt{\log(M)}} R^{-i}(\cup_{j=\ell+1}^{\infty} J_j) \\ &\cup \bigcup_{i=-q_{n_\ell}}^{\sqrt{a_{n_\ell+1}}q_{n_\ell}} R^{-i}(J'_\ell). \end{aligned}$$

Set $B(M) = B_{S^1}(M) \times \mathbb{Z}_2$. The complement of the bad set $B(M)^c$ is a union of disjoint intervals. We define the good set $G(M)$ to be the union of all those intervals in $B(M)^c$ of length at least $\frac{2}{M\sqrt{\log(M)}}$.

Remark 6.3. We chose the neighborhood of the discontinuities in the definition of Q_M to be $\frac{1}{N(\log(N))^{1/12}}$ so that first the measure will be negligible and second, so that

$$\left(\frac{1}{N(\log(N))^{1/12}}\right)^{-1} = o(N\log(N)) = o(M\sqrt{\log(M)})$$

and $\left(\frac{1}{N(\log(N))^{1/12}}\right)^{-2} = o(N^2\log(N)^{1/3})$. These are used in estimates that invoke Corollary 5.4 in the proofs of Claims 3, 4 and 5.

Remark 6.4. The "extra" $5\sqrt{\log(M)}$ in the expression $2M+5\sqrt{\log(M)}$ above is not needed in this section, but will be convenient in Section 8, for example in Lemma 8.2.

Claims (1) and (2): Recall that $\cup_{j=\ell+1}^{\infty} J_j$ is an interval of size

$$\sum_{k=\ell+1}^{\infty} 2\langle\langle q_{n_k}\alpha\rangle\rangle \leq \sum_{k=\ell+1}^{\infty} \frac{2}{q_{n_k+1}} \leq \frac{4}{q_{n_{\ell+1}+1}}$$

by the exponential growth of the q_{n_k}, for example by Assumption 3. Since $\frac{M}{2} < q_{n_{\ell+1}}$, Assumption 4 gives that $\frac{4}{q_{n_{\ell+1}+1}} = o(1/M)$. Hence the measure of the second union in $B_{S^1}(M)$ goes to zero.

The third union has size at most three times

$$\frac{\sqrt{a_{n_\ell+1}}q_{n_\ell}}{q_{n_\ell+1}} = \frac{\sqrt{a_{n_\ell+1}}q_{n_\ell}}{a_{n_\ell+1}q_{n_\ell} + q_{n_\ell-1}} \to 0$$

by Assumption 4.

The first union obviously has size $o(1)$, so in total we see that $B(M)^c$ has measure $1-o(1)$. Note also that $B(M)$ is the disjoint union of $O(M)$ intervals. It remains only to show that the subset covered by intervals of length at least $\frac{2}{M\sqrt{\log(M)}}$ has measure going to 1. This is true since the complement has measure at most $O\left(\frac{M}{M\sqrt{\log(M)}}\right)=o(1)$ (the total number of such intervals, times max length).

Claim (3): Let $g(x)=1/\langle\langle x\rangle\rangle$, so for example $g(-0.1)=\frac{10}{9}$. The difference between $f'(x,j)$ and

$$\begin{aligned}&g(-x+R^{-1}(0)))-g(x-R^{-1}(0))\\&+g(-x+R^{-1}(|J|))-g(x-R^{-1}(|J|))\\&+\chi_{\{1\}}(j)g(1-x)-\chi_{\{1\}}(j)g(x)\end{aligned}$$

is a bounded function, whose first and second derivatives are bounded. Since the derivatives of the difference are bounded, the N^{th} Birkhoff sum of the difference is $O(N)$, and so it suffices to prove claims (3), (4) and (5) for this function. If $x\in G(M)$, we show that Corollary 5.4 gives that the N^{th} Birkhoff sums of the function

$$\begin{aligned}&g(-x+R^{-1}(0))-g(x-R^{-1}(0))\\&+g(-x+R^{-1}(|J|))-g(x-R^{-1}(|J|))\end{aligned}$$

and its derivative are sufficiently small that they may be ignored. Indeed, the consecutive terms have that the $N\log(N)$ terms appear with opposite signs and by our assumption on the set $G(M)$ the other terms are $o(N\log(N))$. In particular, the first term in the union defining $B_{S^1}(M)$ means that g_N is at most $N\log(N)^{\frac{1}{12}}$.

So it remains to consider Birkhoff sums of

$$\chi_{\{1\}}(j)g(-x)-\chi_{\{1\}}(j)g(x),$$

for $x\notin B_{S^1}(M)$. By Lemma 5.6 it suffices to study Birkhoff sums instead of the function

$$(\chi_{\{1\}}(j)g(-x)-\chi_{\{1\}}(j)g(x))\chi_{(0,\langle\langle q_{n_{\ell-1}}\alpha\rangle\rangle]\cup[1-\langle\langle q_{n_{\ell-1}}\alpha\rangle\rangle,1)}(x). \tag{6.0.1}$$

The assumption that

$$B_{S^1}(M)\supset\bigcup_{i=0}^{2M+5\sqrt{\log(M)}}R^{-i}(\cup_{j=\ell+1}^{\infty}J_j)$$

implies that the orbit of p of length $2M$ stays in either U_ℓ or V_ℓ.

Remark 6.5. This will allow us to understand the Birkhoff sum of $p \in G(M)$ by leveraging understanding how V_ℓ and U_ℓ are distributed. Our argument will not directly reference the non-minimal approximants T_ℓ discussed in Section 4, but instead will use estimates from Section 5 together with results from Section 4 on the sets V_ℓ and U_ℓ.

We now prove the following estimate for points $(x,j) \in V_\ell \cap G(M)$,

$$\sum_{i=0}^{N-1} g(R^i(x))\chi_{[0,\langle\langle q_{n_{\ell-1}}\alpha\rangle\rangle)\times\{1\}}(T^i(x,j)) = N\log(N) + o(N\log N), \tag{6.0.2}$$

by explaining how this estimate can be deduced from Corollary 5.4, Lemma 5.6, and Lemma 5.8.

Indeed, first note that our assumption that

$$B_{S^1} \supset \bigcup_{i=0}^{2M+5\sqrt{\log(M)}} R^{-i}(Q_M)\times\mathbb{Z}_2$$

implies that $R^i x \notin Q_M$ for $0 \le i \le N$ and so

$$g_N(x) = O(M(\log M)^{\frac{1}{12}}) = o(N\log(N)).$$

So Corollary 5.4 and Lemma 5.6 imply that

$$\sum_{i=0}^{N-1} g(R^i(x))\chi_{[0,\langle\langle q_{n_{\ell-1}}\alpha\rangle\rangle)}(R^i(x)) = N\log(N) + o(N\log N). \tag{6.0.3}$$

Next, recall that, as we remarked above, the orbit of (x,j) up to time $2M$ remains in V_ℓ.

Finally, denote by S the the projection of

$$\big([0,\langle\langle q_{n_{\ell-1}}\alpha\rangle\rangle)\times\{1\}\big)\setminus V_\ell$$

to S^1, and note that the difference between the left hand side of (6.0.3) and the left hand side of (6.0.2) is

$$\sum_{i=0}^{N-1} g(R^i(x))\chi_S(R^i x). \tag{6.0.4}$$

So to prove equation (6.0.2), it remains to show that (6.0.4) is $o(N\log N)$.

Recall that Theorem 4.4 asserts that

(1) V_ℓ contains all of $[0,\langle\langle q_{n_\ell}\alpha\rangle\rangle)\times\{1\}$, (Conclusion (1)) and
(2) V_ℓ contains all of $[0,\langle\langle q_{n_{\ell-1}}\alpha\rangle\rangle)\times\{1\}$ except a set that is the orbit of length q_{n_ℓ} of an interval of size $\langle\langle q_{n_\ell}\alpha\rangle\rangle$. (Using that $[0,\langle\langle q_{n_{\ell-1}}\alpha\rangle\rangle) \subset V_{\ell-1}$ and Conclusion (2). The interval is J'_ℓ.)

So the S defined above satisfies the assumption on S in Lemma 5.8. Because

$$B_{S^1}(M) \supset \bigcup_{i=-q_{n_\ell}}^{\sqrt{a_{n_\ell+1}}q_{n_\ell}} R^{-i}(J'_\ell),$$

$(x,j) \in G(M)$ implies that x satisfies the assumption on the point in Lemma 5.8. (Note that $S = \cup_{i=-(q_{n_\ell}-1)}^{0} R^{-i} J'_\ell$.) So if $(x,j) \in G(M) \cap V_\ell$, then Lemma 5.8 implies that (6.0.4) is $o(N \log N)$, which concludes the proof of (6.0.2).

Similarly one obtains:

$$\sum_{i=0}^{N-1} g(-R^i x)\chi_{[1-\langle\langle q_{n_{\ell-1}}\alpha\rangle\rangle,1)\times\{1\}}(T^i(x,j)) = o(N \log N). \tag{6.0.5}$$

Indeed, by Theorem 4.4, V_ℓ is disjoint from $[1-\langle\langle q_{n_{\ell-1}}\alpha\rangle\rangle,1)\times\{1\}$ except an orbit of length q_{n_ℓ} of an interval of size $\langle\langle q_{n_\ell}\alpha\rangle\rangle$. Furthermore, by Lemma 4.2, this orbit of length q_{n_ℓ} of an interval of size $\langle\langle q_{n_\ell}\alpha\rangle\rangle$ is disjoint from $[1-\langle\langle q_{n_\ell}\alpha\rangle\rangle,1)\times\{1\}$. So by a corresponding version of Lemma 5.8 for the function $\hat{g}(x) = \frac{1}{1-x}$ the estimate follows.

Combining (6.0.2) and (6.0.5) implies the desired estimate on (6.0.1) for $(x,j) \in V_\ell \cap G(M)$. The case $(x',j') \in U_\ell \cap G(M)$ is similar and we now sketch it to conclude the proof of Claim (3). First, we establish

$$\sum_{i=0}^{N-1} g(-R^i x')\chi_{[1-\langle\langle q_{n_{\ell-1}}\alpha\rangle\rangle,1)\times\{0\}}(T^i(x',j')) = N\log(N) + o(N\log(N)). \tag{6.0.6}$$

Symmetrically to the above, $[1-\langle\langle q_{n_{\ell-1}}\alpha\rangle\rangle,1)\times\{1\}$ is contained in U_ℓ except an orbit of length q_{n_ℓ} of an interval of size $\langle\langle q_{n_\ell}\alpha\rangle\rangle$. Furthermore, by Lemma 4.2, this orbit of length q_{n_ℓ} of an interval of size $\langle\langle q_{n_\ell}\alpha\rangle\rangle$ is disjoint from $[1-\langle\langle q_{n_\ell}\alpha\rangle\rangle,1)\times\{1\}$. So by a corresponding version of Lemma 5.8 for the function $\hat{g}(x) = \frac{1}{1-x}$ we have (6.0.6). Now by symmetry,

$$\sum_{i=0}^{N-1} g(R^i x')\chi_{[0,\langle\langle q_{n_\ell}\alpha\rangle\rangle)\times\{1\}}(T^i(x',j')) = \sum_{i=0}^{N-1} g(R^i x')\chi_S(R^i x')$$

which we have seen is $o(N\log(N))$. This completes Claim (3).

Claim (4): The proof of claim (4) is very similar to that of claim (3).

Claim (5): It suffices to bound the Birkhoff sums of

$$\begin{aligned} & g(-x+R^{-1}(0))) - g(x-R^{-1}(0)) \\ + \quad & g(-x+R^{-1}(|J|)) - g(x-R^{-1}(|J|)) \\ + \quad & g(1-x) - g(x) \end{aligned}$$

which is strictly larger than $|f|$. Thus Corollary 5.4 gives the bound, since the closest hit term contributes at most $M(\log M)^{\frac{1}{12}}$, and the usual main term is of lower order because the orbit segment has length $\sqrt{M}$.

Claim (6): This follows because $B(M)$ contains all points that orbit into a discontinuity in $2M + 5\sqrt{\log(M)}$ iterates of T. $\square$

7. Unique ergodicity

Proposition 7.1. *T is uniquely ergodic.*

It is likely that this theorem can be derived from work of Treviño [**Tre14**, Theorem 3]. (To do this one would find a flat surface ω where T arises as a first return map of the vertical flow. One would then determine the systoles of $g_t\omega$.) We apply a different approach.

Skew products such as T, of circle rotations, skewed over intervals, arise in the study of genus 2 translation surfaces that are covers of tori branched over two points, and their ergodicity has been extensively studied; see for example [**CHM11**] and the references therein.

Define

$$S_k = \bigcup_{i=0}^{q_{n_k-1}-1} R^i(J_k') \times \mathbb{Z}_2. \tag{7.0.1}$$

By Assumption 9,

$$\frac{1}{3}q_{n_k} = \frac{1}{3}(2q_{n_k-1} + q_{n_k-2}) < q_{n_k-1} < \frac{1}{2}q_{n_k}.$$

If $p \in S_k$, then p is either in $U_k \cap V_{k-1}$ or $V_k \cap U_{k-1}$, and by the last inequality and Theorem 4.4 part 2, the orbit of p for time at least q_{n_k-1} remains either exclusively in $U_k \cap V_{k-1}$ or exclusively in $V_k \cap U_{k-1}$.

Recall that the length of J_k' is

$$\langle\langle q_{n_k}\alpha\rangle\rangle \geq \frac{1}{2q_{n_k+1}} \geq \frac{1}{3q_{n_k}a_{n_k+1}}. \tag{7.0.2}$$

Remark 7.2. By Assumption 6 and the inequality above, this gives that the sum of the measures of the S_k is infinite, but $\lim_{k\to\infty} \lambda(S_k) = 0$.

Let λ_{S^1} denote Lebesgue measure on S^1.

Lemma 7.3. *There exists $C > 0, n_0$ such that for $L \neq k$ both at least n_0,*

$$\lambda(S_k \cap S_L) \geq C\lambda(S_k)\lambda(S_L).$$

Proof. Without loss of generality, we assume $L > k$. By definition, the projection of S_k to S^1 is an orbit of length q_{n_k-1} of an interval of size $\langle\langle q_{n_k}\alpha\rangle\rangle$, and the projection of S_L is an orbit of length q_{n_L-1} of an interval of size $\langle\langle q_{n_L}\alpha\rangle\rangle$.

Thus it suffices to show that there is some $C > 0$ such that if I_k is an interval of size $\langle\langle q_{n_k}\alpha\rangle\rangle$, then for any x,

$$\sum_{i=0}^{q_{n_L-1}-1} \chi_{I_k}(R^i(x)) \geq C q_{n_L-1}\lambda_{S^1}(I_k).$$

The desired result then follows by summing as I_k ranges over the intervals of S_k, and integrating x over an interval of size $\langle\langle q_{n_L}\alpha\rangle\rangle$.

We now prove the sufficient condition. By Denjoy-Koksma, the sum is within 2 of $q_{n_L-1}\lambda_{S^1}(I_k)$. Observing

$$\begin{aligned}\frac{q_{n_L-1}\lambda_{S^1}(I_k) - 2}{q_{n_L-1}\lambda_{S^1}(I_k)} &\geq 1 - \frac{2}{q_{n_L-1}\langle\langle q_{n_k}\alpha\rangle\rangle} \\ &\geq 1 - \frac{2q_{n_k+1}}{q_{n_L-1}} \\ &\geq 1 - \frac{6q_{n_k+1}}{q_{n_{k+1}}} \to 1\end{aligned}$$

gives the result. In the last line, we used $L > k$, $q_{n_L-1} > \frac{1}{3}q_{n_L}$ and Assumption 8.

□

Proposition 7.4. *To prove that T is uniquely ergodic, it suffices to show for λ almost every x we have that for each i, (x, i) is in S_k for infinitely many k.*

This will be proved by showing that ergodic measures are absolutely continuous with respect to Lebesgue and points that are in infinitely many S_k cannot be generic for an ergodic measure unless the ergodic measure is Lebesgue.

Lemma 7.5. *If T is not uniquely ergodic then there exist exactly two ergodic probability measures μ, ν with $\mu = \iota(\nu)$, and $\lambda = \mu + \nu$, and both μ and ν are absolutely continuous with respect to Lebesgue.*

Proof. Suppose ν is an ergodic measure that is not Lebesgue. Let $\mu = \iota(\nu)$. Since ι commutes with T, μ must also be ergodic. If $\mu = \nu$, then μ is ι invariant and hence must be Lebesgue, since R is uniquely ergodic.

Since $\mu+\nu$ is ι invariant and R is uniquely ergodic, $\lambda = \mu+\nu$. Hence μ and ν are absolutely continuous with respect to Lebesgue.

If T had a third ergodic measure ν', then $\mu' = \iota(\nu')$ would also be ergodic, and we'd have $\mu' + \nu' = \mu + \nu$. This contradicts uniqueness of ergodic decompositions. □

Lemma 7.6. *U_k is the union of $o(q_{n_{k+1}})$ disjoint intervals. The same result holds with U_k replaced by V_k.*

Proof. By the inductive definition of U_k, it is clear that U_k is the union of at most $O\left(\sum_{i=1}^k q_{n_i}\right)$ disjoint intervals. Thus it suffices to show that $\sum_{i=1}^k q_{n_i}$ is $o(q_{n_{k+1}})$. To do so, note the following crude estimate,

$$q_{n_{k+1}} > q_{n_k+1} > a_{n_k+1}q_{n_k}.$$

Note also that by Assumption 3,

$$\sum_{i=1}^{k} q_{n_i} = q_{n_k} + \sum_{i=1}^{k-1} q_{n_i} < 2q_{n_k}.$$

Thus Assumption 4 gives the result. The argument for V_k is symmetric. □

Lemma 7.7. *Let $A \subset S^1 \times \mathbb{Z}_2$ be any measurable set. For any $\epsilon > 0$,*

$$\lim_{k\to\infty} \lambda\left(\left\{(x,j) \in U_k : \left|\frac{1}{q_{n_k}} \sum_{i=0}^{q_{n_k}-1} \chi_A(T_k^i(x,j)) - \lambda(A\cap U_k)\right| > \epsilon\right\}\right) = 0.$$

The same result is true with U_k replaced by V_k.

Note that T_k appears in the statement, not T.

Proof. It suffices to show this statement for A an interval, since A can be approximated by a union of intervals.

Let A_k be the projection of $A \cap U_k$ to S^1. Since U_k is T_k invariant and projects bijectively to S^1, the lemma is equivalent to

$$\lim_{k\to\infty} \lambda\left(\left\{x \in S^1 : \left|\frac{1}{q_{n_k}} \sum_{i=0}^{q_{n_k}-1} \chi_{A_k}(R^i(x)) - \lambda(A_k)\right| > \epsilon\right\}\right) = 0.$$

Note that since the measure of the symmetric difference of U_k and U_{k-1} goes to zero, because its projection to S^1 is S_k (Remark 7.2), the measure of the symmetric difference of A_k and A_{k-1} must go to zero too.

Thus $|\chi_{A_k} - \chi_{A_{k-1}}|$ has L^1 norm going to zero, so the set of points where the Birkhoff sums for χ_{A_k} and $\chi_{A_{k-1}}$ at time q_{n_k} differ by more than $\epsilon/2$ goes to zero (simply because a function with small L^1 norm can't be big very often: $\lambda(\{x : f(x) > C\}) < \frac{\|f\|_1}{C}$). Hence, it suffices to show the equivalent result with A_k replaced by A_{k-1} and ϵ replaced by $\epsilon/2$.

By the previous lemma, the set $A \cap U_{k-1}$ is a disjoint union of $o(q_{n_k})$ disjoint intervals, and hence $\chi_{A_{k-1}}$ has total variation $o(q_{n_k})$. The statement is now implied by Denjoy-Koksma for the function $\chi_{A_{k-1}}$. □

The proof of Proposition 7.4 will consist of two main steps. Step 1 is to show that if μ is an ergodic measure other than λ it is without loss of generality the weak-* limit of the uniform measure on the U_k. Step 2 is to show that if (x, i) is in S_k for infinitely many k it can not be a generic point for μ.

Proof of Proposition 7.4. Step 1: Assume that λ is not uniquely ergodic. Then Lemma 7.5 gives two ergodic measures μ and ν. Let A be a T invariant set such that $\mu(A) = 1$ and $\nu(A) = 0$. Let $\epsilon > 0$.

We claim that for k sufficiently large

$$\lambda(U_k \cap A) > 1 - 3\epsilon \text{ and } \lambda(V_k \cap A) < 3\epsilon$$

or

$$\lambda(U_k \cap A) < 3\epsilon \text{ and } \lambda(V_k \cap A) > 1 - 3\epsilon.$$

We proceed by contradiction, assuming there exists an infinite sequence of k_i so that

$$3\epsilon < \lambda(U_{k_i} \cap A) < 1 - 3\epsilon.$$

By Lemma 7.7 we have that if k_i is large enough,

$$\lambda\left(\left\{(x, j) \in U_k \cap A : \left|\frac{1}{q_{n_{k_i}}} \sum_{\ell=0}^{q_{n_{k_i}}-1} \chi_A(T^\ell_{k_i}(x, j)) - \lambda(A \cap U_k)\right| < \epsilon\right\}\right)$$
$$> 3\epsilon - \epsilon = 2\epsilon.$$

For all large enough k we have that

$$\lambda(\{(x, j) : T^\ell(x, j) \neq T^\ell_k(x, j) \text{ for some } 0 \leq \ell \leq q_{n_k}\}) < \epsilon.$$

Indeed, this set consists of q_{n_k} intervals of length $\lambda(\cup_{\ell=k+1}^\infty J'_\ell)$, so its measure goes to zero by Inequality (7.0.2) and Assumption 4.

So, for all large enough i, for a set of points in A of measure at least ϵ, we have

$$\sum_{\ell=0}^{q_{n_{k_i}}-1} \chi_A(T^\ell(x,j)) \in (3\epsilon - \epsilon, 1 - 3\epsilon + \epsilon).$$

This contradicts the T invariance of A.

Since $\lim_{k\to\infty} \lambda(U_{k-1}\Delta U_k) = 0$ the property that $\lambda(U_k \cap A)$ is almost 0 or almost 1 is eventually constant (in k). Without loss of generality, suppose $\lambda(U_k \cap A) > 1 - 2\epsilon$ for all large enough k. Since this is true for all $\epsilon > 0$, and since μ projects to Lebesgue, it follows that μ is the weak-* limit of the uniform measure on the U_k.

Step 2: Let $p = (x, i)$ be a point that is μ generic and that is contained in infinitely many S_ℓ. Thus there are infinitely many times ℓ for which the orbit of p is disjoint from U_ℓ for time $q_{n_{\ell+1}-1}$. (As we remarked at the beginning of this section, this is the case when $p \in S_{\ell+1} \cap V_\ell \cap U_{\ell+1}$. Note the "index shift" by one: to get disjointness from U_ℓ for a long time, we use points in $S_{\ell+1}$.)

By the existence of a density point for the set A, there must be some interval $I \in S_1 \times \mathbb{Z}_2$ such that $\mu(I) \geq 0.99\lambda(I)$. Hence for large k, V_k contains at most 0.1 of the λ measure of I. The projection of $V_k \cap I$ to S^1 thus has measure at most $0.1\lambda(I)$, and consists of $o(q_{n_{k+1}})$ intervals (Lemma 7.6). Now assume the orbit segment of length of $q_{n_{k+1}-1}$ of (x, i) is disjoint from U_k. Because $q_{n_{k+1}-1} > \frac{1}{3} q_{n_{k+1}}$ by Denjoy-Koksma, the orbit of x up to time $q_{n_{k+1}-1}$ spends at most $0.2\lambda(I)$ of its time in the projection of I to S_1.

Hence

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \chi_I(T^i p) \leq .2\lambda(I),$$

since $\sum_{i=0}^{N-1} \chi_I(T^i p)$ is bounded by a Birkhoff sum of the characteristic function of the projection of $I \cap V_k$ to S^1.

Thus for p, the liminf of the Birkhoff sums of χ_I is at most $0.2\lambda(I)$, which is a contradiction to Birkhoff's ergodic theorem, and the facts that p is μ generic and $\mu(I) \geq 0.99\lambda(I)$. $\square$

Before completing the proof of Proposition 7.1, we first require a well known and straightforward result, whose proof is included for convenience:

Lemma 7.8 (Quasi-indendent Borel-Cantelli)**.** *Let A_i be measurable subsets of a space with a measure λ of total mass 1. If there exists*

$C > 0$ such that $\lambda(A_i \cap A_j) > C\lambda(A_i)\lambda(A_j)$ and $\sum_{i=1}^{\infty} \lambda(A_i) = \infty$, then $\lambda\left(\bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_i\right) > 0$.

Proof. Let $B_{N,M} = \cup_{i=N}^{M} A_i$. If $\sum_{i=N}^{M} \lambda(A_i) < \frac{1}{2C}$ then for any $j \notin [N, M]$ we have that

$$\lambda(A_j \setminus B_{N,M}) \geq \lambda(A_j) - \sum_{i=M}^{N} C\lambda(A_j)\lambda(A_i) > \frac{1}{2}\lambda(A_j).$$

Because $\sum \lambda(A_i) = \infty$, we have $\lambda(B_{N,\infty}) \geq \frac{1}{4C}$ for all N. Because we are in a finite measure space it follows that $\bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_i$ has positive measure, which proves the claim. $\square$

Proof of Proposition 7.1. Now we complete the proof of the proposition. By Lemmas 7.3 and 7.8, we have $\lambda\left(\bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} S_i\right) > 0$. Thus the set of points in infinitely many S_k has positive Lebesgue measure. Note $(x, i) \in S_k$ depends only on x and being in infinitely many S_i is almost everywhere R invariant. This is because the difference between S_i and $T(S_i)$ is at most two intervals of size $\langle\langle q_{n_i}\alpha\rangle\rangle$, so the difference between $\bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} S_i$ and $T\left(\bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} S_i\right)$ has size bounded by $\sum_{j=i}^{\infty} 2\langle\langle q_{n_j}\alpha\rangle\rangle$ for all i and hence must be measure zero. Indeed, $T\left(\bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} S_i\right) = \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} T(S_i)$ and $\bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} S_i \setminus \left(\bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} T(S_i)\right) \subset \cup_{i=j}^{\infty} S_i \setminus T(S_i)$ for all j.

So, by the ergodicity of R, almost every point is in infinitely many S_i. By Proposition 7.4 this completes the proof. $\square$

8. Mixing of the suspension flow

The purpose of this section is to prove Theorem 1.2, by proving that the flow over T with roof function f, which we denote $F^t : Z \to Z$, is mixing. This section shows that if T is a multi-interval exchange transformation that is ergodic with respect to Lebesgue measure, and f is an integrable function with $f(x) \geq 1$ for all x, and such that f and T satisfy Theorem 6.1, then the suspension flow over T with roof function f is mixing. The fact that estimates like Theorem 6.1 are sufficient for mixing is standard, see for example the sufficient condition for mixing in [**Koc03**, Theorem 2.1]. See also [**Ulc07, Rav17**].

Lemma 8.1. *For all large enough M, there is a set $I(M)$ of disjoint subintervals of $G(M)$ such that each interval has size between $\frac{1}{M\sqrt{\log(M)}}$*

and $\frac{2}{M\sqrt{\log(M)}}$, *and consists entirely of points* p *so that whenever* $N \in [\frac{1}{2}M, 2M]$ *we have*

$$\frac{3}{4}NC \leq \sum_{i=0}^{N-1} f(T^i p) \leq \frac{4}{3}NC,$$

where $C = \int f$, *and such that the union of the intervals of* $I(M)$ *has measure going to 1.*

Proof. For all large enough M, $G(M)$ is the disjoint union of intervals of length at least $\frac{2}{M\sqrt{\log(M)}}$ and it can be divided into disjoint intervals of length between $\frac{1}{M\sqrt{\log(M)}}$ and $\frac{2}{M\sqrt{\log(M)}}$. Let $I(M)$ be the set of these disjoint intervals and consider $I \in I(M)$.

By Theorem 6.1 (5), if $p \in G(M)$, then $|f'(T^i p)| = o(M\sqrt{\log(M)})$ for all $0 \leq i \leq 2M + 5\sqrt{\log(M)}$. Since the poles of f' are of the form $\frac{1}{x}$, it follows that if $p \in G(M)$, then $T^i p$ avoids at $\frac{1}{M\sqrt{\log(M)}}$ neighborhood of the singularities of f for $0 \leq i \leq 2M + 5\sqrt{M}$. Let B denote the $\frac{1}{M\sqrt{\log(M)}}$ neighborhood of the singularities of f. Because the poles of f are logarithmic,

$$\mathrm{Var}(f|_{B^c}) = O(\log(M \log(M)^{\frac{1}{2}})) = O(\log(M)).$$

By an argument as in Corollary 5.4, for any $N \in [\frac{1}{2}M, 2M]$ we have

$$\sum_{i=0}^{N} f(T^i p) - N\int f = O(\log(N)^2).$$

The Mean Value (which we may apply because by our construction of $G(M)$, $\sum_{i=0}^{2M-1} f \circ T^i$ is continuous on I) and Theorem 6.1 part (3) give that if p satisfies the above bound, then all points in I (which by definition of $I(M)$ have distance at most $\frac{2}{M\sqrt{\log(M)}}$ to p) satisfy the weaker bound in the lemma statement for M large enough. $\square$

We now wish to study the curve $F^{CM}(I)$, where I is a fixed interval in $I(M)$. To this end, if $p \in I$, let $N(p, M)$ denote the unique integer such that

$$\sum_{i=0}^{N(p,M)-1} f(T^i p) \leq CM < \sum_{i=0}^{N(p,M)} f(T^i p).$$

So $N(p, M)$ is the number of times $F^t(p)$ returns to the base up to time $t = CM$, counting $t = 0$. In the remainder of the section we will write $N(p)$ in place of $N(p, M)$ because M will be clear from context.

Lemma 8.2. *For M large enough, for every interval $I \in I(M)$ and any point $(p, y) \in Z$ with $p \in I$,*

$$F^{CM}(I)$$

over the interval $I \ni p$ lies within $o(1)$ of a segment of a vertical trajectory of length between $\sqrt{\log(M)}/4$ and $5\sqrt{\log(M)}$. Moreover, this is true in a parametrized sense: the graph is piecewise C^1 with slope within $o(M\sqrt{\log(M)})$ of a constant $s_{M,I}$ at every point.

Remark 8.3. In Lemma 8.2, the variable M plays the role of the "main" variable in the $o(\cdot)$ notation, which was denoted N in Remark 5.3.

Proof. In this proof we assume $p \in I$. The previous lemma gives that $\sum_{i=0}^{\frac{1}{2}M+1} f(T^i p) < CM$ and $\sum_{i=0}^{2M-1} f(T^i p) > CM$ and so we may assume $\frac{1}{2}M \leq N(p) \leq 2M$.

Because F^t is the flow over the piecewise isometry T with roof function f, which is differentiable on $T^i(I)$ for all $0 \leq i \leq 2M$, the vertical length of $F^{CM}(I)$ is

$$\int_I \sum_{i=0}^{N(p)-1} f'(T^i p)dp,$$

which, using Theorem 6.1 (3), that I is an interval in $G(M)$, of length at most $\frac{2}{M\sqrt{\log(M)}}$, and $N(p) \leq 2M$, has size at most

$$\begin{aligned} \int_I \left(\left| \sum_{i=0}^{2M} f'(T^i p)dp \right| \right) &\leq 2\frac{2M\log(2M) + o(2M\log(M))}{M\sqrt{\log(M)}} \\ &\leq 5\sqrt{\log(M)} \end{aligned}$$

for M sufficiently large.

Similarly, because $N(p) \geq \frac{1}{2}M$, if M is sufficiently large

$$\left| \int_I \sum_{i=0}^{N(p)-1} f'(T^i p)dx \right| \geq \frac{\frac{1}{2}M\log(\frac{1}{2}M) + o(M\log(M))}{M\sqrt{\log(M)}} > \frac{\sqrt{\log(M)}}{4}.$$

Now, since $F^{CM}(I)$ has length at most $5\sqrt{\log(M)}$, and the roof function f is always at least 1, it follows that $F^{CM}(I)$ can hit the base at most $5\sqrt{\log(M)}$ times. That is, if p and p' are any two points in I, then

$$|N(p) - N(p')| \leq 5\sqrt{\log(M)}. \tag{8.0.1}$$

Now let $N = N(p)$ for any point $p \in I$, and let $p' \in I$ be another point. Define

$$s = \sum_{i=0}^{N(p)-1} f'(T^i p).$$

By Theorem 6.1 part (3) we have $s = \pm N \log(N) + o(N \log N)$. By Theorem 6.1 part (4) and the Mean Value Theorem, (which we may apply because by our construction of $G(M)$, $\sum_{i=0}^{j} f \circ T^i$ is continuous and differentiable on I for any $0 \leq j \leq 2M$) we have that

$$\left| s - \sum_{i=0}^{N(p)-1} f'(T^i p') \right| \leq \frac{2o(\log(N)^{\frac{1}{3}} N^2)}{M\sqrt{\log(M)}} \leq o((\log(N))^{\frac{1}{3}-\frac{1}{2}} N)$$

and by (8.0.1) and Theorem 6.1 part (5) we have

$$\left| \sum_{i=0}^{N(p')-1} f'(T^i p') - \sum_{i=0}^{N(p)-1} f'(T^i p') \right| = o(N\sqrt{\log(N)}).$$

We get that the given graph is piecewise C^1 with slope within $o(M\sqrt{\log(M)})$ of s at every point. This completes the proof (with $s_{M,I}$ this s). $\square$

Let Λ denote 2 dimensional Lebesgue probability measure on Z, the suspension of T by f. Let R be a rectangle, by which we mean a rectangle contained strictly under the graph of f. Say that a point $(p, y) \in Z$ is (L, ϵ, R) good if every vertical line V through (p, y) of length at least L has

$$|m_V(V \cap R) - \Lambda(R)| \leq \epsilon,$$

where m_V is the 1 dimensional Lebesgue probability measure on V.

Lemma 8.4. *For fixed ϵ and R, the set of points that are (L, ϵ, R) good has measure going to 1 as $L \to \infty$.*

Proof. This follows from ergodicity of the flow. $\square$

Lemma 8.5. *Let R and R' be two rectangles. Then*

$$\lim_{t\to\infty} \Lambda(R \cap F^t(R')) \to \Lambda(R)\Lambda(R').$$

Proof. Let $\epsilon > 0$ be arbitrarily small and in particular much smaller than the height and width of R. Let R_s be the set of points in R that have distance at least $\epsilon/50$ to the boundary of R, and let R_b be an $\epsilon/50$ neighbourhood of R. ("s" and "b" stand for "smaller" and "bigger".)

Let H denote the y-coordinate of the top edge of R'. For notational simplicity, we assume the bottom edge of R' has y-coordinate 0. (The general case is analogous.) Define D_L to be the set of points (p, y) such that all $F^h(p, y)$ with $0 \leq h \leq H$ are both $(L, \epsilon/100, R_s)$ and $(L, \epsilon/100, R_b)$ good. Pick L large enough such that D_L has Λ-measure at least $1 - \epsilon/100$.

Let M_0 be large enough so that

- for all $M \geq M_0$ we have $\sqrt{\log(M)}/4 > \max(L, H)$,
- the $o(1)$ error in Lemma 8.2 is less than $\epsilon/100$,
- the union of the intervals in $I(M)$ has λ-measure at least $1 - \epsilon/100$,
- the error in Theorem 6.1 parts (3) and (5) are less than $\epsilon/1000$,
- $2H/(M_0\sqrt{\log M_0}) < \epsilon/1000$

Claim. If $t \geq 10CM_0$, $M = \lfloor \frac{t}{C} \rfloor$, $I \in I(M)$ and there exists $p \in I$ such that $F^t p \in D_L$, then for any $0 \leq h \leq H$ we have

$$\left| \frac{1}{\lambda(I)} \int_I \chi_R(F^{t+h}p)dp - \Lambda(R) \right| < \varepsilon/10.$$

Proof of claim. By Lemma 8.2, there exists a vertical trajectory, V_0

through $F^{CM}p$, that approximates $F^{CM}(I)$, as a parametrized curve, to within $\epsilon/100$. Now, $0 \leq t + h - CM \leq h + C$ and so by using Theorem 6.1 (5) and the Mean Value Theorem we have that the vertical line $V = F^{h+t-CM}(V_0)$ approximates $F^{CM+h+(t-CM)}(I)$, to within $\epsilon/50$, and observe

$$m_V(R_s \cap V) \leq \frac{1}{\lambda(I)} \int_I \chi_R(F^{t+h}p)dp \leq m_V(R_b \cap V).$$

Since $F^t p \in D_L$, the claim has been established.

Now consider $(p, y) \in R'$. Let $t > 10CM_0$ and let $M = \lfloor \frac{t}{C} \rfloor$. Set

$$P_t = F^{-t}(D_L) \cap \{(p, y) \in Z : p \in I \text{ for some } I \in I(M)\}.$$

Let $I_{R'}(M) = \{I \in I(M) : I \subset R'\}$ and observe that this occupies most of the bottom horizontal segment of R' (for large M) and what is missing is contained in $G(M)^c$. Similarly, for $0 \leq h \leq H$, $\cup_{I \in I_{R'}(M)} F^h(I)$ occupies most of the horizontal coordinate h slice of R'.

$$\Lambda(R \cap F^t R') = \sum_{I \in I_{R'}(M)} \int_0^H \int_I \chi_R(F^t F^h p)dpdh + O\left(H\lambda\left(G(M)^c\right)\right).$$

Furthermore,

$$\Lambda(R\cap F^tR') = \int_0^H \sum_{I\in I_{R'}(M):F^hI\cap P_t\neq\emptyset} \int_I \chi_R(F^tF^hp)dpdh + \\ O\left(H\lambda\left(G(M)^c\right)\right) + O(\Lambda(D_L^c)).$$

Applying the claim, we have that for each summand, the inside integral is $\lambda(I)\Lambda(R)+O(\epsilon\lambda(I))$. This establishes the lemma.

□

Since it is enough to prove mixing for rectangles, this proves Theorem 1.2.

9. A mixing flow on a surface with non-degenerated fixed points

As we remarked after the statement of Theorem 1.2, the T and f constructed thus far do not satisfy the technical conditions to correspond to a smooth flow. In this final section, we modify them to prove Theorem 1.1.

Define $x_\infty = \alpha - \sum_{i=1}^{\infty}\langle\langle q_{n_i-1}\alpha\rangle\rangle$ and $x_k = \alpha - \sum_{i=1}^{k}\langle\langle q_{n_i-1}\alpha\rangle\rangle$. We now define another IET

$$\hat{T}: S^1\times\mathbb{Z}_2\times\mathbb{Z}_2 \to S^1\times\mathbb{Z}_2\times\mathbb{Z}_2$$

by

$$\hat{T}(x,i,j) = (R(x), i+\chi_J(R(x)), j+\chi_{J+x_\infty}(R(x))).$$

Recall the J is defined at the start of Section 4. The key observation about $\hat{T}$ is that its restriction to the first and second coordinates is T, and its restriction to the first and third coordinates is a "translate" of T by x_∞. Our choice of x_∞ was motivated by the proof of Theorem 9.5 below, which gives that $\hat{T}$ is uniquely ergodic. Observe that Assumption 2 implies that the intervals J and $x_\infty + J$ are disjoint.

Define the roof function

$$\begin{aligned}\hat{f}(x,i,j) &= 1+|\log(d(x,R^{-1}0))|+|\log(d(x,R^{-1}|J|)|\\ &+ |\log(d(x,R^{-1}x_\infty))|+|\log(d(x,R^{-1}|J|+x_\infty))|\\ &+ \chi_{\{1\}}(j)|\log(d(x,R^{-1}0))|.\end{aligned}$$

The first two lines put logarithmic singularities of equal weight (coefficient) over all discontinuities of $\hat{T}$, and the third line introduces additional weight to the singularity over one pair of the discontinuities.

The purpose of this section is to show

Theorem 9.1. *The flow over $\hat{T}$ with roof function $\hat{f}$ is mixing.*

At the end of this section, we will use this to conclude Theorem 1.1.

Non-minimal approximants. Define

$$\hat{T}_\ell(x,i,j) = (R(x), \chi_{[0,\sum_{k=1}^{\ell} 2\langle\langle q_{n_k}\alpha\rangle\rangle)}(R(x)), \chi_{[0,\sum_{k=1}^{\ell} 2\langle\langle q_{n_k}\alpha\rangle\rangle)+x_\infty}(R(x))).$$

Define the following subsets of $S^1 \times \mathbb{Z}_2 \times \mathbb{Z}_2$,

$$U_k^1 = \{(x,i,j) : (x,i) \in U_k\}, \qquad V_k^1 = \{(x,i,j) : (x,i) \in V_k\},$$
$$U_k^2 = \{(x,i,j) : (x-x_\infty,j) \in U_k\}, \quad V_k^2 = \{(x,i,j) : (x-x_\infty,j) \in V_k\}.$$

Recall that U_k and V_k are defined after Lemma 4.1. A corollary of Lemma 4.3 is the following.

Corollary 9.2. *The above four sets are all $\hat{T}_k$ invariant.*

Note that all of the sets

$$U_k^1 \cap U_k^2, \quad U_k^1 \cap V_k^2, \quad V_k^1 \cap U_k^2, \quad V_k^1 \cap V_k^2$$

are invariant and project bijectively to S^1. Since R is minimal, we see that each of these four sets is minimal. Since $S^1 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ is the disjoint union of these four sets, the minimal components of $\hat{T}_k$ are exactly these four sets.

Before we continue, we note some estimates.

Lemma 9.3. *The following hold for all large enough ℓ.*

(1) $\langle\langle q_{n_\ell -1}\alpha\rangle\rangle > 4\langle\langle q_{n_\ell}\alpha\rangle\rangle$.
(2) $d(x_\ell, x_\infty) < \frac{1}{2}\langle\langle q_{n_\ell}\alpha\rangle\rangle$.
(3) *If* $N = 1 + \sum_{i=0}^{\ell} q_{n_i -1}$, *then* $\frac{1}{3} q_{n_\ell} \le N < q_{n_\ell}$.

Proof. The first claim follows from Theorem 2.2, since

$$\langle\langle q_{n_\ell -1}\alpha\rangle\rangle > \frac{1}{2q_{n_\ell}} > \frac{4}{q_{n_\ell+1}} > 4\langle\langle q_{n_\ell}\alpha\rangle\rangle,$$

where we have used that, given our assumptions, $q_{n_\ell+1} > 8q_{n_\ell}$.

The second claim follows from the fact that

$$\begin{aligned} d(x_\ell, x_\infty) &= \sum_{k=\ell+1}^{\infty} 2\,|\langle\langle q_{n_k-1}\alpha\rangle\rangle| \\ &\le \sum_{k=\ell+1}^{\infty} \frac{2}{q_{n_k}} \le \frac{4}{q_{n_{\ell+1}}} \le \frac{1}{2}\langle\langle q_{n_\ell}\alpha\rangle\rangle \end{aligned}$$

and the exponential growth of the q_j (compare to the start of proof of Theorem 6.1).

The upper bound in the third claim follows by noting that Assumption 3 from Section 3 gives that $q_{n_{\ell-1}} \ge \sum_{i=1}^{\ell-1} q_{n_i-1}$ and so $N <$

$q_{n_{\ell-1}} + q_{n_\ell - 1}$. The lower bound is obtained by noting $N \geq q_{n_\ell - 1}$ and $q_{n_\ell} \leq (a_{n_\ell} + 1) q_{n_\ell - 1}$ and then using Assumption 9. ☐

Unique ergodicity. Before proving unique ergodicity we need the following lemma. Recall $J'_\ell = \left[\sum_{k=1}^{\ell-1} 2\langle\langle q_{n_k}\alpha\rangle\rangle, \sum_{k=1}^{\ell-1} 2\langle\langle q_{n_k}\alpha\rangle\rangle + \langle\langle q_{n_\ell}\alpha\rangle\rangle\right)$.

Lemma 9.4. *For all large enough ℓ the intervals*

$$R^i(x_\infty + J'_\ell), \quad i = 0, \dots, q_{n_\ell}$$

are disjoint from each other and $[0, \frac{1}{2}\langle\langle q_{n_\ell}\alpha\rangle\rangle)$ and $[1 - \langle\langle q_{n_\ell}\alpha\rangle\rangle, 1)$.

For large enough ℓ, $\cup_{i=0}^{q_{n_\ell - 1}} R^i(J'_\ell)$ is disjoint from $x_\infty + J'_\ell$ and $x_\infty + J''_\ell$.

Proof. Compare to the proof of Lemma 4.2. We will use Lemma 9.3 several times.

Set $N = 1 + \sum_{i=1}^{\ell} q_{n_i - 1}$ and $N_0 = 2\sum_{k=1}^{\ell-1} q_{n_k}$. Note that the intervals

$$R^i(x_\ell + J'_\ell), \quad i = 0, \dots, q_{n_\ell}$$

and $[0, \langle\langle q_{n_\ell}\alpha\rangle\rangle)$ and $[1 - \langle\langle q_{n_\ell}\alpha\rangle\rangle, 1)$ are contained in an orbit of length $N_0 + N + 2q_{n_\ell}$ of an interval of size $\langle\langle q_{n_\ell}\alpha\rangle\rangle$.

By Assumption 3, $N_0 < q_{n_\ell}$, so $N_0 + N + 2q_{n_\ell} < 4q_{n_\ell}$. By Assumption 4, we have $4q_{n_\ell} < a_{n_\ell+1} q_{n_\ell} < q_{n_\ell+1}$, so in particular we conclude that $N_0 + N + 2q_{n_\ell} < q_{n_\ell+1}$.

Hence, by the separation property, the above intervals are disjoint. Since $x_\ell - \frac{1}{2}\langle\langle q_{n_\ell}\alpha\rangle\rangle < x_\infty < x_\ell$, the intervals $R^i(x_\infty + J'_\ell)$ are disjoint from each other and $[0, \frac{1}{2}\langle\langle q_{n_\ell}\alpha\rangle\rangle)$ and $[1 - \langle\langle q_{n_\ell}\alpha\rangle\rangle, 1)$.

We now prove the final claim. Note that by Theorem 2.2 (1), $x_\ell + J'_\ell = R^N(J'_\ell)$. The intervals in $\{R^i(J'_\ell)\}_{i=0}^{q_{n_\ell}-1}$ are $\langle\langle q_{n_\ell - 1}\alpha\rangle\rangle - \langle\langle q_{n_\ell}\alpha\rangle\rangle$ separated because their left endpoints are $\langle\langle q_{n_\ell - 1}\alpha\rangle\rangle$ separated by Theorem 2.2 (4) and the intervals have length $\langle\langle q_{n_\ell}\alpha\rangle\rangle$). So, $R^N(J'_\ell)$ is the unique element of this orbit segment within $\langle\langle q_{n_\ell - 1}\alpha\rangle\rangle - \langle\langle q_{n_\ell}\alpha\rangle\rangle$ of $x_\ell + J'_\ell$. Combining these two facts and the fact that $d(x_\infty, x_\ell) < \langle\langle q_{n_\ell - 1}\alpha\rangle\rangle - \langle\langle q_{n_\ell}\alpha\rangle\rangle$, N is the unique $i \in \{0, ..., q_{n_\ell} - 1\}$ such that $R^i(J'_\ell) \cap (x_\infty + J'_\ell) \neq \emptyset$. Since $N > q_{n_\ell - 1}$ we have treated $x_\infty + J'_\ell$. Because $J''_\ell = \langle\langle q_{n_\ell}\alpha\rangle\rangle + J'_\ell$, we have that $x_\infty + J''_\ell$ is a subset of a $3\langle\langle q_{n_\ell}\alpha\rangle\rangle$ neighborhood of $R^N(J''_\ell)$. Since any orbit of length $q_{n_\ell - 1}$ is $\langle\langle q_{n_\ell - 1}\alpha\rangle\rangle$ separated and $\langle\langle q_{n_\ell - 1}\alpha\rangle\rangle - 3\langle\langle q_{n_\ell}\alpha\rangle\rangle > 0$ we have that $R^i(J'_\ell) \cap (x_\infty + J''_\ell) = \emptyset$ for all i satisfying $0 < |N - i| < q_{n_\ell}$, completing the proof of the lemma. ☐

Theorem 9.5. *$\hat{T}$ is uniquely ergodic.*

Proof. This is similar to the proof that T is uniquely ergodic. We will outline the additional considerations that are required.

Since T is uniquely ergodic, $\hat{T}$ has at most two ergodic components, each of which project to Lebesgue under the projection $(x, i, j) \mapsto (x, i)$ to the first two coordinates, and also under the projection $(x, i, j) \mapsto (x, j)$ to the first and third coordinates. If there are two ergodic measures, they must be exchanged under the involutions $(x, i, j) \mapsto (x, i+1, j)$ and $(x, i, j) \mapsto (x, i, j+1)$, and hence invariant under their product $(x, i+1, j+1)$. In fact one of the two measures must be the weak-* limit of Lebesgue on $E_\ell = (U_\ell^1 \cap V_\ell^2) \cup (V_\ell^1 \cap U_\ell^2)$, and the other must be the limit of $F_\ell = (U_\ell^1 \cap U_\ell^2) \cup (V_\ell^1 \cap V_\ell^2)$.

The proof now follows as for T, with S_ℓ replaced by

$$\cup_{i=0}^{q_{n_\ell - 2}-1} R^i(J'_\ell) \times \mathbb{Z}_2 \times \mathbb{Z}_2.$$

Note that, as in the beginning of Section 7, we have $\frac{1}{3} q_{n_\ell - 1} \leq q_{n_\ell - 2} \leq \frac{1}{2} q_{n_\ell - 1}$. Thus by the previous lemma, this set consists of points p such that the orbit of p of length $q_{n_\ell - 2}$ is entirely in $E_\ell \cup F_{\ell-1}$ or entirely in $F_\ell \cup E_{\ell-1}$. So we may repeat the last half of the proof of Proposition 7.4. $\square$

Birkhoff sums. We now outline the relevant changes to the proof of Theorem 6.1.

To set up context, the appropriate version of Lemma 4.2 is Lemma 9.4. With this, Theorem 4.4 can be modified to

Theorem 9.6. *For each integer $\ell \geq 1$, the following properties hold.*

(1) U_ℓ^2 contains

$$[0, \frac{1}{2}\langle\langle q_{n_\ell}\alpha\rangle\rangle) \times \mathbb{Z}_2 \times \{0\} \quad \textit{and} \quad [1 - \langle\langle q_{n_\ell}\alpha\rangle\rangle, 1) \times \mathbb{Z}_2 \times \{1\}.$$

(2)

$$\pi_{S_1}(U_\ell^2 \setminus U_{\ell-1}^2) = \cup_{i=0}^{q_{n_\ell}-1} R^i(J'_\ell + x_\infty).$$

We next observe that Lemma 5.6 can be modified to be

Lemma 9.7. *Set $g(x) = 1/x$. For any k with $q_{n_k} \leq N$ and any $x \in [0, 1)$ we have,*

$$\sum_{i=0}^{N-1} g(R^i x)\chi_{[\frac{1}{2}\langle\langle q_{n_{k-1}}\alpha\rangle\rangle, 1)}(R^i x) = o(N\log(N))$$

and

$$\left|\sum_{i=0}^{N-1} g'(R^i x)\chi_{[\frac{1}{2}\langle\langle q_{n_{k-1}}\alpha\rangle\rangle, 1)}(R^i x)\right| = O(N^2).$$

With Lemma 9.4 in mind, we can modify Lemma 5.8 to be

Lemma 9.8. *Set $g(x) = 1/x$. For any ℓ such that $q_{n_\ell} < N$, let S be an orbit of length q_{n_ℓ} of an interval of length $\langle\langle q_{n_\ell}\alpha\rangle\rangle$. Assume S is disjoint from $[0, \frac{1}{2}\langle\langle q_{n_\ell}\alpha\rangle\rangle)$, and x is any point disjoint from $\cup_{i=0}^{\sqrt{a_{n_\ell+1}}q_{n_\ell}} R^{-i}(S)$. Then*
$$\sum_{i=0}^{N-1} g(R^i x)\chi_S(R^i x) = o(N\log(N))$$
and
$$\left|\sum_{i=0}^{N-1} g'(R^i x)\chi_S(R^i x)\right| = o(N^2\log(N)^{\frac{1}{3}}).$$

We now outline the straightforward modifications to Theorem 6.1. Let $\hat{Q}_M$ be the points within distance $\frac{1}{M(\log(M))^{\frac{1}{12}}}$ of the singularities of $\hat{f}$. Let
$$\begin{aligned}\hat{B}_{S^1}(M) &= \bigcup_{i=0}^{2M+5\sqrt{\log(M)}} R^{-i}(\hat{Q}_M) \\ &\cup \bigcup_{i=0}^{2M+5\sqrt{\log(M)}} R^{-i}(x_\infty + \cup_{i=\ell+1}^{\infty} J_\ell) \\ &\cup \bigcup_{i=-q_{n_\ell}}^{\sqrt{a_{n_\ell+1}}q_{n_\ell}} R^{-i}(x_\infty + J'_\ell).\end{aligned}$$

The changes to the proof of Claims 1 and 2 are obvious changes to the measure estimates.

Claim 3 requires the most substantive changes. We assume $p \in V_\ell^2$ (the case of $p \in U_\ell^2$ is analogous). As before, because of cancellations and our choice of $\hat{B}_{S^1}$ we may restrict our attention to $\chi_{\{1\}}(j)(g(-x) - g(x))$. Recall $g(x) = \frac{1}{\langle\langle x\rangle\rangle}$. By Lemma 9.7 we may restrict to $[0, \frac{1}{2}\langle\langle q_{n_{\ell-1}}\alpha\rangle\rangle) \times \mathbb{Z}_2 \times \{1\}$ which by Theorem 9.6 is entirely contained in $V_{\ell-1}^2$. Then in place of Lemma 5.8 we invoke Lemma 9.8.

Claim 4 is analogous to Claim 3.

The change to Claim 5 is straightforward.

Claim 6 is straightforward.

This proves the analogue of Theorem 6.1. Since we have already proven unique ergodicity of $\hat{T}$ (Theorem 9.5), mixing for the suspension flow now follows as in the previous section for T and f.

A flow on a surface. The flat surface pictured in Figure 9.1 has 8 cone points, each with angle 4π. This flow is C^∞ away from the cone

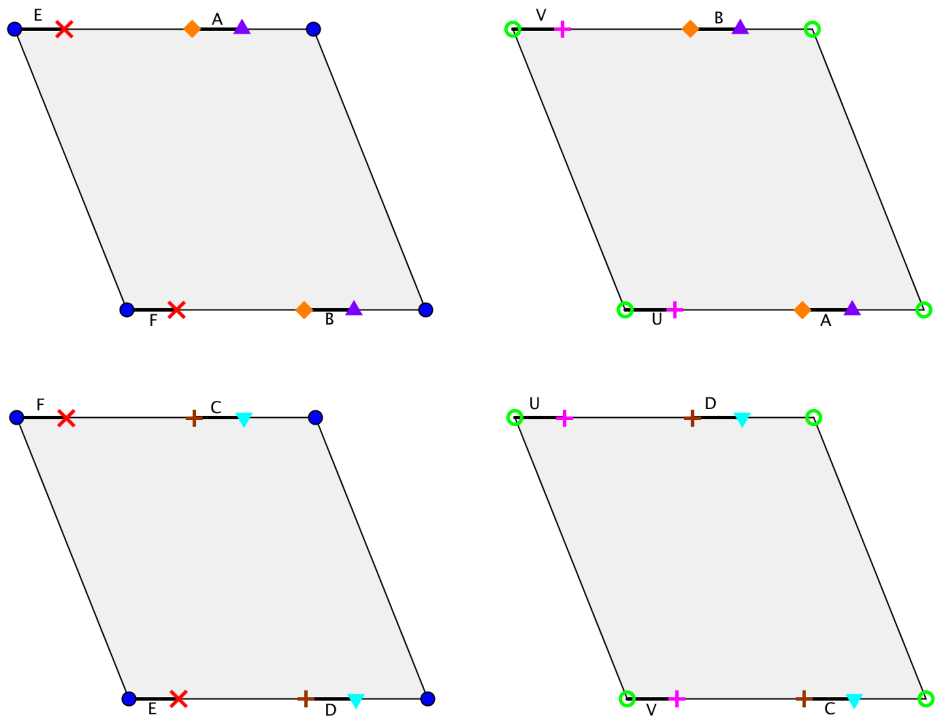


FIGURE 9.1. A union of slit tori. Unmarked opposite sides are identified. The height and base of each parallelogram is one, and the shear is by α. All of the labelled intervals have length $|J|$, and in each parallelogram the second interval is equal to the first translated by x_∞.

points. By appropriately slowing down the flow near these fixed points–a standard procedure explained in detail in **[CF11**, Section 7]–one can obtain a C^∞ flow on this surface that has non-degenerated fixed points at the 8 distinguished points points, and such that the first return time function h satisfies that $h - \hat{f}$, $h' - \hat{f}'$, $h'' - \hat{f}''$ are bounded. Because Birkhoff sums of a bounded function over orbit segments of length N are $O(N)$, all estimates in this paper hold with $\hat{f}$ replaced with h. Hence the above arguments show that the C^∞ flow we have produced is mixing. Note that the fixed point represented by an open circle contributes twice the shearing as the others.

The first return map of the flow (either the straight line flow on the translation surface or the C^∞ flow that it is a reparametrization of it) to the union of the four intervals at the bottom of the four parallelograms is $\hat{T}$.

A saddle connection is a trajectory of the flow that connects singularities of the flat surface. By the definition of the skew product, points in $S^1 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ with the same S^1 coordinate cannot have that their forward orbits intersect. So to show that there does not exist a saddle connection, it suffices to show that the forward $\hat{T}$ orbits each element of $\{x_\infty - \alpha, x_\infty + |J| - \alpha, -\alpha, 1 - \alpha + |J|\} \times \mathbb{Z}_2 \times \mathbb{Z}_2$ are infinite and distinct. This is straightforward to check from the construction. This verifies that the flow does not have saddle connections and completes the proof of Theorem 1.1.

Remark 9.9. In the introduction, we remarked that there are non-minimal smooth flows with finitely many non-degenerate fixed points on surfaces of genus 3 and 4 such that one minimal component sees only one side of a fixed point. To obtain such a flow, start with two flat tori (translation surfaces of genus 1), and cut a slit in each, parallel and of the same length. Each torus now has a boundary circle consisting of two intervals. We may obtain a genus 3 surface as follows: Glue one boundary interval from one of the tori to a boundary interval on the other torus. Glue the remaining two intervals to each other using an IET with permutation on $\{1, 2, 3, 4, 5\}$ that fixes 1, 3 and 5 and swaps 2 and 4. This gives a genus three translation surface with 4 singularities of cone angle 4π. A genus 4 translation surface can be obtained using a second slit disjoint from and parallel to the first, using the typical gluing without non-trivial IETs. One can obtain the desired smooth flows as a time change from straight line flow in the direction of the slits.

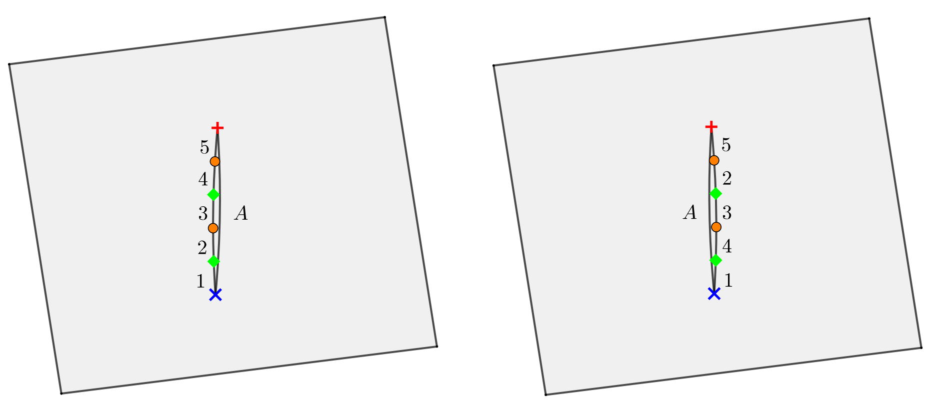


FIGURE 9.2. The genus 3 surface constructed in Remark 9.9.

References

[Arn91] V. I. Arnol′d, *Topological and ergodic properties of closed 1-forms with incommensurable periods*, Funktsional. Anal. i Prilozhen. **25** (1991), no. 2, 1–12, 96.

[Blo72] A. A. Blohin, *Smooth ergodic flows on surfaces*, Trudy Moskov. Mat. Obšč. **27** (1972), 113–128.

[CF11] Jean-Pierre Conze and Krzysztof Frączek, *Cocycles over interval exchange transformations and multivalued Hamiltonian flows*, Adv. Math. **226** (2011), no. 5, 4373–4428.

[CHM11] Yitwah Cheung, Pascal Hubert, and Howard Masur, *Dichotomy for the Hausdorff dimension of the set of nonergodic directions*, Invent. Math. **183** (2011), no. 2, 337–383.

[Fay06] Bassam Fayad, *Smooth mixing flows with purely singular spectra*, Duke Math. J. **132** (2006), no. 2, 371–391.

[FFK] Bassam Fayad, Giovanni Forni, and Adam Kanigowski, *Lebesgue spectrum for area preserving flows on the two torus*, preprint, arXiv 1609.03757 (2016).

[FK] Bassam Fayad and Adam Kanigowski, *Multiple mixing for a class of conservative surface flows*, preprint, arXiv 1409.1054 (2014).

[FL03] K. Frączek and M. Lemańczyk, *On symmetric logarithm and some old examples in smooth ergodic theory*, Fund. Math. **180** (2003), no. 3, 241–255.

[FL04] ———, *A class of special flows over irrational rotations which is disjoint from mixing flows*, Ergodic Theory Dynam. Systems **24** (2004), no. 4, 1083–1095.

[FL05] Krzysztof Frączek and Mariusz Lemańczyk, *On disjointness properties of some smooth flows*, Fund. Math. **185** (2005), no. 2, 117–142.

[FL06] K. Frączek and M. Lemańczyk, *On mild mixing of special flows over irrational rotations under piecewise smooth functions*, Ergodic Theory Dynam. Systems **26** (2006), no. 3, 719–738.

[FL09] ———, *Smooth singular flows in dimension 2 with the minimal self-joining property*, Monatsh. Math. **156** (2009), no. 1, 11–45.

[For02] Giovanni Forni, *Deviation of ergodic averages for area-preserving flows on surfaces of higher genus*, Ann. of Math. (2) **155** (2002), no. 1, 1–103.

[Her79] Michael-Robert Herman, *Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations*, Inst. Hautes Études Sci. Publ. Math. (1979), no. 49, 5–233.

[Kat67] A. B. Katok, *Spectral properties of dynamical systems with an integral invariant on the torus*, Funkcional. Anal. i Priložen. **1** (1967), no. 4, 46–56.

[Kat73] ———, *Invariant measures of flows on orientable surfaces*, Dokl. Akad. Nauk SSSR **211** (1973), 775–778.

[Kat80] Anatole Katok, *Interval exchange transformations and some special flows are not mixing*, Israel J. Math. **35** (1980), no. 4, 301–310.

[Khi64] A. Ya. Khinchin, *Continued fractions*, The University of Chicago Press, Chicago, Ill.-London, 1964.

[KKP] Adam Kanigowski and Joanna Kuaga-Przymus, *Ratners property and mild mixing for smooth flows on surfaces*, preprint, arXiv 1409.2987 (2014).

[KKPU] Adam Kanigowski, Joanna Kuaga-Przymus, and Corinna Ulcigrai, *Multiple mixing and parabolic divergence in smooth area-preserving flows on higher genus surfaces*, preprint, arXiv 1606.09189 (2016).

[Koč72] A. V. Kočergin, *The absence of mixing in special flows over a rotation of the circle and in flows on a two-dimensional torus*, Dokl. Akad. Nauk SSSR **205** (1972), 515–518.

[Koč75] ______, *Mixing in special flows over a rearrangement of segments and in smooth flows on surfaces*, Mat. Sb. (N.S.) **96(138)** (1975), 471–502, 504.

[Koč76] ______, *Nondegenerate saddles, and the absence of mixing*, Mat. Zametki **19** (1976), no. 3, 453–468.

[Koc02] A. V. Kochergin, *A mixing special flow over a rotation of the circle with an almost Lipschitz function*, Mat. Sb. **193** (2002), no. 3, 51–78.

[Koc03] ______, *Nondegenerate fixed points and mixing in flows on a two-dimensional torus*, Mat. Sb. **194** (2003), no. 8, 83–112.

[Koc04a] A. Kochergin, *Well-approximable angles and mixing for flows on $\mathbb{T}^2$ with nonsingular fixed points*, Electron. Res. Announc. Amer. Math. Soc. **10** (2004), 113–121 (electronic).

[Koc04b] A. V. Kochergin, *Hölder time change and the mixing rate in a flow on a two-dimensional torus*, Tr. Mat. Inst. Steklova **244** (2004), no. Din. Sist. i Smezhnye Vopr. Geom., 216–248.

[Koc04c] ______, *Nondegenerate fixed points and mixing in flows on a two-dimensional torus. II*, Mat. Sb. **195** (2004), no. 3, 15–46.

[Koc04d] ______, *Some generalizations of theorems on mixing flows with nondegenerate saddles on a two-dimensional torus*, Mat. Sb. **195** (2004), no. 9, 19–36.

[Koc07a] ______, *Nondegenerate saddles, and absence of mixing. II*, Mat. Zametki **81** (2007), no. 1, 145–148.

[Koc07b] Andrey Kochergin, *Causes of stretching of Birkhoff sums and mixing in flows on surfaces*, Dynamics, ergodic theory, and geometry, Math. Sci. Res. Inst. Publ., vol. 54, Cambridge Univ. Press, Cambridge, 2007, pp. 129–144.

[Kol53] A. N. Kolmogorov, *On dynamical systems with an integral invariant on the torus*, Doklady Akad. Nauk SSSR (N.S.) **93** (1953), 763–766.

[KSS75] A. B. Katok, Ja. G. Sinaĭ, and A. M. Stepin, *The theory of dynamical systems and general transformation groups with invariant measure*, Mathematical analysis, Vol. 13 (Russian), Akad. Nauk SSSR Vsesojuz. Inst. Naučn. i Tehn. Informacii, Moscow, 1975, pp. 129–262. (errata insert).

[KT06] Anatole Katok and Jean-Paul Thouvenot, *Spectral properties and combinatorial constructions in ergodic theory*, Handbook of dynamical systems. Vol. 1B, Elsevier B. V., Amsterdam, 2006, pp. 649–743.

[Lem00] M. Lemańczyk, *Sur l'absence de mélange pour des flots spéciaux au-dessus d'une rotation irrationnelle*, Colloq. Math. **84/85** (2000), no. part 1, 29–41, Dedicated to the memory of Anzelm Iwanik.

[Lim] Yuri Lima, *CF1: uniform distribution and Denjoy-Koksma's inequality*, http://matheuscmss.wordpress.com/2012/02/16/cf1-uniform-distribution-and-denjoy-koksmas-inequality/.

[Mas82] Howard Masur, *Interval exchange transformations and measured foliations*, Ann. of Math. (2) **115** (1982), no. 1, 169–200.

[Nov82] S. P. Novikov, *The Hamiltonian formalism and a multivalued analogue of Morse theory*, Uspekhi Mat. Nauk **37** (1982), no. 5(227), 3–49, 248.

[Ost22] Alexander Ostrowski, *Bemerkungen zur Theorie der Diophantischen Approximationen*, Abh. Math. Sem. Univ. Hamburg **1** (1922), no. 1, 77–98.

[Rav17] Davide Ravotti, *Quantitative Mixing for Locally Hamiltonian Flows with Saddle Loops on Compact Surfaces*, Ann. Henri Poincaré **18** (2017), no. 12, 3815–3861.

[Sat75] E. A. Sataev, *The number of invariant measures for flows on orientable surfaces*, Izv. Akad. Nauk SSSR Ser. Mat. **39** (1975), no. 4, 860–878.

[Sch09a] Dmitri Scheglov, *Absence of mixing for smooth flows on genus two surfaces*, J. Mod. Dyn. **3** (2009), no. 1, 13–34.

[Sch09b] ______, *Absence of mixing for smooth flows on genus two surfaces*, ProQuest LLC, Ann Arbor, MI, 2009, Thesis (Ph.D.)–The Pennsylvania State University.

[SK92] Ya. G. Sinaĭ and K. M. Khanin, *Mixing of some classes of special flows over rotations of the circle*, Funktsional. Anal. i Prilozhen. **26** (1992), no. 3, 1–21.

[Tre14] Rodrigo Treviño, *On the ergodicity of flat surfaces of finite area*, Geom. Funct. Anal. **24** (2014), no. 1, 360–386.

[Ulc07] Corinna Ulcigrai, *Mixing of asymmetric logarithmic suspension flows over interval exchange transformations*, Ergodic Theory Dynam. Systems **27** (2007), no. 3, 991–1035.

[Ulc09] ______, *Weak mixing for logarithmic flows over interval exchange transformations*, J. Mod. Dyn. **3** (2009), no. 1, 35–49.

[Ulc11] ______, *Absence of mixing in area-preserving flows on surfaces*, Ann. of Math. (2) **173** (2011), no. 3, 1743–1778.

[Vee69] William A. Veech, *Strict ergodicity in zero dimensional dynamical systems and the Kronecker-Weyl theorem* mod 2, Trans. Amer. Math. Soc. **140** (1969), 1–33.

[Vee82] ______, *Gauss measures for transformations on the space of interval exchange maps*, Ann. of Math. (2) **115** (1982), no. 1, 201–242.

[vN32] J. von Neumann, *Zur Operatorenmethode in der klassischen Mechanik*, Ann. of Math. (2) **33** (1932), no. 3, 587–642.

Department of Mathematics
University of Utah
155 S 1400 E, Room 233
Salt Lake City, UT 84112
E-mail address: chaika@math.utah.edu

Department of Mathematics
Stanford University
Palo Alto, CA 94305
E-mail address: amwright@stanford.edu