Ergod. Th. & Dynam. Sys. (2007), **27**, 991–1035
doi:10.1017/S0143385706000836

Mixing of asymmetric logarithmic suspension flows over interval exchange transformations

CORINNA ULCIGRAI

Department of Mathematics, Princeton University, Fine Hall, Washington Road, Princeton, NJ 08544-1000, USA
(e-mail: ulcigrai@math.princeton.edu)

(*Received* 13 *April* 2006 *and accepted in revised form* 3 *September* 2006)

Abstract. We consider suspension flows built over interval exchange transformations with the help of roof functions having an asymmetric logarithmic singularity. We prove that such flows are strongly mixing for a full measure set of interval exchange transformations.

1. *Introduction*

1.1. *Motivation and main references.* Hamiltonian systems with multi-valued Hamiltonian functions on two-dimensional tori give rise to area-preserving flows which can be decomposed into a finite number of components filled with periodic trajectories and one ergodic component (see [**1**]). The flow on this ergodic component is isomorphic to a suspension flow built over a rotation of the circle with the help of a roof function which has asymmetric logarithmic singularities (see also §1.2 for precise definitions).

The question about the mixing of such flows, first mentioned in the same paper [**1**], was answered by Sinai and Khanin in [**23**], where it was proved that, under a generic Diophantine condition on the rotation angle, suspension flows with asymmetric singularities over a rotation are strongly mixing (see also [**11**]). The Diophantine condition of [**23**] was weakened by Kochergin in a series of works [**16–19**].

Mixing in these flows is produced by different deceleration rates near the singular points. Neighboring points on a Poincaré transversal have different return times and this causes a phenomenon sometimes called stretching of the Birkhoff sums (the idea of how this stretching leads to mixing is explained in §1.3.3). A similar stretching of Birkhoff sums was also used by Fayad in [**5**] to construct mixing reparametrization of flows on $\mathbb{T}^3$.

Mixing does not arise in suspension flows over rotations in the case of bounded variation roof functions [**12**]. The presence of a symmetric logarithmic singularity is also not enough, as was shown by Kochergin in [**14**]. Lemańczyk [**20**] proved the absence of mixing if the Fourier coefficients of the roof function are of order $O(1/|n|)$ and showed with

Frączek that these flows are disjoint in the sense of Furstenberg from all mixing flows [**6**]. This condition is essentially sharp, see [**15**].

Consider, instead of $\mathbb{T}^2$, a compact orientable surface M_g of higher genus ($g \geq 2$). A closed Morse 1-form ω generates a Hamiltonian flow determined by the multi-valued Hamiltonian H locally defined by $dH = \omega$. The corresponding area-preserving flow on M_g can be decomposed into components filled by periodic orbits and components on which the flow is metrically isomorphic to a suspension flow over an interval exchange transformation (IET) (see, e.g., [**27**]). IETs are piecewise orientation-preserving isometries of an interval which appear naturally as first return maps of such flows on a transversal, as rotations do in the case of $\mathbb{T}^2$.

It was proved by Katok in [**8**] that suspension flows over IETs under roof functions of bounded variation do not mix and (see [**7**]) are disjoint from mixing flows. On the other hand, Kochergin (see [**13**]) proves mixing for a class of roof functions over IETs which includes power-like singularities, which arise when the fixed points on the corresponding surface flow are degenerate. The presence of non-degenerate fixed points give rise to logarithmic singularities. Frączek and Lemańczyk prove in [**7**] that in the case of symmetric logarithmic singularities and typical IETs of two or three intervals the suspension flows are also disjoint from mixing flows.

In this paper we consider suspension flows over IETs of an arbitrary number of intervals with roof functions having a single asymmetric logarithmic singularity. We prove that for typical IETs such flows are strongly mixing. The case of several asymmetric singularities will be treated in another paper.

As mentioned above, the main mechanism of mixing is the stretching of Birkhoff sums. The proof of stretching in our case uses the Rauzy–Veech renormalization algorithm for IETs (see §2). The condition on the IET which guarantees mixing is typical in view of a recent result in [**2**].

1.2. *Definitions and main result*

1.2.1. *IETs.* Let $I^{(0)} = [0, 1)$ and let $T : I^{(0)} \to I^{(0)}$ be an IET of d subintervals, i.e. a piecewise orientation-preserving isometry of $I^{(0)}$ defined in the following way. Assign a permutation $\pi \in \mathcal{S}_d$ and a partition of $I^{(0)}$ into d subintervals, $I_1^{(0)}, I_2^{(0)}, \ldots, I_d^{(0)}$, defined by a lengths vector $\underline{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_d), \lambda_i > 0, \sum_{i=1}^{d} \lambda_i = 1$, such that $\lambda_i = |I_i^{(0)}|$. Then T permutes the subintervals according to π so that under the action the transformation $I_i^{(0)}$ becomes the $\pi(i)$th interval, i.e. the order of the subintervals after applying T is $I_{\pi^{-1}(1)}^{(0)}, I_{\pi^{-1}(2)}^{(0)}, \ldots, I_{\pi^{-1}(d)}^{(0)}$. More precisely

$$I_j^{(0)} \doteqdot \left[\sum_{i=1}^{j-1} \lambda_i, \sum_{i=1}^{j} \lambda_i\right), \quad j = 1, \ldots, d,$$

$$T(x) = x - \sum_{i=1}^{j-1} \lambda_i + \sum_{i=1}^{j-1} \lambda_{\pi^{-1}i} \quad \text{for } x \in I_j^{(0)},\ j = 1, \ldots, d.$$

We shall often use the notation $T = (\lambda, \pi)$.

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

1.2.2. *Suspension flows.* Let $f \in L^1(I^{(0)}, dx)$ be a strictly positive function $f \geq m_f > 0$ and assume $\int_{I^{(0)}} f(x)\, dx = 1$. Further assumptions on f will be formulated in §1.2.3. The *phase space* X_f of the suspension flow is defined as

$$X_f \doteqdot \{(x, y) \mid x \in I^{(0)},\ 0 \leq y < f(x)\}$$

and can be depicted as the set of points below the graph of the roof function f. We introduce the normalized measure μ which is the restriction to X_f of the Lebesgue measure $dx\, dy$.

The *suspension flow built over T with the help of the roof function f* is a one-parameter group $\{\varphi_t\}_{t\in\mathbb{R}}$ of μ-measure-preserving transformations of X_f whose action is generated by the following two relations:

$$\begin{cases} \varphi_t(x, y) = (x, y + t) & \text{if } 0 \leq y + t < f(x), \\ \varphi_{f(x)}(x, 0) = (Tx, 0). \end{cases} \tag{1}$$

Under the action of the flow, a point of $(x, y) \in X_f$ moves with unit velocity along the vertical line up to the point $(x, f(x))$ and then jumps instantly to the point $(T(x), 0)$, according to the base transformation. Afterwards it continues its motion along the vertical line until the next jump and so on (see, e.g., [**4**]).

Let $S_0(f, T)(x) \doteqdot 0$. We will denote by

$$S_r(f, T)(x) = S_r(f)(x) \doteqdot \sum_{i=0}^{r-1} f(T^i(x)), \quad x \in I^{(0)},\ r \in \mathbb{N}^+,$$

the rth non-renormalized *Birkhoff sum* of f along the trajectory of x under T. The dependence on T is omitted when there is no ambiguity.

Let $t > 0$. Given $x \in I^{(0)}$ denote by $r(x, t)$ the integer uniquely defined by

$$r(x, t) \doteqdot \max\{r \in \mathbb{N} \mid S_r(f)(x) \leq t\}, \tag{2}$$

which describes the number of *discrete iterations* of the IET that the point $(x, 0)$ undergoes before time t. According to this notation the flow φ_t defined by (1) acts as

$$\varphi_t(x, 0) = (T^{r(x,t)}(x), t - S_{r(x,t)}(f)(x)). \tag{3}$$

For $t < 0$, the action of the flow is defined as the inverse map.

1.2.3. *Single asymmetric logarithmic singularity.* Assume that $f \in \mathcal{C}^2((0, 1))$ and there exist two positive constants $C^+ > 0$, $C^- > 0$, such that

$$\lim_{x\to 0^+} \frac{f''(x)}{1/x^2} = C^+, \quad \lim_{x\to 1^-} \frac{f''(x)}{(1/(1-x))^2} = C^-. \tag{4}$$

It is easy to see that (4) implies that

$$\lim_{x\to 0^+} \frac{f(x)}{|\log x|} = C^+, \quad \lim_{x\to 1^-} \frac{f(x)}{|\log (1-x)|} = C^-. \tag{5}$$

Hence we say in this case that f has a *logarithmic singularity* at the origin. The singularity is said to be *asymmetric* if $C^+ \neq C^-$.

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

1.2.4. *Mixing.* Recall that a flow $\{\varphi_t\}_{t\in\mathbb{R}}$ preserving the measure μ is said to be *mixing* if, for each pair of measurable sets A, B, one has

$$\lim_{t\to\infty} \mu(\varphi_t(A) \cap B) = \mu(A)\mu(B). \tag{6}$$

1.2.5. *Main result.* The main result of this paper is the following.

THEOREM 1.1. *The suspension flow $\{\varphi_t\}_{t\in\mathbb{R}}$ built over a typical IET T with the help of a roof function f having a single asymmetric logarithmic singularity at the origin is mixing.*

The notion of *typical* IET is understood from the measure theoretical point of view. More precisely, for every irreducible π, Theorem 1.1 holds for almost every length vector $\underline{\lambda} \in \Delta_{d-1}$ with respect to the Lebesgue measure on the simplex Δ_{d-1}.

1.3. *A criterion for mixing*

1.3.1. *Partial partitions and rectangles.* By a *partial partition* η of $I^{(0)}$ into intervals we mean a collection of disjoint intervals $I = [a, b)$. We do not require that the union of these intervals is the whole $I^{(0)}$. All of the partitions in this paper will be partial partitions into a finite number of intervals. Denote by Leb the Lebesgue measure on the Borel subsets of $I^{(0)}$. By using the notation $\text{Leb}(\eta)$ we mean the total measure of a partition η, i.e. $\text{Leb}(\eta) \doteqdot \sum_{I\in\eta} \text{Leb}(I)$. The *mesh* of the partition η is given by $\text{mesh}(\eta) \doteqdot \sup_{I\in\eta} \text{Leb}(I)$. We will consider one-parameter families of partial partitions $\eta(t)$, $t \in \mathbb{R}$.

We call a *rectangle* of base $b(R) \subset I^{(0)}$ and height $h = h(R) < m_f$ the set R of points (x, y) such that $0 \le y \le h$ and $x \in b(R)$. Rectangles and their shifts $\varphi_t(R)$ generate the Borel σ-algebra of (X_f, μ).

1.3.2. *Mixing criterion.* In order to demonstrate mixing it is sufficient to verify the following criterion, similar to that used in [**5, 13**].

LEMMA 1.1. (Mixing criterion) *If, given any rectangle R, any $\epsilon > 0$ and any $\delta > 0$, one can find $t_0 > 0$ such that for each $t \ge t_0$ one can define a partial partition $\eta(t)$ of $I^{(0)}$ into intervals such that*

$$\text{Leb}(\eta(t)) > 1 - \delta, \quad \text{mesh}(\eta(t)) \le \delta \tag{7}$$

and for each $I \in \eta(t)$

$$\text{Leb}(I \cap \varphi_{-t}(R)) \ge (1 - \epsilon)\,\text{Leb}(I)\mu(R), \tag{8}$$

then the flow $\{\varphi_t\}_{t\in\mathbb{R}}$ is mixing.

Proof. Mixing means that for any two measurable sets A and B and any $\epsilon > 0$, for sufficiently large positive t,

$$\mu(A \cap \varphi_{-t}(B)) > (1 - \epsilon)\mu(A)\mu(B), \tag{9}$$

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

and by applying (9) to A^C and B one obtains

$$\mu(A \cap \varphi_{-t}(B)) < (1+\epsilon)\mu(B)\mu(A) + \epsilon\mu(B)$$

and therefore (6). For $t < 0$, it is sufficient to exchange the roles of A and B and use μ-invariance of φ_t. Moreover, it is sufficient to verify (9) for A and B rectangles, since any measurable set can be approximated by a finite union of rectangles and their shifts under the flow.

Let $b(A)$ be the base of a rectangle A. For each $\delta > 0$ and $t \geq t_0$, there exists a finite number of intervals $I_k^{(t)} \in \eta(t)$, $k = 0, \ldots, K(t)$, such that $\text{Leb}\big(b(A)\Delta \bigcup_{k=0}^{K(t)} I_k^{(t)}\big) \leq 3\delta$. Here Δ denotes the symmetric difference of sets. To see this, consider all intervals of $\eta(t)$ which intersect $b(A)$ and use (7). Let

$$A' \doteqdot \bigcup_{0 \leq y \leq h(A)} \bigg(\bigcup_{k_y=0}^{K(t+y)} I_{k_y}^{(t+y)} \times \{y\}\bigg).$$

Choosing $\delta \leq (\epsilon/3)\,\text{Leb}(b(A))\mu(B)$, by the Fubini theorem, $\mu(A\Delta A') \leq \epsilon\mu(A)\mu(B)$. Noting the inequality $y \leq h(A) < m_f$, we have, for each slice of A',

$$\bigg(\bigcup_{k_y=0}^{K(t+y)} I_{k_y}^{(t+y)} \times \{y\}\bigg) \cap \varphi_{-t}(B) = \varphi_y\bigg(\bigg(\bigcup_{k_y=0}^{K(t+y)} I_{k_y}^{(t+y)} \times \{0\}\bigg) \cap \varphi_{-t-y}(B)\bigg).$$

Moreover, φ_y preserves Leb on each slice and therefore one can assume that the hypothesis (8) in which we set $B = R$ holds for all slices. Thus, combining these estimates and again applying the Fubini theorem, we obtain, for $t \geq t_0$,

$$\begin{aligned}
\mu(A \cap \varphi_{-t}(B)) &\geq \mu(A' \cap \varphi_{-t}(B)) - 3\delta h(A) \\
&\geq \int_0^{h(A)} (1-\epsilon)\,\text{Leb}\bigg(\bigcup_{k_y=0}^{K(t+y)} I_{k_y}^{(t+y)}\bigg)\mu(B)\,dy - \epsilon\mu(A)\mu(B) \\
&\geq (1-3\epsilon)\mu(A)\mu(B),
\end{aligned}$$

hence proving the lemma. □

1.3.3. *Intuitive explanation of the mixing mechanism.* Consider a sufficiently small segment $I = [a, b] \subset I^{(0)}$ and let us consider its image under the flow, $\varphi_t(I)$, for very large t. We claim that $\varphi_t(I)$ will consist of many almost vertical curves, as shown in Figure 1(a).

Assume as a simple example that f has only a one-sided logarithmic singularity at the origin and is monotonically decreasing, as in Figure 1(b). Notice first that until $t < m_f, \varphi_t(I)$ is still a horizontal segment, while as $t = f(x_0)$ for some $x_0 \in I$, $\varphi_t(I)$ splits into two curves: one is a still a horizontal segment, while the other curve will project over $T([x_0, b])$ and be a translation of the graph of $-f|_{[x_0,b]}$, as can be seen by (3); see Figure 1(b). More generally, from (3), each of the curves in which $\varphi_t(I)$ will split is a graph of a translation of the Birkhoff sum $S_r(f)$ restricted over a small interval of the form $T^r([x_i, x_{i+1}))$, where $[x_i, x_{i+1}) \subset I$. Noticing that $f' < 0$ and the integral of f'

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

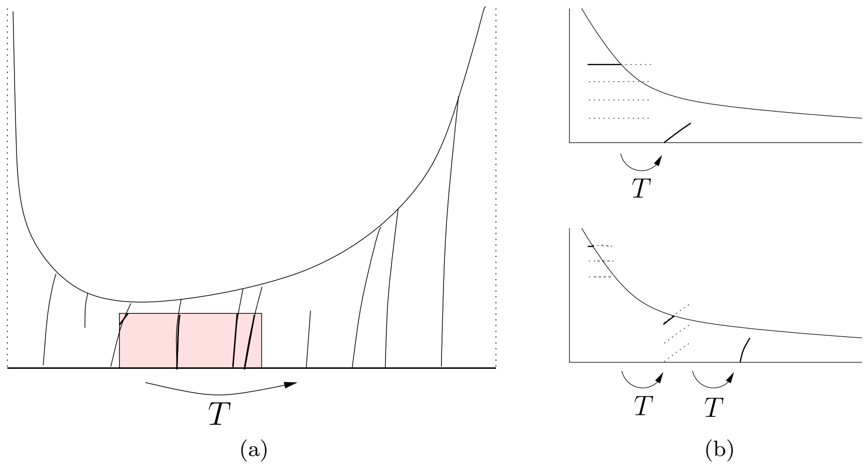

FIGURE 1. The evolution of $I \subset I^{(0)}$ under φ_t: (a) $\varphi_t(I)$ for large t; (b) a simple model of the initial evolution.

is divergent, one can prove in this simple model that the slopes of these curves, which are given by $-S_r(f')$, are growing to infinity, i.e. they are becoming almost vertical. Hence the increasingly large delay between different points causes $\varphi_t(I)$ to split into many curves, which are distributed over the orbit of $T^n(I)$. Using the unique ergodicity of T on the base and the fact that each strip can be approximated by a straight line, one can show that the fraction of $x \in I$ such that $\varphi_t(x) \cap R \neq \emptyset$ is proportional to $\mu(R)$.

When the singularity is asymmetric the same phenomenon occurs and one can show that delays accumulated from visits on one side are stronger than the delays accumulated from the opposite side, causing $S_r(f')(x)$ to diverge as if it were in the presence of a one-sided singularity for most of the points.

1.3.4. *Outline of the proof of Theorem 1.1.* In order to prove mixing for the suspension flow, we use the criterion in Lemma 1.1. Given a rectangle R and $\epsilon, \delta > 0$, our goal is to construct, for any sufficiently large t, a partial partition $\eta(t)$ of $I^{(0)}$ into intervals which satisfy (7) and (8). Each of the intervals of these partitions under the flow exhibits the behavior as explained in the previous paragraph. The construction of the partition is carried out in several steps, explained in §4. In order to obtain the final estimate (8), in §4.2, the key step is to obtain a good estimate of the rate of growth of the first two derivatives of $S_r(f)$. Such estimates, presented in §3, are based on some property of the renormalization cocycle for IETs first introduced by Rauzy [**22**], and developed by Veech [**24**] and Zorich [**26**]. The definition and some properties of this cocycle are recalled in §2.

2. *Renormalization algorithms for IETs*

Rauzy [**22**], Veech [**24**] and Zorich [**26**] developed a renormalization algorithm for IETs which is a multi-dimensional generalization of the continued fraction algorithm.

In the following, let $T = (\underline{\lambda}, \pi)$ be an IET. We assume that π is *irreducible*, i.e. if the subset $\{1, 2, \ldots, i\}$ is π-invariant, then $i = d$, since this is a necessary condition for minimality. We also assume that T satisfies the infinite distinct orbit condition (IDOC) introduced by Keane in [**9**]. We say that T has the IDOC if, denoting by $\beta_0 = 0$ and $\beta_j \doteqdot \sum_{i=1}^{j} \lambda_i$ for $j = 1, \ldots, d$ the discontinuities of T, the orbits $\mathcal{O}(\beta_j) \doteqdot \{T^n(\beta_j) \mid n \in \mathbb{N}\}$, $1 \le j \le d-1$, are infinite and disjoint, i.e. $\mathcal{O}(\beta_j) \cap \mathcal{O}(\beta_i) = \emptyset$ for any $i \neq j$. As it was shown by Keane in [**9**], the IDOC implies minimality.

2.1. *Rauzy–Veech and Zorich algorithms and cocycles.* Starting with $T = T^{(0)}$, the Rauzy–Veech algorithm produces a sequence of IET $T^{(r)}$ which are induced maps of T onto a sequence of nested subintervals $I^{(r)} \subset I$. It is easy to see that in general the induced first return map of T on a subinterval $I' \subset I$ is again an IET, of at most $d+2$ intervals. One Rauzy step is defined so that $T^{(1)}$ is an exchange of exactly the same number d of subintervals.

2.1.1. *One step of the Rauzy–Veech algorithm.* At the first step, compare the lengths of I_d and $I_{\pi^{-1}(d)}$, i.e. of the last subintervals before and after the transformation. It follows from the IDOC that $\lambda_d \neq \lambda_{\pi^{-1}d}$. Hence there can be two cases.

(a) $\lambda_d < \lambda_{\pi^{-1}d}$. In this case we consider the new interval $I^{(1)} \doteqdot [0, 1-\lambda_d)$. Define $T^{(1)}$ to be the induced map, i.e. the first return map of $T^{(0)}$ onto $I^{(1)}$. It is important that it is again an IET of the same number d of exchanged intervals.

(b) $\lambda_d > \lambda_{\pi^{-1}d}$. In this case we consider the new interval $I^{(1)} \doteqdot [0, 1-\lambda_{\pi^{-1}d})$ and, as before, define $T^{(1)}$ to be the induced map on $I^{(1)}$. Also in this case $T^{(1)}$ is again an IET of d intervals.

Since $T^{(1)}$ is again an exchange of d intervals, we can write $T^{(1)} = (\lambda^{(1)}, \pi^{(1)})$, defining in this way a new lengths vector and a new permutation. One can explicitly write the expressions for two combinatorial operators a and b on $\mathcal{S}_d$, where $\mathcal{S}_d$ is the space of permutations of d elements, such that $\pi^{(1)} = a\pi$ or $b\pi$, respectively. Explicitly,

$$a\pi(j) = \begin{cases} \pi(j), & j \le \pi^{-1}(d), \\ \pi(d), & j = \pi^{-1}(d) + 1, \\ \pi(j-1), & \text{otherwise}, \end{cases}$$

$$b\pi(j) = \begin{cases} \pi(j), & j \le \pi(d), \\ \pi(j)+1, & \pi(d) < \pi(j) < d, \\ \pi(d)+1, & \pi(j) = d. \end{cases}$$

We introduce the following matrices to describe the new lengths. We denote by Id the identity $d \times d$ matrix and by $E_{i,j}$ the matrix whose only non-zero entry is $(E_{i,j})_{ij} = 1$. Let us introduce the auxiliary permutation $\tau_s \in \mathcal{S}_d$, $\tau_s = (12 \ldots s s+2 \ldots d s+1)$ if $1 \le s < d-1$ and $\tau_{d-1} = \text{Id}$, which rotates cyclically all elements after the sth element. Denote by $P(\tau_s)$ the matrix associated to the permutation, i.e. $P(\tau_s)_{ij} = \delta_{i\tau_s(i)}$.

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

The two *Rauzy–Veech elementary matrices* associated to $(\underline{\lambda}, \pi)$ are defined by

$$\begin{cases} A(\pi, a) = (\mathrm{Id} + E_{\pi^{-1}(d),d}) \cdot P(\tau_{\pi^{-1}(d)}), \\ A(\pi, b) = \mathrm{Id} + E_{d,\pi^{-1}(d)}. \end{cases} \tag{10}$$

The induced IET $T^{(1)}$ is then given by

$$(\underline{\lambda}^{(1)}, \pi^{(1)}) \doteqdot \begin{cases} (A^{-1}(\pi, a) \cdot \underline{\lambda}, a(\pi)), & \lambda_d < \lambda_{\pi^{-1}(d)}, \\ (A^{-1}(\pi, b) \cdot \underline{\lambda}, b(\pi)), & \lambda_d > \lambda_{\pi^{-1}(d)}. \end{cases} \tag{11}$$

Note that both $A(\pi, a)$ and $A(\pi, b)$ belong to $SL(d, \mathbb{Z})$ and have non-negative entries.

Define inductively $T^{(r)} = (\underline{\lambda}^{(r)}, \pi^{(r)})$ to be the induced map of $T^{(r-1)}$ on $I^{(r)}$. It can be seen that the IDOC assures that the algorithm is well defined at each step, i.e. that $\lambda_d^{(r-1)} \neq \lambda^{(r-1)}_{(\pi^{(r-1)})^{-1}(d)}$ for any $r \in \mathbb{N}$.

2.1.2. *Renormalized Rauzy–Veech map.* The *Rauzy class* of π, denoted by $\mathcal{R}(\pi) \subset \mathcal{S}_d$, is the set of all permutations obtained iterating the operators a and b starting from π. Using the norm $|\underline{\lambda}| = \sum_{i=1}^{d} \lambda_i$, assume that the initial lengths belong to the simplex Δ_{d-1} of vectors $\underline{\lambda} \in \mathbb{R}^d_+$ such that $|\underline{\lambda}| = 1$. Let us denote by $\Delta(\mathcal{R}) = \Delta_{d-1} \times \mathcal{R}(\pi)$ the space of IETs on the unit interval corresponding to a given Rauzy class $\mathcal{R}$.

Consider the map on $\Delta(\mathcal{R})$ which associates to T the induced IET after one step of the algorithm including the following renormalization:

$$\mathcal{R}((\underline{\lambda}^{(0)}, \pi^{(0)})) \doteqdot \left(\frac{\underline{\lambda}^{(1)}}{|\underline{\lambda}^{(1)}|}, \pi^{(1)} \right).$$

Let us call it the *Rauzy–Veech map*. Veech proved that $\mathcal{R}$ admits an invariant measure $\mu_{\mathcal{V}}$, absolutely continuous with respect to the Lebesgue measure, which is infinite. The main result proved by Veech in [**25**] is that the map $\mathcal{R}$ is conservative. As a consequence, Veech proves that, given any irreducible $\pi \in \mathcal{S}_d$, for almost every $\underline{\lambda} \in \Delta_{d-1}$, the IET $T = (\underline{\lambda}, \pi)$ is uniquely ergodic.

2.1.3. *Zorich acceleration.* Take an IET T and consider its Rauzy–Veech orbit $\{\mathcal{R}^n T\}_{n \in \mathbb{N}}$. In a typical situation one can find an integer $z_o = z_o(T) > 0$ so that $T, \mathcal{R}T, \ldots, \mathcal{R}^{z_0 - 1}(T)$ all correspond to the same case (a) or (b) while $\mathcal{R}^{z_o}(T)$ corresponds to the other. Grouping together these z_0 steps of Rauzy induction, we obtain a new transformation $\mathcal{Z}$ on the space of IET, where the letter $\mathcal{Z}$ is chosen in honor of A. Zorich who introduced this map in [**26**]. Zorich showed in [**26**] that $\mathcal{Z}$ has an absolutely continuous *finite* invariant measure. We will denote the *Zorich invariant measure* by $\mu_{\mathcal{Z}}$.

2.1.4. *Rauzy–Veech lengths cocycle.* As was explained above, to each T one can associate an elementary matrix $A(T)$ in $SL(d, \mathbb{Z})$ defining $A(T) \doteqdot A(\pi, a)$ or $A(T) \doteqdot A(\pi, b)$, respectively. Let $A_r = A_r(T) \doteqdot A(\mathcal{R}^r T)$. Then for each r we can associate to T the product

$$A^{(r)} \doteqdot A_0 \cdot \cdots \cdot A_{r-1}.$$

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

We can easily see that the map $A^{-1} : \Delta(\mathscr{R}) \to SL(d, \mathbb{Z})$ is a cocycle over $\mathcal{R}$, which we call the *Rauzy–Veech lengths cocycle*. Iterating the lengths relation in (11) we obtain the formula for the lengths vector of $T^{(r)}$:

$$\underline{\lambda}^{(r)} = (A^{(r)})^{-1}\underline{\lambda}. \tag{12}$$

Let us also introduce the following notation which is useful when considering more general products of Rauzy–Veech cocycle matrices from m to n, $m < n$:

$$A^{(m,n)} \doteqdot A_m \cdot A_{m+1} \cdot \cdots \cdot A_{n-2} \cdot A_{n-1}.$$

2.1.5. *Hilbert metric and projective contractions.* Consider on the simplex $\Delta_{d-1} \subset \mathbb{R}^d_+$ the *Hilbert distance* d_{H}, defined as follows:

$$d_{\mathrm{H}}(\lambda, \lambda') \doteqdot \log \frac{\max_{i=1,\dots,d} \lambda_i/\lambda'_i}{\min_{i=1,\dots,d} \lambda_i/\lambda'_i}. \tag{13}$$

We denote the diameter with respect to d_{H} of a projective subset $\Lambda \subset \Delta_{d-1}$ by

$$\mathrm{diam}_{\mathrm{H}}(\Lambda) \doteqdot \sup_{\lambda,\lambda' \in \Lambda} d_{\mathrm{H}}(\lambda, \lambda'). \tag{14}$$

Note that if its closure $\overline{\Lambda} \subset \Delta_{d-1}$, then $\mathrm{diam}_{\mathrm{H}}(\Lambda)$ is finite.

Let us write $A \geq 0$ if A has non-negative entries and $A > 0$ if A has strictly positive entries. Recall that to each $A \in SL(d, \mathbb{Z})$, $A \geq 0$, one can associate a projective transformation $\widetilde{A} : \Delta_{d-1} \to \Delta_{d-1}$ given by

$$\widetilde{A}\lambda = \frac{A\lambda}{|A\lambda|}.$$

When $A \geq 0$, $d_{\mathrm{H}}(\widetilde{A}\lambda, \widetilde{A}\lambda') \leq d_{\mathrm{H}}(\lambda, \lambda')$. Furthermore, if $A > 0$, then we obtain a contraction. More precisely, $A > 0$ is equivalent to the closure $\overline{\widetilde{A}(\Delta_{d-1})}$ being contained in Δ_{d-1}; hence defining

$$D(A) \doteqdot \mathrm{diam}_{\mathrm{H}}(\overline{\widetilde{A}(\Delta_{d-1})}), \tag{15}$$

we have $D(A) < \infty$. Then

$$d_{\mathrm{H}}(\widetilde{A}\lambda, \widetilde{A}\lambda') \leq (1 - e^{-D(A)})\, d_{\mathrm{H}}(\lambda, \lambda'). \tag{16}$$

2.1.6. *Paths on Rauzy classes.* Rauzy classes can be visualized in terms of directed labeled graphs, the *Rauzy graphs*. Vertices are in one-to-one correspondence with permutations of $\mathscr{R}(\pi)$; arrows connect permutations obtained from one vertex to the other by applying a or b and are labeled according to the type a or b, respectively. Each vertex is the starting point and the ending point of exactly two arrows, one of each type. We will denote by $\gamma(\pi', a)$ ($\gamma(\pi', b)$) the arrow of type a (type b) coming out from the vertex π'.

A *path* $\underline{\gamma} = (\gamma_1, \dots, \gamma_r)$ is a sequence of compatible arrows on the Rauzy graph, i.e. the starting vertex of γ_{i+1} is the ending vertex of γ_i, $i = 1, \dots, r-1$. Given a path $\underline{\gamma}$, we can associate to it a matrix

$$A(\underline{\gamma}) \doteqdot A(\gamma_1) \cdot \cdots \cdot A(\gamma_r),$$

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

where $A(\gamma_i) = A(\pi_i, a)$ if $\gamma_i = \gamma(\pi_i, a)$ and $A(\pi_i, b)$ if $\gamma_i = \gamma(\pi_i, b)$. Associate to $\underline{\gamma}$ also the subsimplex

$$\Delta(\underline{\gamma}) \doteqdot \{\widetilde{A(\underline{\gamma})\underline{\lambda}} \mid \underline{\lambda} \in \Delta_{d-1}\} \subset \Delta_{d-1}. \tag{17}$$

Using induction, one can easily verify the following.

Remark 2.1. If $T = (\underline{\lambda}, \pi)$ and $\underline{\lambda} \in \Delta(\underline{\gamma})$ where $\underline{\gamma} = (\gamma_1, \ldots, \gamma_r)$ is a path starting at π, $\gamma_i = \gamma(\pi_i, c_i)$, $c_i \in \{a, b\}$, the sequence of types and permutations obtained in the first r steps of the Rauzy–Veech induction is determined by $\underline{\gamma}$, i.e. $A_i(T) = A(\gamma_i)$ and $\pi^{(i)} = \pi_i$.

2.1.7. *The natural extension.* The natural extension $\hat{\mathcal{R}}$ of the map $\mathcal{R}$ was introduced by Veech [**25**] and admits a geometric interpretation in terms of the space of zippered rectangles. We use the simpler choice of coordinates for zippered rectangles, adopted by [**2, 3, 21**].

Consider the following polyhedral cones $\Theta_\pi \subset \mathbb{R}^d$, where $\pi \in \mathfrak{R}$:

$$\Theta_\pi \doteqdot \left\{ \underline{\tau} = (\tau_1, \ldots, \tau_d) \in \mathbb{R}^d \;\middle|\; \sum_{i=1}^{k} \tau_i > 0, \sum_{i=1}^{k} \tau_{\pi^{-1}i} < 0, k = 1, \ldots, d-1 \right\},$$

which is non-empty since if $\tau_i \doteqdot \pi(i) - i$, then $\underline{\tau} \in \Theta_\pi$.

The real-valued function Area$(\cdot)$ associates to $(\underline{\lambda}, \pi, \underline{\tau}) \in \Delta_{d-1} \times \{\pi\} \times \Theta_\pi$,

$$\mathrm{Area}(\underline{\lambda}, \pi, \underline{\tau}) \doteqdot \sum_{k=1}^{d} \lambda_k \left(\sum_{i=1}^{k-1} \tau_i - \sum_{i=1}^{\pi(k)-1} \tau_{\pi^{-1}i} \right).$$

Area$(\cdot)$ has a geometric interpretation as the area of the zippered rectangle associated to the data $(\underline{\lambda}, \pi, \underline{\tau})$ (see, e.g., [**21**]).

Consider the following space as a domain of the natural extension:

$$\hat{\Upsilon}^{(1)}_{\mathfrak{R}} \doteqdot \{(\underline{\lambda}, \pi, \underline{\tau}) \mid (\underline{\lambda}, \pi) \in \Delta(\mathfrak{R}), \underline{\tau} \in \Theta_\pi, \mathrm{Area}((\underline{\lambda}, \pi, \underline{\tau})) = 1\}.$$

The map $\hat{\mathcal{R}} : \hat{\Upsilon}^{(1)}_{\mathfrak{R}} \to \hat{\Upsilon}^{(1)}_{\mathfrak{R}}$ is defined as follows (more precisely $\hat{\mathcal{R}}$ is defined on triples $(\underline{\lambda}, \pi, \underline{\tau})$ such that $(\underline{\lambda}, \pi)$ belong to the domain of $\mathcal{R}$):

$$\hat{\mathcal{R}}((\underline{\lambda}^{(0)}, \pi^{(0)}, \underline{\tau}^{(0)})) = (\mathcal{R}(\underline{\lambda}, \pi), |\underline{\lambda}^{(1)}|\underline{\tau}^{(1)}) = \left(\frac{\underline{\lambda}^{(1)}}{|\underline{\lambda}^{(1)}|}, \pi^{(1)}, |\underline{\lambda}^{(1)}|\underline{\tau}^{(1)} \right),$$

where $(\underline{\lambda}^{(1)}, \pi^{(1)})$ is defined in (11) and, analogously,

$$\underline{\tau}^{(1)} \doteqdot \begin{cases} A^{-1}(\pi, a) \cdot \underline{\tau}, & \lambda_d < \lambda_{\pi^{-1}(d)}, \\ A^{-1}(\pi, b) \cdot \underline{\tau}, & \lambda_d > \lambda_{\pi^{-1}(d)}. \end{cases}$$

The map $\hat{\mathcal{R}}$ preserves an invariant measure $\hat{m}$ which is the restriction to $\hat{\Upsilon}^{(1)}_{\mathfrak{R}}$ of the Lebesgue measure. Denote by p the projection

$$p : \hat{\Upsilon}^{(1)}_{\mathfrak{R}} \to \Delta(\mathfrak{R}), \quad p(\underline{\lambda}, \pi, \underline{\tau}) = (\underline{\lambda}, \pi).$$

The measure $p\hat{m}$ is absolutely continuous with respect to Lebesgue on $\Delta(\mathfrak{R})$ and it is exactly the $\mathcal{R}$-invariant measure $\mu_{\mathcal{V}}$ constructed by Veech.

If $\underline{\gamma}$ is a path on $\mathcal{R}$, starting at π, denote

$$\Theta(\underline{\gamma}) \doteqdot \{A(\gamma)^{-1}\underline{\tau} \mid \underline{\tau} \in \Theta_\pi\} \subset \mathbb{R}^d.$$

If γ is an arrow starting at π and ending at π', then $\hat{\mathcal{R}}$ maps

$$(\Delta(\gamma) \times \{\pi\} \times \Theta_\pi) \cap \hat{\Upsilon}^{(1)}_{\mathcal{R}} \xrightarrow{\hat{\mathcal{R}}} (\Delta_{d-1} \times \{\pi'\} \times \Theta(\gamma)) \cap \hat{\Upsilon}^{(1)}_{\mathcal{R}}.$$

As $\underline{\lambda}$ determines the *future* induction steps (see Remark 2.1), similarly $\underline{\tau}$ determines the *past* steps. More precisely, let $(\underline{\lambda}^{(-i)}, \pi^{(-i)}, \underline{\tau}^{(-i)}) \doteqdot \hat{\mathcal{R}}^{-i}(\underline{\lambda}, \pi, \underline{\tau})$, for $i \in \mathbb{N}$.

Remark 2.2. If π' is the ending vertex of $\underline{\gamma} = (\gamma_1, \ldots, \gamma_r)$ where $\gamma_i = \gamma(\pi_i, c_i)$, $c_i \in \{a, b\}$, and $\underline{\tau} \in \Theta(\underline{\gamma})$, the sequence of types and permutations obtained in the past r steps of $\hat{\mathcal{R}}$ is determined by $\underline{\gamma}$, i.e. $A(\lambda^{(-i)}, \pi^{(-i)}) = A(\gamma_{r-i+1})$ and $\pi^{(-i)} = \pi_{r-i+1}$ for $i = 1, \ldots, r$.

2.2. *Towers construction and heights vectors.* The initial interval exchange T can be seen as a suspension over each of the induced $T^{(r)}$ obtained at the rth step of the Rauzy–Veech algorithm. In this subsection we define the towers which allow one to retrieve T from $T^{(r)}$ and $A^{(r)}$.

Note that the entries of $A^{(r)}$ have a *dynamical meaning* in terms of return times. Namely, denote by $I^{(r)}_j$, $1 \le j \le d$, the subintervals of $T^{(r)}$.

Remark 2.3. The entry $A^{(r)}_{ij}$ is equal to the number of visits of the orbit of any point $x \in I^{(r)}_j$ to the interval $I^{(0)}_i$ of the original partition before its first return in $I^{(r)}$.

Therefore, the norm $h^{(r)}_j$ of the jth column of $A^{(r)}$, i.e. $h^{(r)}_j \doteqdot \sum_{i=1}^d A^{(r)}_{ij}$, gives the return time of any $x \in I^{(r)}_j$ to $I^{(r)}$.

2.2.1. *The towers.* Define

$$Z^{(r)}_j \doteqdot \bigcup_{l=0}^{h^{(r)}_j - 1} T^l I^{(r)}_j. \tag{18}$$

When T is ergodic, $\bigcup_{j=1}^d Z^{(r)}_j$ is a non-trivial T-invariant set, and therefore the sets $Z^{(r)}_j$, $1 \le j \le d$, give a partition of the whole I. Each $Z^{(r)}_j$ can be visualized as a tower over $I^{(r)}_j \subset I^{(r)}$, of height $h^{(r)}_j$ (see Figure 2). A floor of the tower, denoted by $Z^{(r)}_{j,l}$, is defined by $Z^{(r)}_{j,l} \doteqdot T^l I^{(r)}_j$, $l = 0, \ldots, h^{(r)}_j - 1$. The original T is an integral map over $I^{(r)}$; under the action of T every floor $Z^{(r)}_{j,l}$, excluding the top floor ($l \neq h^{(r)}_j$), moves one step up, while $TZ^{(r)}_{j,h^{(r)}_j - 1} = T^{(r)} I^{(r)}_j$.

Let ϕ_s be the partition of $I^{(0)}$ into floors of step s, i.e. whose elements are $Z^{(s)}_{j,l}$ with $1 \le j \le d$ and $0 \le l < h^{(s)}_j$. Partitions $\phi_{s'}$ with $s' > s$ are refinements of ϕ_s.

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

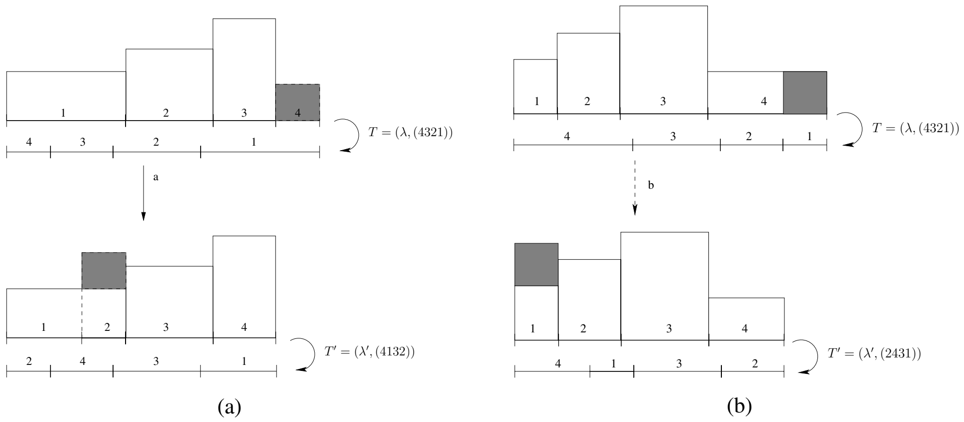

FIGURE 2. Stacking for $T = (\lambda, (4321))$: (a) $\lambda_d < \lambda_{\pi^{-1}(d)}$; (b) $\lambda_d > \lambda_{\pi^{-1}(d)}$.

2.2.2. *Heights cocycle.* Let $\underline{h}^{(0)}$ be the column vector $\underline{e} \doteqdot (1, \ldots, 1)^{\mathrm{T}} \in \mathbb{Z}^d$ and $\underline{h}^{(n)}$ the column vector whose components are the heights $(h_1^{(1)}, \ldots, h_d^{(1)})^{\mathrm{T}}$ of the towers after the first step of the Rauzy–Veech algorithm. If we write $\underline{h}^{(1)} = R\underline{h}^{(0)}$, where $R = R(T)$ is a matrix in $SL(d, \mathbb{Z})$, it is easy to see that $R(T) = A(\pi, a)^{\mathrm{T}}$ or $R(T) = A(\pi, b)^{\mathrm{T}}$, depending on whether the Rauzy step is of type a or b.

Hence, comparing with (12), the cocycle that determines how the vectors of the heights transform is given by the inverse transpose of the Rauzy–Veech cocycle. More precisely, if $\underline{h}^{(r)}$ is the vector of the heights after r iterations of the algorithm, then

$$\underline{h}^{(r)} = (A^{(r)})^{\mathrm{T}}\underline{e}, \quad \underline{h}^{(s+r)} = (A^{(r)}(\mathcal{R}^s T))^{\mathrm{T}}\underline{h}^{(s)}. \tag{19}$$

2.2.3. *Algorithm action on towers.* The Rauzy–Veech algorithm can be visualized as acting on the towers, in terms of *stacking* towers. One step corresponds to cutting the last tower before the permutation, i.e. $Z_d^{(r)}$, and stacking it over $Z_{\pi^{-1}d}^{(r)}$. In the case a, when $\lambda_d^{(r)} < \lambda_{\pi^{-1}d}^{(r)}$, $Z_d^{(r)}$ is completely cut and stacked above $Z_{\pi^{-1}d}^{(r)}$, at its right end (see Figure 2(a)). In the case b, $\lambda_d^{(r)} > \lambda_{\pi^{-1}d}^{(r)}$, only the right portion of $Z_d^{(r)}$ of width $\lambda_{\pi^{-1}d}^{(r)}$ is cut and stacked completely above $Z_{\pi^{-1}d}^{(r)}$ (see Figure 2(b)). It is clear from the stacking description of the algorithm that each tower $Z_{j_0}^{(r)}$ consists of pieces of towers $Z_j^{(s)}$.

2.2.4. *Towers partitions.* Define also the following system of measurable partitions $\xi_s = \xi_s(Z_{j_0}^{(r)})$ of the tower $Z_{j_0}^{(r)}$ in terms of the subtowers $Z_i^{(s)}$, $0 \le s \le r$. The elements of the partition are complete blocks of floors of $Z_{j_0}^{(r)}$ which are all contained inside the same tower $Z_j^{(s)}$: namely, for each floor $Z_{j_0,l}^{(r)}$ which is contained in $I^{(s)}$, construct an element $Z \in \xi_s$ in the following way. If $Z_{j_0,l}^{(r)} \subset I_j^{(s)}$,

$$Z \doteqdot \bigcup_{i=0}^{h_j^{(s)}-1} T^i Z_{j_0,l}^{(r)}. \tag{20}$$

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

The set of all such Z gives a partition ξ_s of $Z^{(r)}_{j_0}$. Clearly for each $Z \in \xi_s$ there is a unique j such that $Z \subset Z^{(s)}_j$. Partitions $\xi_{s'}$, $s' < s$, are refinements of ξ_s.

The entries of $A^{(m,n)}$ have the following meaning for the partition $\xi_m(Z^{(n)}_j)$. For $m < n$, $A^{(m,n)}_{ij}$ gives the number of visits of $x \in I^{(n)}_j$ to $I^{(m)}_i$ under the action of $T^{(m)}$ before the first return to $I^{(n)}$. Hence

$$A^{(m,n)}_{ij} = \#\{Z \in \xi_m(Z^{(n)}_j) \mid Z \subset Z^{(m)}_i\}. \tag{21}$$

3. *Growth of Birkhoff sums of derivatives.*

Let us now introduce two auxiliary functions u, v defined on $I^{(0)}$:

$$u(x) \doteqdot \frac{1}{x}, \quad v(x) \doteqdot \frac{1}{1-x}.$$

PROPOSITION 3.1. *Assume T is uniquely ergodic. There exists a sequence α_r such that $\alpha_r \to 0$ as $r \to \infty$ and for all x distinct from singularities of $S_r(f)$,*

$$S_r(f')(x) = (-C^+ + \alpha_r^+)S_r(u)(x) + (C^- + \alpha_r^-)S_r(v)(x),$$

where $|\alpha_r^\pm| \le \alpha_r$.

Proof. See Theorem 3.1 in [**16**]. The same proof applies also for uniquely ergodic IETs. □

In §3.2 we prove estimates on the growth of the Birkhoff sums for u and v for a typical IET and then we use them in §3.3 to derive some information about the growth of $S_r(f')$ and $S_r(f'')$. It is sufficient to obtain estimates from u, since estimates from v can be easily derived from the following observation. Let $\mathcal{I}(x) = 1 - x$ be the reflection on the interval $I^{(0)}$. Since $v(x) = u(\mathcal{I}x)$, $v \cdot T^n = u \cdot (\mathcal{I} \cdot T \cdot \mathcal{I}^{-1})^n \cdot \mathcal{I}$. Let us denote $T^{\mathcal{I}} \doteqdot \mathcal{I} \cdot T \cdot \mathcal{I}^{-1}$. Hence the Birkhoff sums for v with respect to T and those for u with respect to $T^{\mathcal{I}}$ are related by

$$S_r(v, T)(x) = S_r(u, T^{\mathcal{I}})(1 - x). \tag{22}$$

Note that if $T = ((\lambda_1, \lambda_2, \dots, \lambda_n), \pi)$, then $T^{\mathcal{I}} = ((\lambda_n, \lambda_{n-1}, \dots, \lambda_1), \pi^{\mathcal{I}})$ where $\pi^{\mathcal{I}} \doteqdot (n \quad n-1 \quad \dots \quad 2 \quad 1) \cdot \pi \cdot (n \quad n-1 \quad \dots \quad 2 \quad 1)$. Hence the map $T \mapsto T^{\mathcal{I}}$ from $\Delta_{d-1} \times \mathcal{R}(\pi) \to \Delta_{d-1} \times \mathcal{R}(\pi^{\mathcal{I}})$ preserves the Lebesgue measure.

3.1. *A Diophantine-type condition for IETs.*

In this section we define the set of full measure of IETs for which we prove Theorem 1.1. Proposition 3.2 shows that for typical T one can find a subsequence $\{n_l\}_{l \in \mathbb{N}}$ of induction times such that the corresponding IETs $\{\mathcal{R}^{n_l}T\}_{l\in\mathbb{N}}$ in the Rauzy orbit $\{\mathcal{R}^n T\}_{n\in\mathbb{N}}$ enjoy some properties (listed in Proposition 3.2 below), which we call *balance*; moreover, it gives a control over their frequencies. In the following sections we will use these balanced induction times in order to estimate the growth of $S_r(u)$.

Balanced times are related to occurrences of some positive matrices in the renormalization cocycle. For interval exchanges, conditions on the frequencies of the occurrence of such matrices play an analogous role to Diophantine conditions for rotations. Different types of estimates in this spirit appear in the works of [**2, 3, 10, 21**]. Full measure of the condition that we use is derived from [**2**].

3.1.1. *Existence of balanced return times.*

PROPOSITION 3.2. *Let $1 < \tau < 2$. For each irreducible $\pi \in \mathcal{S}_d$ and for Lebesgue almost every $\lambda \in \Delta_{d-1}$, there exists a subsequence $\{n_l\}_{l\in\mathbb{N}}$ of induction times, $\nu > 1$, $\kappa > 1$, $0 < D < \infty$ and $\bar{l} \in \mathbb{N}$, such that the following hold for all $l \in \mathbb{N}$:*

(1) *ν-balance of lengths*

$$\frac{1}{\nu} \leq \frac{\lambda_i^{(n_l)}}{\lambda_j^{(n_l)}} \leq \nu, \quad \text{for all } 1 \leq i, j \leq d; \tag{23}$$

(2) *κ-balance of heights*

$$\frac{1}{\kappa} \leq \frac{h_i^{(n_l)}}{h_j^{(n_l)}} \leq \kappa, \quad \text{for all } 1 \leq i, j \leq d; \tag{24}$$

(3) *positivity*

$$A^{(n_l, n_{l+\bar{l}})} > 0, \quad \text{and } \dagger \quad D(A^{(n_l, n_{l+\bar{l}})}) \leq D; \tag{25}$$

(4) *integrability*

$$\lim_{l\to+\infty} \frac{\|A^{(n_l, n_{l+1})}\|}{l^\tau} = 0. \tag{26}$$

A return time which satisfies properties (1) and (2) will be called a *balanced return time*. A balanced return time occurs when the lengths and the heights of the induction towers are approximately of the same size. Property (3) gives some uniform distribution of subintervals of time $n_{l+\bar{l}}$ inside the subintervals of the previous balanced time n_l. Property (4) is the Diophantine condition which guarantees some control of the frequencies of occurrence of balanced times. It will be deduced from the power integrability of a certain induced cocycle, proved in [**2**].

We remark that (26) implies for $d = 2$ to the Diophantine condition used for rotations in [**23**], i.e. $k_l = o(l^\tau)$, where $\{k_l\}_{l\in\mathbb{N}}$ are the entries of the continued fraction and the exponent τ satisfies the same assumption $1 < \tau < 2$.

In our proof of mixing we need the condition $\tau < 2$ (see §3.2.2). It would be interesting to know whether mixing also holds for flows over IETs which satisfy properties (1)–(4) in Proposition 3.2 for $\tau > 2$.

Definition 3.1. Let $\mathcal{M}^+ = \mathcal{M}^+(\Delta_{d-1} \times \mathcal{R}(\pi))$ be the set of IETs in $\Delta_{d-1} \times \mathcal{R}(\pi)$ such that Proposition 3.2 holds and $\mathcal{M}^- = \mathcal{M}^-(\Delta_{d-1} \times \mathcal{R}(\pi))$ is the set of $T \in \Delta_{d-1} \times \mathcal{R}(\pi)$ such that $T^{\mathcal{I}} \in \mathcal{M}^+(\Delta_{d-1} \times \mathcal{R}(\pi^{\mathcal{I}}))$. Denote $\mathcal{M} = \mathcal{M}^+ \cap \mathcal{M}^-$.

The IETs in $\mathcal{M}$ are those for which we prove mixing of the suspension flows having one asymmetric logarithmic singularity.

Remark 3.1. The set $\mathcal{M}$ has full measure. Indeed, $\mathcal{M}^+$ has full measure by Proposition 3.2 and also $\mathcal{M}^-$ has full measure since, as already remarked, $T \mapsto T^{\mathcal{I}}$ preserves the Lebesgue measure.

Remark 3.2. The IETs in $\mathcal{M}$ are uniquely ergodic, as follows from property (3) in Proposition 3.2 with the help of techniques used by Veech in [**24, 25**].

† Recall that $D(A)$ was defined in (15).

3.1.2. *Proof of Proposition 3.2.* If $Y \subset \hat{\Upsilon}^{(1)}_{\mathcal{R}}$, let A_Y denote the induced cocycle of the Rauzy–Veech lengths cocycle associated to first returns to Y under $\hat{\mathcal{R}}$, i.e. for $(\underline{\lambda}, \pi, \underline{\tau}) \in Y$,

$$A_Y((\underline{\lambda}, \pi, \underline{\tau})) \doteqdot A^{(r_Y)}((\underline{\lambda}, \pi)), \quad \text{where } r_Y \doteqdot \min\{r \in \mathbb{N}^+ \mid \hat{\mathcal{R}}^r(\underline{\lambda}, \pi, \underline{\tau}) \in Y\}.$$

The following result is proved in [**2**].

THEOREM 3.1. (Avila, Gouëzel, Yoccoz [**2**]) *For every $\delta > 0$ there exists a finite union*

$$\hat{Z}^{(1)} \doteqdot \left(\bigcup_{i=1}^{n} \Delta(\underline{\gamma}_{s_i}) \times \{\pi_i\} \times \Theta(\underline{\gamma}_{e_i})\right) \cap \hat{\Upsilon}^{(1)}_{\mathcal{R}},$$

where π_i is both the initial permutation of the path $\underline{\gamma}_{s_i}$ and the final permutation of the path $\underline{\gamma}_{e_i}$ and where $A(\underline{\gamma}_{s_i}) > 0$ and $A(\underline{\gamma}_{e_i}) > 0$ for all $i = 1, \ldots, n$, such that

$$\int_{\hat{Z}^{(1)}} \|A_{\hat{Z}^{(1)}}\|^{1-\delta} \, d\hat{m} < \infty. \tag{27}$$

Theorem 3.1 is a reformulation of Theorem 4.10 in [**2**]. The original statement claims the integrability of $e^{(1-\delta)r_{\hat{Z}^{(1)}}}$, where $r_{\hat{Z}^{(1)}}$ is the first return time of $(\underline{\lambda}, \pi, \underline{\tau}) \in \hat{Z}^{(1)}$ under the Veech flow, which is given by

$$r_{\hat{Z}^{(1)}}((\underline{\lambda}, \pi, \underline{\tau})) \doteqdot -\log|A^{-1}_{\hat{Z}^{(1)}}\underline{\lambda}| = \log|A_{\hat{Z}^{(1)}}\underline{\lambda}'|,$$

where $(\underline{\lambda}', \pi') = \mathcal{R}^{r_{\hat{Z}^{(1)}}}(\underline{\lambda}, \pi)$. The second equality follows by taking norms of $A_{\hat{Z}^{(1)}}\underline{\lambda}' = \underline{\lambda}/|A^{-1}_{\hat{Z}^{(1)}}\underline{\lambda}|$. Since $\underline{\lambda}'$ belongs to the compact set $\bigcup_{i=1}^{n} \Delta(\underline{\gamma}_{s_i})$,

$$\log|A_{\hat{Z}^{(1)}}\underline{\lambda}'| \geq \log\left(\min_i \lambda'_i \|A_{\hat{Z}^{(1)}}\|\right) \geq \text{const} + \log\|A_{\hat{Z}^{(1)}}\|.$$

Hence (27) follows from the integrability of $e^{(1-\delta)r_{\hat{Z}^{(1)}}}$. Positivity of $A(\underline{\gamma}_{s_i})$ and $A(\underline{\gamma}_{e_i})$ is clear from the proof of Theorem 4.10, in which $\underline{\gamma}_{s_i}$ and $\underline{\gamma}_{e_i}$ are chosen minimal and $(2d-3)$-complete and hence positive by Lemma 3.3.

We recall that Bufetov, by different techniques, obtained in [**3**] a result analogous to (27) for some $\delta < 1$. We need the result of [**2**] for $\delta < 1/2$, since it assures that Proposition 3.2 holds under the condition $\tau < 2$.

Proof of Proposition 3.2. Given $1 < \tau < 2$, let $\delta \doteqdot 1 - \tau^{-1} > 0$. Let $\hat{Z}^{(1)}$ be the corresponding set given by Theorem 3.1. Let $\bar{l}$ be the maximum length of the paths $\underline{\gamma}_{s_i}$ and $\underline{\gamma}_{e_i}$ for $i = 1, \ldots, n$.

Given $(\underline{\lambda}, \pi)$, choose any $\underline{\tau} \in \Theta_\pi$. Let $\{n_l\}_l \in \mathbb{N}$ be the subsequence of visits of the $\hat{\mathcal{R}}$ orbit of $(\underline{\lambda}, \pi, \underline{\tau})$ to $\hat{Z}^{(1)}$ given by

$$n_0 \doteqdot \min\{n \in \mathbb{N}^+ \mid n \geq \bar{l}, \hat{\mathcal{R}}^n(\underline{\lambda}, \pi, \underline{\tau}) \in \hat{Z}^{(1)}\}, \tag{28}$$

$$n_{l+1} \doteqdot \min\{n \in \mathbb{N}^+ \mid n > n_l, \hat{\mathcal{R}}^n(\underline{\lambda}, \pi, \underline{\tau}) \in \hat{Z}^{(1)}\}, \quad l \in \mathbb{N}^+. \tag{29}$$

Notice that (28) is independent of $\underline{\tau}$, since as soon as n is greater than the maximum length $\bar{l}$ of the paths $\underline{\gamma}_{e_i}$, visits to $\hat{Z}^{(1)}$ are determined by $\underline{\lambda}$ only (see Remarks 2.1 and 2.2). Let us show that properties (1), (2) (balance) and (3) (positivity) of Proposition 3.2 automatically hold for $(\underline{\lambda}, \pi)$ and the sequence $\{n_l\}_{l\in\mathbb{N}}$.

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

Since by definition $\hat{\mathcal{R}}^{n_l}(\underline{\lambda}, \pi, \underline{\tau}) \in \Delta(\underline{\gamma}_{s_j}) \times \{\pi_j\} \times \Theta(\underline{\gamma}_{s_j})$ for some j, in particular $\lambda^{(n_l)}/|\lambda^{(n_l)}| \in \Delta(\underline{\gamma}_{s_j})$. By positivity of the $A(\underline{\gamma}_{s_i}) > 0$, the union $\bigcup_i \Delta(\underline{\gamma}_{s_i})$ is compact and hence (see §2.1.5) contained in a ball for the Hilbert metric d_{H}, centered at $(1/d, \ldots, 1/d)$, of some radius $r_s > 0$. Hence,

$$d_{\mathrm{H}}\bigg(\frac{\underline{\lambda}^{(n_l)}}{|\underline{\lambda}^{(n_l)}|}, \bigg(\frac{1}{d}, \ldots, \frac{1}{d}\bigg)\bigg) \leq r_s \quad \text{or equivalently} \quad \frac{\max_i \lambda_i^{(n_l)}}{\min_i \lambda_i^{(n_l)}} \leq e^{r_s},$$

which, setting $\nu \doteqdot e^{r_s} > 1$, is ν-balance of lengths.

Similarly, from $\underline{\tau}^{(n_l)} \in \Theta(\underline{\gamma}_{e_j})$ we obtain by Remark 2.2 that $\underline{\lambda}^{(n_l)} = A(\underline{\gamma}_{e_j})^{-1}\underline{\lambda}^{(n_l-L)}$, where L is the length of $\underline{\gamma}_{e_j}$. Since the heights transform according to (19), $\underline{h}^{(n_l)} = A(\underline{\gamma}_{e_j})^{\mathrm{T}}\underline{h}^{(n_l-L)}$. Arguing as above, by compactness, the union $\bigcup_i \widetilde{A(\underline{\gamma}_{e_i})^{\mathrm{T}}}\Delta_{d-1}$ is contained in a ball centered at $(1/d, \ldots, 1/d)$ of some radius $r_e > 0$ and this gives $\kappa \doteqdot e^{r_e}$ balance of the heights.

For property (3), since $\overline{l}$ is the maximum length of the paths $\underline{\gamma}_{s_i}$ and $n_{l+\overline{l}} \geq n_l + \overline{l}$, by Remark 2.1 we have $A^{(n_l, n_{l+\overline{l}})} = A(\underline{\gamma}_{s_j})A$ for some $A \geq 0$. Hence $A^{(n_l, n_{l+\overline{l}})} > 0$ and $D(A^{(n_l, n_{l+\overline{l}})}) \leq D(A(\underline{\gamma}_{s_j})) \leq 2r_s$.

Let us show first that property (4) holds for typical $(\underline{\lambda}, \pi, \underline{\tau}) \in \hat{Z}^{(1)}$. Note that $A^{(n_l, n_{l+1})}(\underline{\lambda}, \pi, \underline{\tau}) = A_{\hat{Z}^{(1)}}(\hat{\mathcal{R}}^{n_l}(\underline{\lambda}, \pi, \underline{\tau}))$. For each $\epsilon_k \doteqdot 1/k$, by $\hat{\hat{\mathcal{R}}}$ invariance of $\hat{m}$,

$$\begin{aligned}&\hat{m}\{(\underline{\lambda}, \pi, \underline{\tau}) \in \hat{Z}^{(1)} \mid \|A^{(n_l, n_{l+1})}(\underline{\lambda}, \pi, \underline{\tau})\| \geq \epsilon_k l^{\tau}\} \\ &\quad = \hat{m}\{(\underline{\lambda}, \pi, \underline{\tau}) \in \hat{Z}^{(1)} \mid \|A_{\hat{Z}^{(1)}}(\underline{\lambda}, \pi, \underline{\tau})\|^{\tau^{-1}}\epsilon_k^{-\tau^{-1}} \geq l\}.\end{aligned}$$

Since we chose $\tau^{-1} = 1 - \delta$, the integrability condition (27) implies that $\sum_l \hat{m}\{\|A_{\hat{Z}^{(1)}}\|^{\tau^{-1}}\epsilon_k^{-\tau^{-1}} \geq l\} < \infty$ for each ϵ_k. Hence, it follows by a Borel–Cantelli type of argument that there exists a subset $\hat{Z}' \subset \hat{Z}^{(1)}$ with $\hat{m}(\hat{Z}') = \hat{m}(\hat{Z}^{(1)})$ such that for $(\underline{\lambda}, \pi, \underline{\tau}) \in \hat{Z}'$ the sequence n_l satisfies (26).

Consider the projection $p\hat{Z}' \subset \Delta(\mathcal{R})$. By independence of the definition of $\{n_l\}$ on τ, for each $(\underline{\lambda}, \pi) \in p\hat{Z}'$ we have (26). Moreover, $p\hat{Z}'$ has $p\hat{m}$-full measure in $p\hat{Z}^{(1)} = \bigcup_i \Delta(\underline{\gamma}_{s_i}) \times \{\pi_i\}$ and, in particular, positive $p\hat{m} = \mu_{\mathcal{Y}}$-measure and hence also $\mu_{\mathcal{Z}}$-positive measure.

To conclude, let $\mathcal{M}^+$ be the set of $T \in \Delta(\mathcal{R})$ such that there exists $\overline{n}$, for which $\mathcal{R}^{\overline{n}}T \in p\hat{Z}'$. Clearly, if $T \in \mathcal{M}^+$, all the properties are satisfied by the sequence $n_l \doteqdot \widetilde{n}_l + \overline{n}$, where $\widetilde{n}_l$ is the sequence associated to $\mathcal{R}^{\overline{n}}T$. To see that $\mathcal{M}^+$ has full measure, it is enough to use the ergodicity of $\mathcal{Z}$ and the fact that $\mu_{\mathcal{Z}}(p\hat{Z}') > 0$, remarking that $\mathcal{Z}$ orbits are subsets of $\mathcal{R}$ orbits. The formulation in Proposition 3.2 follows by absolute continuity of $\mu_{\mathcal{Z}}$ with respect to Lebesgue. □

3.1.3. *Some consequences of balance.* We will also frequently use the following simple lemmas. Recall that $A > 0$ means that A has strictly positive entries.

LEMMA 3.1. *Let $A_i > 0$, $A_i \in SL(d, \mathbb{Z})$ for $i = 0, \ldots, n$. If $\underline{\lambda} = A_0 \cdot \underline{\lambda}'$, then $\sum_j \lambda_j > d \sum_j \lambda'_j$.*

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

If $\underline{h} = A_0 \cdot \underline{h}'$, *then* $\min_j h_j > d \min_j h'_j$. *In particular, if* $\underline{h} = A_1 \cdots A_n \underline{e}$, *then* $\min_j h_j \geq d^n$.

Proof. All the properties follow easily from $A_{ij} \geq 1$. □

For convenience, let us denote $\lambda^{(n_l)} \doteqdot \sum_j \lambda_j^{(n_l)}$ and $h^{(n_l)} \doteqdot \max_j h_j^{(n_l)}$.

COROLLARY 3.1. *For each* $L \in \mathbb{N}$, $\log(h^{(n_{L\bar{l}})}) \geq L \log d$. *In particular,*

$$\lim_{l \to +\infty} \frac{1}{\log h^{(n_l)}} = 0. \tag{30}$$

Proof. For the first property, apply Lemma 3.1 to $h^{(n_{L\bar{l}})} = {A^{(n_{L\bar{l}})}}^{\mathrm{T}} \underline{e}$ and note that ${A^{(n_{L\bar{l}})}}^{\mathrm{T}}$ is the product of at least L positive matrices by property (3) of Proposition 3.2 (see (25)) of the sequence $\{n_l\}_{l \in \mathbb{N}}$. It follows that $\log h^{(n_l)} \geq [l/\bar{l}] \log d$, where $[\cdot]$ denotes the integer part, and hence we obtain (30). □

LEMMA 3.2. *If* n_l *has* ν*-balanced lengths (23) and* κ*-balanced heights (24), then, for each* $j = 0, \ldots, d$,

$$\frac{1}{\kappa \lambda^{(n_l)}} \leq h_j^{(n_l)} \leq \frac{\kappa}{\lambda^{(n_l)}}, \tag{31}$$

$$\frac{1}{d \kappa \nu h^{(n_l)}} \leq \lambda_j^{(n_l)} \leq \frac{\kappa}{h^{(n_l)}}. \tag{32}$$

Hence, if n_l is a balanced time, $\lambda^{(n_l)} \approx 1/h^{(n_l)}$ up to constants.

Proof. Since by (12) and (19) we have $\sum_i h_i^{(n_l)} \lambda_i^{(n_l)} = 1$, we obtain $\min_i h_i^{(n_l)} \lambda^{(n_l)} \leq 1$ and $\max_i h_i^{(n_l)} \lambda^{(n_l)} \geq 1$. In conjunction with balanced heights (24), this gives

$$\frac{1}{\kappa \lambda^{(n_l)}} \leq \frac{1}{\kappa} \max_i h_i^{(n_l)} \leq h_j^{(n_l)} \leq \kappa \min_i h_i^{(n_l)} \leq \frac{\kappa}{\lambda^{(n_l)}}.$$

To show (32), let i be such that $h_i^{(n_l)} \lambda_i^{(n_l)} = \max_j h_j^{(n_l)} \lambda_j^{(n_l)} \geq 1/d$. Also, for each j, $h_j^{(n_l)} \lambda_j^{(n_l)} < 1$. Furthermore, when the lengths balance (23),

$$\frac{1}{\kappa \nu d h^{(n_l)}} \leq \frac{1}{\nu d h_i^{(n_l)}} \leq \frac{\lambda_i^{(n_l)}}{\nu} \leq \lambda_j^{(n_l)} \leq \frac{1}{h_j^{(n_l)}} \leq \frac{\kappa}{h^{(n_l)}}.$$ □

LEMMA 3.3. *For each fixed* $L \in \mathbb{N}$,

$$\lim_{l \to +\infty} \frac{\log \|A^{(n_l, n_{l+L})}\|}{\log h^{(n_l)}} = \lim_{l \to +\infty} \frac{\log \|A^{(n_{l-L}, n_l)}\|}{\log h^{(n_l)}} = 0. \tag{33}$$

Proof. By Corollary 3.1, $\log h^{(n_l)} \geq [l/\bar{l}] \log d$. Hence

$$\lim_{l \to +\infty} \frac{\log \|A^{(n_l, n_{l+L})}\|}{\log h^{(n_l)}} \leq \lim_{l \to +\infty} \frac{\sum_{i=l}^{l+L-1} \log \|A^{(n_i, n_{i+1})}\|}{(l/\bar{l} - 1) \log d}. \tag{34}$$

Using the property (26), each of the L terms in the sum on the right-hand side can be bounded for $l \gg 1$ by $\tau \log i \leq \tau \log(l + L - 1)$ and hence the first limit is zero. The second limit is analogous. □

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

3.2. *Growth of Birkhoff sums of u.* Assume in this section that the roof function is $u(x) = 1/x$.

3.2.1. *Growth of Birkhoff sums along a balanced tower.* In order to understand the asymptotic growth of $S_r(u)(x)$, we first consider $S_r(u)(x)$ when $x \in I_j^{(n)}$ is a point in the base of the tower $Z_j^{(n)}$, $r = h_j^{(n)}$ is exactly the height of the same tower and n is one of the balanced return times constructed in §3.2. This preliminary estimate is used in §3.2.2 as a building block to obtain an estimate for any r and most of the other points x.

PROPOSITION 3.3. *Assume $T \in \mathcal{M}^+$. Let n_{l_0} be a balanced return time given by Proposition 3.2. Let $x_0 \in I_{j_0}^{(n_{l_0})}$ be a point belonging to the base of the tower $Z_{j_0}^{(n_{l_0})}$ and let $r_0 = h_{j_0}^{(n_{l_0})}$ be the corresponding tower height. Given $\varepsilon > 0$, there exists $l(\varepsilon)$ such that, for $l_0 \geq l(\varepsilon)$,*

$$(1-\varepsilon) h_{j_0}^{(n_{l_0})} \ln(h^{(n_{l_0})}) \leq S_{r_0}(u)(x_0) - \frac{1}{x_0} \leq (1+\varepsilon) h_{j_0}^{(n_{l_0})} \ln(h^{(n_{l_0})}). \tag{35}$$

A Birkhoff sum of the form $S_{r_0}(u)(x_0)$ where x_0 and r_0 satisfy the hypotheses of Proposition 3.3 will be referred as the *Birkhoff sum along a tower*. The proposition shows that each sum along a tower gives a contribution of order $r_0 \log(r_0)$, plus the contribution of the closest point to the singularity, x_0, which could be arbitrary large and will be estimated separately when using these sums as building blocks in §3.2.2.

Proof of Proposition 3.3. Consider the inducing intervals $I^{(n_l)} = [0, \lambda^{(n_l)})$ where $\{n_l\}_{l\in\mathbb{N}}$ is the sequence of balanced induction times constructed in §3.1. Given any $\epsilon > 0$, let D and $\bar{l}$ be given by Proposition 3.2. Choose $L_1 \in \mathbb{N}$ such that

$$(1 - e^{-D})^{L_1 - 1} D < \epsilon. \tag{36}$$

We remark that $\operatorname{diam}(\widetilde{A^{(n_{\bar{l}})}\Delta_{d-1}}) < \infty$ by (25). Choose also $L_2 \in \mathbb{N}$ such that $1/d^{L_2} < \epsilon$. Assume $l_0 \geq \bar{l}(1 + L_1 + L_2)$. For convenience, introduce the notation

$$l_{-1} \doteqdot l_0 - L_1\bar{l}, \quad l_{-2} \doteqdot l_{-1} - L_2\bar{l} = l_0 - (L_1 + L_2)\bar{l}. \tag{37}$$

The past induction times $n_{l_{-1}}$ and $n_{l_{-2}}$ will play the following role in the proof: $n_{l_{-1}}$ is such that the elements of the orbit $\{T^r(x_0)\}_{0\leq r<r_0}$ are uniformly distributed inside the elements of the partition $\xi_{n_{l_{-1}}}$; $n_{l_{-2}}$ is such that the main contribution to $S_{r_0}(u)(x_0)$ comes from visits to $[\lambda^{(n_{l_{-2}})}, 1)$.

We denote the points along the T-orbit of x_0 by $x_i = T^i(x_0)$, $0 \leq i < r_0$. Since the original interval can be partitioned as

$$I^{(0)} = I^{(n_{l_0})} \cup I^{(n_{l_{-2}})} \backslash I^{(n_{l_0})} \cup [\lambda^{(n_{l_{-2}})}, 1),$$

and $x_i \notin I^{(n_{l_0})}$, for $1 \leq i < r_0$ because r_0 is by definition the first return time of x_0 to $I^{(n_{l_0})}$, the Birkhoff sums can be decomposed as follows:

$$S_{r_0}(u)(x_0) = \sum_{i=0}^{r_0-1} \frac{1}{x_i} = \frac{1}{x_0} + \sum_{x_i \in I^{(n_{l_{-2}})} \backslash I^{(n_{l_0})}} \frac{1}{x_i} + \sum_{x_i \in [\lambda^{(n_{l_{-2}})}, 1)} \frac{1}{x_i}. \tag{38}$$

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

We will refer to the first term on the right-hand side of (38) as the *singular error*, to the sum which appears as the second term on the right-hand side as the *gap error*, while the sum which appears as the last term determines the *main contribution*.

3.2.1.1. Uniform ergodic convergence. Recall that $A_{ij}^{(n_{l_{-1}},n_{l_0})}$ gives the number of visits of $x \in I_j^{(n_{l_0})}$ to $I_i^{(n_{l_{-1}})}$ before the time $h_j^{(n_{l_0})}$ of first return to $I^{(n_{l_0})}$.

LEMMA 3.4. (Uniform distribution) *For each* $1 \leq i,\ j \leq d$,

$$e^{-2\epsilon}\lambda_i^{(n_{l_{-1}})} \leq \frac{A_{ij}^{(n_{l_{-1}},n_{l_0})}}{h_j^{(n_{l_0})}} \leq e^{2\epsilon}\lambda_i^{(n_{l_{-1}})}. \tag{39}$$

Proof. Consider the sets $\widetilde{A^{(n_{l_{-1}},n)}}\Delta_{d-1} \subset \Delta_{d-1}$, for $n > n_{l_{-1}}$, which form a nested sequence of compact sets. By the transformation formula (12) for lengths vectors $\underline{\lambda}^{(n_{l_{-1}})} = A^{(n_{l_{-1}},n)}\underline{\lambda}^{(n)}$, the normalized vector

$$\frac{\underline{\lambda}^{(n_{l_{-1}})}}{\lambda^{(n_{l_{-1}})}} \in \bigcap_{n>n_{l_{-1}}} \widetilde{A^{(n_{l_{-1}},n)}}\Delta_{d-1}.$$

When $n = n_{l_0}$, since $l_0 = l_{-1} + L_1\bar{l}$, applying L_1 times property (3) (positivity) in Proposition 3.2 through the contraction property (16), we obtain

$$\operatorname{diam}_{\mathrm{H}}(\widetilde{A^{(n_{l_{-1}},n_{l_0})}}\Delta_{d-1}) \leq (1-e^{-D})^{L_1-1}D \leq \epsilon, \tag{40}$$

where the last inequality follows by the choice (36) of L_1.

Denote by $\underline{e}_j$ the unit vector $(\underline{e}_j)_i = \delta_{ij}$ (δ is the Kronecker symbol). Since both the vectors $\widetilde{A^{(n_{l_{-1}},n_{l_0})}}\underline{e}_j$ and $\underline{\lambda}^{(n_{l_{-1}})}/\lambda^{(n_{l_{-1}})}$ belong to the closure of $\widetilde{A^{(n_{l_{-1}},n_{l_0})}}\Delta_{d-1}$, it follows by (40), using compactness, that

$$d_{\mathrm{H}}\bigg(\frac{\underline{\lambda}^{(n_{l_{-1}})}}{\lambda^{(n_{l_{-1}})}}, \widetilde{A^{(n_{l_{-1}},n_{l_0})}}\underline{e}_j\bigg) = \log\frac{\max_{i=1,\ldots,d} A_{ij}^{(n_{l_{-1}},n_{l_0})}/\lambda_i^{(n_{l_{-1}})}}{\min_{i=1,\ldots,d} A_{ij}^{(n_{l_{-1}},n_{l_0})}/\lambda_i^{(n_{l_{-1}})}} \leq \epsilon,$$

where we also used the invariance of the distance expression by multiplication of the arguments by a scalar. Equivalently, for each $1 \leq i, k \leq d$,

$$e^{-\epsilon}(A_{kj}^{(n_{l_{-1}},n_{l_0})}\lambda_i^{(n_{l_{-1}})}) \leq A_{ij}^{(n_{l_{-1}},n_{l_0})}\lambda_k^{(n_{l_{-1}})} \leq e^{\epsilon}(A_{kj}^{(n_{l_{-1}},n_{l_0})}\lambda_i^{(n_{l_{-1}})}) \tag{41}$$

and summing over k we get

$$e^{-\epsilon} \leq \frac{A_{ij}^{(n_{l_{-1}},n_{l_0})}\lambda^{(n_{l_{-1}})}}{\sum_k A_{kj}^{(n_{l_{-1}},n_{l_0})}\lambda_i^{(n_{l_{-1}})}} \leq e^{\epsilon}. \tag{42}$$

If we multiply (41) by $h_i^{(n_{l_{-1}})}$ and then also sum over i, using $\sum_i h_i^{(n_{l_{-1}})}\lambda_i^{(n_{l_{-1}})} = 1$ and $\underline{h}^{(n_{l_0})} = A^{(n_{l_{-1}},n_{l_0})^{\mathrm{T}}}\underline{h}^{(n_{l_{-1}})}$,

$$e^{-\epsilon} \leq \frac{h_j^{(n_{l_0})}\lambda^{(n_{l_{-1}})}}{\sum_k A_{kj}^{(n_{l_{-1}},n_{l_0})}} \leq e^{\epsilon}. \tag{43}$$

The combination of (42) and (43) gives (39). □

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

3.2.1.2. Estimate of the main contribution. The following lemma shows that the main contribution in (38) determines the order of the Birkhoff sum in Proposition 3.3.

LEMMA 3.5. (Main contribution) *For each* $\epsilon > 0$, *if* $l_0 > l_m(\epsilon)$,

$$e^{-2\epsilon}(1-\epsilon)^2 h_{j_0}^{(n_{l_0})} \log h^{(n_{l_0})} \leq \sum_{x_i \in [\lambda^{(n_{l-2})},1)} \frac{1}{x_i} \leq e^{2\epsilon}(1+\epsilon)^2 h_{j_0}^{(n_{l_0})} \log h^{(n_{l_0})}. \tag{44}$$

Proof. Consider the partition $\phi_{n_{l_{-1}}}$, introduced in §2.2.1, restricted to $[\lambda^{(n_{l-2})}, 1)$, which is measurable with respect to $\phi_{n_{l_{-1}}}$. Recall that the elements $F_\alpha \in \phi_{n_{l_{-1}}}$ are floors $F_\alpha = T^k(I_{j_\alpha}^{(n_{l-1})})$ for some $1 \leq j_\alpha \leq d$ and $0 \leq k < h_{j_\alpha}^{(n_{l-1})}$. In particular, $\mathrm{Leb}(F_\alpha) = \lambda_{j_\alpha}^{(n_{l-1})}$. For each F_α choose, by the mean value theorem, a point $\bar{x}_\alpha$ such that

$$\frac{1}{\bar{x}_\alpha} \doteqdot \frac{1}{\lambda_{j_\alpha}^{(n_{l-1})}} \int_{F_\alpha} \frac{1}{s}\, ds. \tag{45}$$

LEMMA 3.6. *If* $x_i \in F_\alpha$ *and* $F_\alpha \subset [\lambda^{(n_{l-2})}, 1)$,

$$1-\epsilon \leq \frac{1/x_i}{1/\bar{x}_\alpha} \leq 1+\epsilon.$$

Proof. Let $F_\alpha = [a, b)$; then $a \leq x_i, \bar{x}_\alpha < b$. Since by assumption $b - a \leq \lambda^{(n_{l-1})}$ and $a \geq \lambda^{(n_{l-2})}$,

$$\frac{\bar{x}_\alpha}{x_i} \leq \frac{b}{a} = \frac{a+(b-a)}{a} \leq 1 + \frac{\lambda^{(n_{l-1})}}{\lambda^{(n_{l-2})}},$$
$$\frac{\bar{x}_\alpha}{x_i} \geq \frac{a}{b} = \frac{b-(b-a)}{b} \geq 1 - \frac{\lambda^{(n_{l-1})}}{\lambda^{(n_{l-2})}}.$$

Let us show that $\lambda^{(n_{l-1})}/\lambda^{(n_{l-2})} < \epsilon$. Since $l_{-1} = l_{-2} + L_2\bar{l}$,

$$\underline{\lambda}^{(n_{l-2})} = \prod_{i=0}^{L_2-1} A^{(n_{l_{-2}+i\bar{l}}, n_{l_{-2}+(i+1)\bar{l}})} \underline{\lambda}^{(n_{l-1})}$$

and each of the matrices in the product has positive entries by property (3) in Proposition 3.2. Hence by iterated application of Lemma 3.1 and by the choice of L_2, we obtain $\lambda^{(n_{l-1})}/\lambda^{(n_{l-2})} < 1/d^{L_2} < \epsilon$. □

Rearranging the main contribution in (38) by floors, i.e.

$$\sum_{x_i \in [\lambda^{(n_{l-2})},1)} \frac{1}{x_i} = \sum_{\substack{F_\alpha \in \phi_{n_{l_{-1}}},\\ F_\alpha \subset [\lambda^{(n_{l-2})},1)}} \sum_{x_i \in F_\alpha} \frac{1}{x_i},$$

and applying Lemma 3.6, we obtain

$$\sum_{F_\alpha \subset [\lambda^{(n_{l-2})},1)} \sum_{x_i \in F_\alpha} \frac{1-\epsilon}{\bar{x}_\alpha} \leq \sum_{F_\alpha \subset [\lambda^{(n_{l-2})},1)} \sum_{x_i \in F_\alpha} \frac{1}{x_i} \leq \sum_{F_\alpha \subset [\lambda^{(n_{l-2})},1)} \sum_{x_i \in F_\alpha} \frac{1+\epsilon}{\bar{x}_\alpha}. \tag{46}$$

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

Now consider

$$\sum_{x_i \in F_\alpha} \frac{1}{\bar{x}_\alpha} = \#\{x_i \in F_\alpha\} \frac{1}{\bar{x}_\alpha}. \tag{47}$$

Recall that $x_0 \in I_{j_0}^{(n_{l_0})}$; if $F_\alpha = T^k(I_{j_\alpha}^{(n_{l_{-1}})})$ is a floor of the j_αth tower $Z_{j_\alpha}^{(n_{l_{-1}})}$,

$$\#\{x_i \in F_\alpha\} = \#\{x_i \in I_{j_\alpha}^{(n_{l_{-1}})}\} = A_{j_\alpha j_0}^{(n_{l_{-1}}, n_{l_0})}. \tag{48}$$

In (48) we used the dynamical meaning of $A_{j_\alpha j_0}^{(n_{l_{-1}}, n_{l_0})}$ together with the fact that $Z_{j_0}^{(n_{l_0})}$ is decomposed into a whole number of elements of $\xi_{n_{l_{-1}}}$ corresponding to towers of the previous step $n_{l_{-1}}$; hence visits to a floor $T^k I_{j_\alpha}^{(n_{l_{-1}})}$ of the tower are in one-to-one correspondence with visits to its base $I_{j_\alpha}^{(n_{l_{-1}})}$.

From Lemma 3.4 and (48),

$$e^{-2\epsilon} \lambda_{j_\alpha}^{(n_{l_{-1}})} h_{j_0}^{(n_{l_0})} \le \#\{x_i \in F_\alpha\} \le e^{2\epsilon} \lambda_{j_\alpha}^{(n_{l_{-1}})} h_{j_0}^{(n_{l_0})}. \tag{49}$$

Using this bound and recalling the definition (45) of $\overline{x}_\alpha$, we obtain

$$e^{-2\epsilon} h_{j_0}^{(n_{l_0})} \int_{F_\alpha} \frac{1}{s}\, ds \le \#\{x_i \in F_\alpha\} \frac{1}{\bar{x}_\alpha} \le e^{2\epsilon} h_{j_0}^{(n_{l_0})} \int_{F_\alpha} \frac{1}{s}\, ds.$$

Summing it over $F_\alpha \subset [\lambda^{(n_{l_{-2}})}, 1)$ and using

$$\sum_{F_\alpha \subset [\lambda^{(n_{l_{-2}})}, 1)} \int_{F_\alpha} \frac{1}{s}\, ds = \int_{[\lambda^{(n_{l_{-2}})}, 1)} \frac{1}{s}\, ds = \log \frac{1}{\lambda^{(n_{l_{-2}})}},$$

we get by (46)

$$e^{-2\epsilon}(1-\epsilon) h_{j_0}^{(n_{l_0})} \log \frac{1}{\lambda^{(n_{l_{-2}})}} \le \sum_{F_\alpha \subset [\lambda^{(n_{l_{-2}})}, 1)} \sum_{x_i \in F_\alpha} \frac{1}{x_i} \le e^{2\epsilon}(1+\epsilon) h_{j_0}^{(n_{l_0})} \log \frac{1}{\lambda^{(n_{l_{-2}})}}. \tag{50}$$

In order to obtain the estimate (44) of Lemma 3.5 from (50) it is sufficient to compare $1/\lambda^{(n_{l_{-2}})}$ with $h^{(n_{l_0})}$. Since $\lambda^{(n_{l_{-2}})} > \lambda^{(n_{l_0})} \ge 1/(\kappa h^{(n_{l_0})})$ by κ-balance of heights (see Lemma 3.2), we obtain

$$\log(1/\lambda^{(n_{l_{-2}})}) \le \log(h^{(n_{l_0})})(1 + \log \kappa / \log(h^{(n_{l_0})}))$$

and, if $l_0 \ge l_m$ for some $l_m(\epsilon) > 0$, the upper estimate in (44) by Corollary 3.1.

For the lower bound, adding and subtracting $h_{j_0}^{(n_{l_0})} \log h^{(n_{l_0})}$,

$$h_{j_0}^{(n_{l_0})} \log \frac{1}{\lambda^{(n_{l_{-2}})}} = h_{j_0}^{(n_{l_0})} \log h^{(n_{l_0})} \left(1 - \frac{\log(h^{(n_{l_0})} \lambda^{(n_{l_{-2}})})}{\log h^{(n_{l_0})}}\right). \tag{51}$$

In order to estimate the very last term in (51), notice that, again by Lemma 3.2 and balance, $\lambda^{(n_{l_{-2}})} \le \kappa / h^{(n_{l_{-2}})}$. Hence, using the fact that $h^{(n_{l_0})} / h^{(n_{l_{-2}})} \le \|A^{(n_{l_{-2}}, n_{l_0})}\|$,

$$\frac{\log(h^{(n_{l_0})} \lambda^{(n_{l_{-2}})})}{\log h^{(n_{l_0})}} \le \frac{\log \kappa + \log \|A^{(n_{l_{-2}}, n_{l_0})}\|}{\log h^{(n_{l_0})}}. \tag{52}$$

Enlarging l_m if necessary, the right-hand side is less than ϵ for $l_0 \ge l_m$ by Lemma 3.3 (recall that the difference $l_0 - l_{-2} = (L_1 + L_2)\bar{l}$ is fixed) and Corollary 3.1.

Combining (52) and (51) to estimate the left-hand side of (50) from below, we obtain the lower estimate that completes the proof of the lemma. □

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

3.2.1.3. Estimate of the gap error.

LEMMA 3.7. (Gap error) *For each ϵ, if $l_0 > l_g(\epsilon)$,*

$$0 \leq \sum_{x_i \in I^{(n_{l-2})} \setminus I^{(n_{l_0})}} \frac{1}{x_i} \leq \epsilon (h_{j_0}^{(n_{l_0})} \log h^{(n_{l_0})}). \tag{53}$$

Proof. The bound below is trivial since $1/x_i > 0$. Since we are considering $0 \leq i < r_0 = h_{j_0}^{(n_{l_0})}$, it follows from the tower construction (see §2.2.1) that the points $x_i = T^i x_0$ of the orbit of $x_0 \in I_{j_0}^{(n_{l_0})}$ belong to different floors of the tower $Z_{j_0}^{(n_{l_0})}$ and that their minimum distance is bounded from below by

$$\min_{0 \leq i, j < r_0} |x_i - x_j| \geq \lambda_{j_0}^{(n_{l_0})} \geq \frac{1}{d\kappa \nu h_{j_0}^{(n_{l_0})}},$$

where in the last inequality we used the fact that n_{l_0} is balanced and Lemma 3.2.

Noting also that x_0 is the closest point to the singularity, it follows that if we rearrange the x_i in increasing order and relabel them $\tilde{x}_i$ ($\tilde{x}_i < \tilde{x}_{i+1}$), we have

$$\tilde{x}_i \geq x_0 + \frac{i}{d\kappa \nu h_{j_0}^{(n_{l_0})}}, \quad i = 0, \ldots, r_0 - 1.$$

Since the roof function $1/x$ is monotonically decreasing, the gap error can be bounded from above by

$$\sum_{x_i \in I^{(n_{l-2})} \setminus I^{(n_{l_0})}} \frac{1}{x_i} \leq \sum_{k=1}^{K} \frac{1}{x_0 + k/d\kappa \nu h_{j_0}^{(n_{l_0})}}, \tag{54}$$

where $K = \#\{x_i \in I^{(n_{l-2})} \setminus I^{(n_{l_0})}\}$ and $k \geq 1$ since $x_0 \in I^{(n_{l_0})}$.

The following lemma is proved by Kochergin in [**16**] as Lemma 5.1.

LEMMA 3.8. *Let $h > 0$ and $x > 0$.*

$$\sum_{k=1}^{K} \frac{1}{x + kh} = \frac{1}{h} \log\left(\frac{t_0 + K}{t_0 + 1}\right) + \frac{1}{h} R_K(t_0),$$

where $t_0 = x/h$ and $0 < R_K(t_0) < 1/(t_0 + 1)$.

Applying Lemma 3.8 to (54) and using the fact that $\log((t_0 + K)/(t_0 + 1))$ is decreasing in t_0, so it reaches its maximum $\log K$ at $t_0 = 0$,

$$\sum_{x_i \in I^{(n_{l-2})} \setminus I^{(n_{l_0})}} \frac{1}{x_i} \leq d\kappa \nu h_{j_0}^{(n_{l_0})} (\log K + 1). \tag{55}$$

The cardinality K of points $x_i \in I^{(n_{l-2})} \setminus I^{(n_{l_0})}$ can be bounded by Remark 2.3 in terms of the cocycle matrices by

$$K \leq \sum_{j=1}^{d} A_{j j_0}^{(n_{l-2}, n_{l_0})} \leq \| A^{(n_{l-2}, n_{l_0})} \|. \tag{56}$$

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

Hence, applying (55) and (56) we obtain

$$\frac{\sum_{x_i \in I^{(n_{l-2})} \setminus I^{(n_{l_0})}} 1/x_i}{h_{j_0}^{(n_{l_0})} \log h^{(n_{l_0})}} \leq d\kappa\nu \frac{\log K + 1}{\log h^{(n_{l_0})}} \leq d\kappa\nu \frac{\log \|A^{(n_{l_0 - (L_1+L_2)\bar{l}}, n_{l_0})}\| + 1}{\log h^{(n_{l_0})}}. \tag{57}$$

The right-hand side can be made smaller than ϵ by again using Lemma 3.3 and Corollary 3.1 as long as $l_0 \geq l_g$ for some $l_g(\epsilon) \in \mathbb{N}$. □

Recalling the decomposition (38) of the Birkhoff sums, the estimates of the main contribution and of the gap error in Lemma 3.5 and Lemma 3.7 combine together, for $l_0 \geq l(\epsilon) \doteqdot \max\{l_m, l_g\}$, to yield the estimate in Proposition 3.3. □

3.2.2. *Growth of Birkhoff sums for other points.* In this section we obtain an estimate for $S_r(u)(x)$ using the estimate found in §3.2.1 as a fundamental block, i.e. decomposing $S_r(u)(x)$ into pieces which correspond to Birkhoff sums along a tower. It turns out that singular errors from points in the bottom floors of the towers (see the terminology introduced just after (38)) could prevent us from obtaining an estimate of order $r \log r$. In order to obtain this type of asymptotic behaviour, it is necessary to throw away a set of initial points x which has an arbitrarily small measure. The integrability condition (26) of the sequence of balanced times is used in its full strength only in this part.

3.2.2.1. Preliminary notation. Let $\{n_l\}_{l\in\mathbb{N}}$ be the sequence of balanced times in Proposition 3.2. Assume $h^{(n_l)} \leq r < h^{(n_{l+1})}$. Define the sequence $\{\sigma_l\}_{l\in\mathbb{N}}$ used in the proof of Proposition 3.4 as a threshold to determine whether r is closer to $h^{(n_l)}$ or to $h^{(n_{l+1})}$. Let τ' be such that $\tau/2 < \tau' < 1$, where τ is the Diophantine exponent in (26) given by Proposition 3.2 and τ' is well defined since $\tau < 2$. Let

$$\sigma_l = \sigma_l(T) \doteqdot \left(\frac{\log \|A^{(n_l, n_{l+1})}\|}{\log h^{(n_l)}}\right)^{\tau'}, \qquad \frac{\tau}{2} < \tau' < 1. \tag{58}$$

Clearly σ_l depends on the IET T we start with, since the sequence $\{n_l\}_{l\in\mathbb{N}}$ does.

LEMMA 3.9. *The sequence $\{\sigma_l\}_{l\in\mathbb{N}}$ satisfies the following properties:*

(1) $\lim_{l\to+\infty} \sigma_l = 0$;
(2) $\lim_{l\to+\infty} (\log\|A^{(n_l, n_{l+1})}\| / \log h^{(n_l)})(1/\sigma_l) = 0$;
(3) $\lim_{l\to+\infty} \sigma_l \log h^{(n_l)} = +\infty$;
(4) $\lim_{l\to+\infty} \sigma_l^2 h^{(n_{l+1})} \lambda^{(n_l)} = 0$.

Proof. Both (1) and (2) follow from Lemma 3.3. To show (3), note that $\log \|A^{(n_l, n_{l+1})}\| \geq \log d \geq 1$, so $\sigma_l \log h^{(n_l)} \geq (\log h^{(n_l)})^{1-\tau'}$ and apply Corollary 3.1. For (4), using in order balance (see Lemma 3.2), the transformation relation for heights, the definition of σ_l, the Diophantine property (26) in Proposition 3.2 and Corollary 3.1, we obtain

$$\sigma_l^2 h^{(n_{l+1})} \lambda^{(n_l)} \leq \sigma_l^2 \kappa \frac{h^{(n_{l+1})}}{h^{(n_l)}} \leq \sigma_l^2 \kappa \|A^{(n_l, n_{l+1})}\|.$$

Hence, substituting for σ_l the explicit expression in (58),

$$\lim_{l\to+\infty} \frac{\kappa (\log \|A^{(n_l, n_{l+1})}\|)^{2\tau'} \|A^{(n_l, n_{l+1})}\|}{(\log h^{(n_l)})^{2\tau'}} \leq \lim_{l\to+\infty} \text{const} \frac{(\log l)^{2\tau'} o(l^{\tau})}{l^{2\tau'}} = 0,$$

where the last limit is zero since $2\tau' > \tau$. □

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

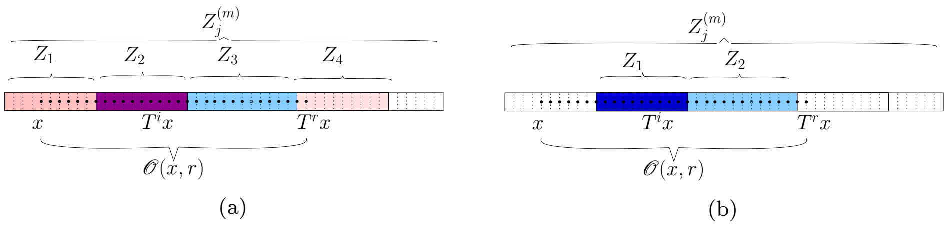

FIGURE 3. $\mathcal{O}_r(x) \prec Z_j^{(m)}$ (a) $\mathcal{O}_r(x) \prec Z_1 \wedge \cdots \wedge Z_4$, (b) $Z_1 \wedge Z_2 \prec \mathcal{O}_r(x)$.

Definition 3.2. Let $\Sigma_l^+ = \Sigma_l^+(T)$ be the following set, where $[\cdot]$ denotes the fractional part:

$$\Sigma_l^+(T) \doteqdot \bigcup_{i=0}^{[\sigma_l h^{(n_{l+1})}]} T^{-i}[0, \sigma_l \lambda^{(n_l)}]. \tag{59}$$

Note that, by property (4) of Lemma 3.9,

$$\mathrm{Leb}(\Sigma_l^+) \leq (\sigma_l h^{(n_{l+1})})(\sigma_l \lambda^{(n_l)}) \xrightarrow{l \to +\infty} 0. \tag{60}$$

PROPOSITION 3.4. (Growth of Birkhoff sums for general points) *Let $T \in \mathcal{M}^+$. For any $\varepsilon > 0$ there exists $l_o > 0$ such that for $l \geq l_o$, whenever $r \in \mathbb{N}$ and $x \in I^{(0)}$ satisfy*

$$h^{(n_l)} \leq r < h^{(n_{l+1})} \quad \text{and} \quad x \notin \Sigma_l^+(T), \tag{61}$$

we have

$$(1-\varepsilon) r \log r \leq S_r(u)(x) \leq (1+\varepsilon) r \log r + \frac{\kappa + 1}{x_m}, \tag{62}$$

where $x_m \doteqdot \min_{0 \leq i < r} T^i x$ and κ is given by Proposition 3.2.

By adding a small measure set to the excluded set Σ_l^+ of initial points, as in §4, one can also take into account the term κ / x_m and obtain the asymptotic behaviour $r \log r$ for $S_r(u)(x)$.

The following notation is used in the proof of Proposition 3.4.

3.2.2.2. Notation for approximation by towers. Denote by $\mathcal{O}_r(x)$ the orbit segment $\{T^i x, 0 \leq i < r\}$. Consider a tower $Z_j^{(m)}$. In what follows, we write

$$\mathcal{O}_r(x) \prec Z_j^{(m)} \quad \text{if and only if} \quad \exists k \mid 0 \leq k \leq h_j^{(m)} - r, \quad T^i x \in T^{k+i} I_j^{(m)}, \quad 0 \leq i < r.$$

When $\mathcal{O}_r(x) \prec Z_j^{(m)}$, each point of $\mathcal{O}_r(x)$ is contained in a different floor of $Z_j^{(m)}$ and T acts on the orbit points $T^i x$ $(0 \leq i < r-1)$ by shifting them to the next floor (see Figure 3, where the tower $Z_j^{(m)}$ is drawn horizontally).

Assume $\mathcal{O}_r(x) \prec Z_j^{(m)}$. We write $\mathcal{O}_r(x) \prec Z_1 \wedge Z_2 \wedge \cdots \wedge Z_N$, where $Z_i \in \xi_n(Z_j^{(m)})$, $n \leq m$ (see §2.2.4) for $i = 1, \ldots, N$, if $\mathcal{O}_r(x) \subset \bigcup_{i=1}^N Z_i$, $\mathcal{O}_r(x) \cap Z_i \neq \emptyset$ for $1 \leq i \leq N$ and moreover Z_i are *consecutive* partition elements of $\xi_n(Z_j^{(m)})$, i.e. if h_i denotes the height

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

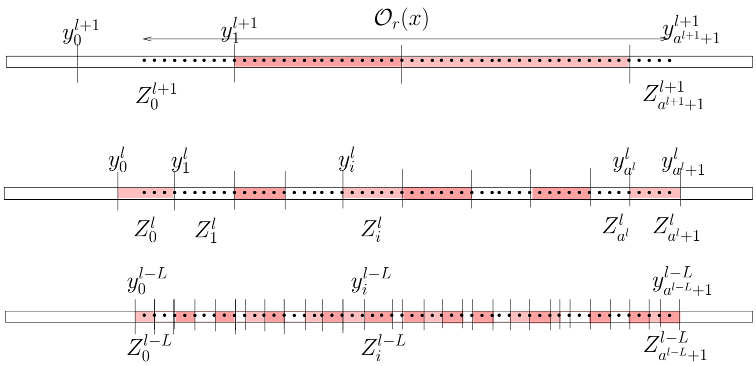

FIGURE 4. Approximation of $\mathcal{O}_r(x)$ by elements of $\xi_{n_{l+1}}$, ξ_{n_l} and $\xi_{n_{l-L}}$.

of Z_i, noting that $Z_i \cap I^{(n)}$ is the base of Z_i, we have $T^{h_i} Z_i \cap I^{(n)} = Z_{i+1} \cap I^{(n)}$ for $i = 1, \dots, N-1$ (see, e.g., Figure 3(a)).

On the other hand, we write $Z_1 \wedge Z_2 \wedge \cdots \wedge Z_N \prec \mathcal{O}_r(x)$, where $Z_i \in \xi_n(Z_j^{(m)})$, $n < m$, if $\#\{Z_i \cap \mathcal{O}_r(x)\} = h_i$ for all $i = 1 \dots, N$, i.e. there is exactly one point of $\mathcal{O}_r(x)$ in each floor of each Z_i and moreover Z_i are, as above, consecutive partition elements (see Figure 3(b)).

Proof of Proposition 3.4. It is always possible to assume that $\mathcal{O}_r(x) \prec \overline{Z} \doteqdot Z_{\bar{j}}^{(n_{\bar{l}})}$ for some $\bar{l} \geq l+1$, since, by choosing $\bar{l}$ such that $\lambda^{(n_{\bar{l}})} < x_m = \min_{0 \leq i < r} T^i x$, we assure that $\mathcal{O}_r(x) \cap I^{(n_{\bar{l}})} = \emptyset$.

3.2.2.3. Orbit decomposition into sums along towers. Let us approximate $\mathcal{O}_r(x)$ with elements of $\xi_{n_{l+1}}(\bar{Z})$ and $\xi_{n_l}(\bar{Z})$. Using the assumption $r < h^{(n_{l+1})}$, the cardinality $\#\mathcal{O}_r(x) \cap I^{(n_{l+1})}$ is bounded by $[\kappa]+1$: since the return time to $I^{(n_{l+1})}$ is at least $\min_j h^{(n_{l+1})} \geq h^{(n_{l+1})}/\kappa$ by κ-balance of heights, there cannot be more than $r/\min_j h_j^{(n_{l+1})} + 1 \leq [\kappa]+1$ returns.

Hence there exists $Z_0^{l+1}, Z_1^{l+1}, \dots, Z_{a^{l+1}+1}^{l+1} \in \xi_{n_{l+1}}(\bar{Z})$, with $a^{l+1} \leq [\kappa]$ such that (see Figure 4)

$$\mathcal{O}_r(x) \prec Z_0^{l+1} \wedge Z_1^{l+1} \wedge \cdots \wedge Z_{a^{l+1}+1}^{l+1}. \tag{63}$$

Approximating $\mathcal{O}_r(x)$ also with elements $Z_i^l \in \xi_{n_l}(\bar{Z})$ (again, see Figure 4),

$$Z_1^l \wedge Z_2^l \wedge \cdots \wedge Z_{a^l}^l \prec \mathcal{O}_r(x) \prec Z_0^l \wedge Z_1^l \wedge \cdots \wedge Z_{a^l}^l \wedge Z_{a^l+1}^l. \tag{64}$$

We still need another level of approximation. Let $L \doteqdot l_1 \bar{l} \in \mathbb{N}$ where l_1 is such that $2\kappa/d^{l_1} < \epsilon$. We can find elements $Z_i^{l-L} \in \xi_{n_{l-L}}(\bar{Z})$ such that (see Figure 4)

$$Z_1^{l-L} \wedge Z_2^{l-L} \wedge \cdots \wedge Z_{a^{l-L}}^{l-L} \prec \mathcal{O}_r(x) \prec Z_0^{l-L} \wedge Z_1^{l-L} \wedge \cdots \wedge Z_{a^{l-L}}^{l-L} \wedge Z_{a^{l-L}+1}^{l-L}. \tag{65}$$

Denote by h_i^l and h_i^{l-L} the heights of Z_i^l and Z_i^{l-L}, respectively. Note that

$$\sum_{i=1}^{a^l} h_i^l \leq \sum_{i=1}^{a^{l-L}} h_i^{l-L} \leq r \leq \sum_{i=0}^{a^{l-L}+1} h_i^{l-L} \leq \sum_{i=0}^{a^l+1} h_i^l. \tag{66}$$

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

Let us truncate $\mathcal{O}_r(x)$ into segments contained in different elements $Z_i^{l-L} \in \xi_{l-L}(\bar{Z})$. Noting that $I^{(n_{l-L})}$ contains all the bases of the towers Z_i^{l-L}, let us denote

$$y_i^{l-L} \doteqdot \mathcal{O}_r(x) \cap Z_i^{l-L} \cap I^{(n_{l-L})}, \quad i = 1, \ldots, a^{l-L}. \tag{67}$$

Since h_i^{l-L} is exactly the first return time of y_i^{l-L} to $I^{(n_{l-L})}$, $y_{i+1}^{l-L} = T^{h_i^{l-L}} y_i^{l-L} = T^{(n_{l-L})} y_i^{l-L}$. Add also the two auxiliary points:

$$y_0^{l-L} \doteqdot (T^{(n_{l-L})})^{-1} y_1^{l-L}, \quad y_{a^{l-L}+1}^{l-L} \doteqdot T^{(n_{l-L})} y_{a^{l-L}}^{l-L}. \tag{68}$$

From (65), $\bigcup_{i=1}^{a^{l-L}} \mathcal{O}_{h_i^{l-L}}(y_i^{l-L}) \subset \mathcal{O}_r(x) \subset \bigcup_{i=0}^{a^{l-L}+1} \mathcal{O}_{h_i^{l-L}}(y_i^{l-L})$.

As a consequence, since $u > 0$, we obtain the following estimate for $S_r(u)(x)$:

$$\sum_{i=1}^{a^{l-L}} S_{h_i^{l-L}}(u)(y_i^{l-L}) \leq S_r(u)(x) \leq \sum_{i=0}^{a^{l-L}+1} S_{h_i^{l-L}}(u)(y_i^{l-L}). \tag{69}$$

Each term in the summations in (69) is a Birkhoff sum along a tower of step n_{l-L}. Hence we can apply Proposition 3.3 to each term and find $l_0 \doteqdot l(\epsilon) + L$ such that for each $l \geq l_0$ we obtain

$$S_r(u)(x) \geq (1-\epsilon) \sum_{i=1}^{a^{l-L}} h_i^{l-L} \log h^{(n_{l-L})} + \sum_{i=1}^{a^{l-L}} \frac{1}{y_i^{l-L}}, \tag{70}$$

$$S_r(u)(x) \leq (1+\epsilon) \sum_{i=0}^{a^{l-L}+1} h_i^{l-L} \log h^{(n_{l-L})} + \sum_{i=0}^{a^{l-L}+1} \frac{1}{y_i^{l-L}}. \tag{71}$$

Let us refer to the first term on the right-hand side of (70) or (71) as the *ergodic term* and to the last term, i.e. the contributions of points in the bottom floors, as the *resonant term*. There is an analogy with the terminology used by [**16, 17**].

Ergodic term. Taking the ratio of the ergodic term over $r \log r$ and applying the bounds (66) for r,

$$\frac{\left(\sum_{i=1}^{a^{l-L}} h_i^{l-L}\right) \log h^{(n_{l-L})}}{r \log r} \geq \left(1 - \frac{2h^{(n_{l-L})}}{r}\right) \frac{\log h^{(n_{l-L})}}{\log r}, \tag{72}$$

$$\frac{\left(\sum_{i=0}^{a^{l-L}+1} h_i^{l-L}\right) \log h^{(n_{l-L})}}{r \log r} \leq \left(1 + \frac{2h^{(n_{l-L})}}{r}\right) \frac{\log h^{(n_{l-L})}}{\log r}. \tag{73}$$

By assumption (61) on r, property (3) of Proposition 3.2, Lemma 3.1, balance and choice of $L = l_1 \bar{l}$,

$$\frac{2h^{(n_{l-L})}}{r} \leq \frac{2h^{(n_{l-l_1\bar{l}})}}{h^{(n_l)}} \leq \frac{2\kappa}{d^{l_1}} < \epsilon.$$

Hence the first factors on the right-hand side of (72) and (73) are bounded, respectively, below by $(1-\epsilon)$ and above by $(1+\epsilon)$. The second factor on the right-hand side of (73) is trivially less than one. From $r < h^{(n_{l+1})}$,

$$\frac{\log h^{(n_{l-L})}}{\log r} \geq \frac{\log h^{(n_{l+1})} - \log h^{(n_{l+1})}/h^{(n_{l-L})}}{\log h^{(n_{l+1})}}. \tag{74}$$

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

Since by the heights transformation formula and Lemma 3.3

$$\frac{\log(h^{(n_{l+1})}/h^{(n_{l-L})})}{\log h^{(n_{l+1})}} \leq \frac{\log \|A^{(n_{l-L},n_{l+1})}\|}{\log h^{(n_{l+1})}} \xrightarrow{l\to\infty} 0,$$

also the second factor on the right-hand side of (73) is bounded from below by $(1-\epsilon)$ if $l \geq l_o$ for some $l_o \geq l_0$.

So far, combining (70) and (71) with (72), (73) and (74), and by using the fact that the resonant term is positive, we have proved that

$$(1-\epsilon)^2 r \log r \leq S_r(u)(x) \leq (1+\epsilon) r \log r + \sum_{j=0}^{a^{l-L}+1} \frac{1}{y_j^{l-L}}. \tag{75}$$

Resonant term. We want to prove the following estimate for the resonant term:

$$0 \leq \sum_{j=0}^{a^{l-L}+1} \frac{1}{y_j^{l-L}} \leq \epsilon r \log r + \frac{\kappa+1}{x_m}. \tag{76}$$

Let us first group $\{y_j^{l-L}\}_{j=0,\dots,a^{l-L}+1}$ according to visits to different elements of the partition $\xi_{n_l}(\bar{Z})$. Since

$$\bigcup_{j=0}^{a^{l-L}+1} \{y_j^{l-L}\} \subset \bigcup_{j=0}^{a^{l-L}+1} Z_j^{l-L} \subset \bigcup_{i=0}^{a^l+1} Z_i^l,$$

we have the estimate

$$\sum_{j=0}^{a^{l-L}+1} \frac{1}{y_j^{l-L}} \leq \sum_{i=0}^{a^l+1} \sum_{\substack{y_j^{l-L}\in Z_i^l \\ j=0,\dots,a^{l-L}+1}} \frac{1}{y_j^{l-L}}.$$

Each of the points $y_j^{l-L} \in Z_i^l$ belongs to a different floor of a tower of step n_l. Hence, using an argument similar to that of Lemma 3.7 for the gap error, each of the terms $\sum_{y_j^{l-L}\in Z_i^l}(1/y_j^{l-L})$ can be bounded from above by applying Lemma 3.8 to an auxiliary arithmetic progression with step $1/d\kappa\nu h_i^l$. The cardinality of points in each group, by (21), is bounded by $\#\{y_j^{l-L} \mid y_j^{l-L} \in Z_i^l\} \leq \#\{Z_j^{l-L} \in \xi_{n_{l-L}}(Z_i^l)\} \leq \|A^{(n_{l-L},n_l)}\|$. The initial point $\min\{y_j^{l-L} \mid y_j^{l-L} \in Z_i^l\}$ is given by the only visit to the base $Z_i^l \cap I^{(n_l)}$. Denote, as above (see (67) and (68)),

$$y_i^l \doteqdot \mathcal{O}_r(x) \cap Z_i^l \cap I^{(n_l)}, \quad i = 1,\dots,a^l; \quad y_0^l \doteqdot (T^{(n_l)})^{-1} y_1^l; \quad y_{a^l+1}^l \doteqdot T^{(n_l)} y_{a^l}^l.$$

Hence we obtain

$$\sum_{j=0}^{a^{l-L}+1} \frac{1}{y_j^{l-L}} \leq \sum_{i=0}^{a^l+1} d\kappa\nu h_i^l (\log \|A^{(n_{l-L},n_l)}\| + 1) + \sum_{i=0}^{a^l+1} \frac{1}{y_i^l}. \tag{77}$$

Comparing the first term on the right-hand side of (77) to $r\log r$ and recalling (66) we obtain

$$\frac{d\kappa\nu\left(\sum_{i=0}^{a^l+1} h_i^l\right)(\log \|A^{(n_{l-L},n_l)}\| + 1)}{r\log r} \leq d\kappa\nu\left(1 + \frac{2h^{(n_l)}}{r}\right)\frac{\log \|A^{(n_{l-L},n_l)}\| + 1}{\log h^{(n_l)}}$$

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

where $(1 + 2h^{(n_l)}/r) \leq 3$, so the last term, enlarging l_o if necessary, is less than ϵ when $l \geq l_o$ by Lemma 3.3 and Corollary 3.1.

The second term on the right-hand side of (77) is bounded in two different ways, according to the ratio between r and $h^{(n_{l+1})}$, using the quantity σ_l defined in (58) as a threshold.

Case 1. Assume $\sigma_l h^{(n_{l+1})} \leq r < h^{(n_{l+1})}$. Recalling (63),

$$\{y_i^l \mid i = 0, \ldots, a^l + 1\} \subset \bigcup_{j=0}^{a^{l+1}+1} \{y_i^l \mid y_i^l \in Z_j^{l+1}, i = 0, \ldots, a^l + 1\}.$$

To estimate the contribution from each of the sets on the right-hand side of (77), arguing as above, consider an auxiliary arithmetic progression of step $d\kappa\nu h_j^{l+1}$. The closest point of each set is given by the visit to $I^{(n_{l+1})}$. Noting that the number of points in each is bounded by $\#\{y_i^l \mid y_i^l \in Z_j^{l+1}, i = 0, \ldots, a^l + 1\} \leq \#\xi_{n_l}(Z_j^{l+1}) \leq \|A^{(n_l, n_{l+1})}\|$, we obtain

$$\sum_{i=0}^{a^l+1} \frac{1}{y_i^l} \leq \sum_{j=0}^{a^{l+1}+1} d\kappa\nu h_j^{l+1}(\log \|A^{(n_l, n_{l+1})}\| + 1) + \sum_{\substack{y_i^l \in I^{(n_{l+1})} \\ i=0,\ldots,a^l+1}} \frac{1}{y_i^l}. \tag{78}$$

For the first term on the right-hand side of (78), by the assumptions on r, $h_j^{l+1}/r \leq h^{(n_{l+1})}/r \leq 1/\sigma_l$ and $a^{l+1} \leq \kappa$,

$$\frac{\sum_{j=0}^{a^{l+1}} d\kappa\nu h_j^{l+1}(\log \|A^{(n_l, n_{l+1})}\| + 1)}{r \log r} \leq \frac{(\kappa + 2)d\kappa\nu(\log \|A^{(n_l, n_{l+1})}\| + 1)}{\sigma_l \log h^{(n_l)}}.$$

The last expression, again enlarging l_o if necessary, is smaller than ϵ if $l \geq l_o$ by properties (2) and (3) in Lemma 3.9.

The second term on the right-hand side of (78) can be just estimated with $(\kappa + 1)/x_m$ since, as remarked at the beginning of this proof, $\#\mathcal{O}_r(x) \cap I^{(n_{l+1})} \leq [\kappa]$. This completes the proof of (76) in this case.

Case 2. Assume $h^{(n_l)} \leq r < \sigma_l h^{(n_{l+1})}$. In this case, use the trivial estimate

$$\sum_{i=0}^{a^l+1} \frac{1}{y_i^l} \leq (a^l + 2)\frac{1}{x_m}.$$

Since $x_m = T^i x$ for some $0 \leq i < r$ and in this case $r < \sigma_l h^{(n_{l+1})}$, by the assumption (61) and the definition (59) of Σ_l^+,

$$x_m \geq \sigma_l \lambda^{(n_l)} \geq \sigma_l \frac{1}{\kappa h^{(n_l)}},$$

where the last inequality uses the balance of n_l (see Lemma 3.2). Moreover, from (66) and κ-balance of heights,

$$a^l \leq \frac{r}{\min_j h_j^{(n_l)}} \leq \frac{\kappa r}{h^{(n_l)}}.$$

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

Hence,

$$\frac{\sum_{i=0}^{a^l+1} 1/y_i^l}{r\log r} \leq \frac{(\kappa r/h^{(n_l)}+2)\kappa h^{(n_l)}/\sigma_l}{r\log r} \leq \frac{\kappa^2+2\kappa}{\sigma_l \log h^{(n_l)}},$$

which, again enlarging l_o, is smaller than ϵ for $l \geq l_o$ by property (3) in Lemma 3.9.

In both cases we have proved the estimate (76) for the resonant term. Together with (75), for an appropriate choice of ϵ, this completes the proof of Proposition 3.4. □

COROLLARY 3.2. *Let $T \in \mathcal{M}^+$. For each $\varepsilon > 0$ there exists r_0 such that, for all $r \geq r_0$, $x \in I^{(0)}$,*

$$S_r(u)(x) \leq \varepsilon r^2 + \frac{\kappa+1}{x_m}. \tag{79}$$

The estimate in the corollary is worse than (62) in Proposition 3.4, but holds for all points and is used in §4.3.

Proof. Let $h^{(n_l)} \leq r < h^{(n_{l+1})}$. Note that (75), (77) and (78) in the proof of Proposition 3.4 were obtained without using the assumption (61) and hence still hold if $r \geq r_0 \doteqdot h^{(n_{l_o})}$. Terms estimated by $r\log r$ are clearly less than εr^2 choosing r_0 large enough. The second term on the right-hand side of (78) is estimated by $(\kappa+1)/x_m$. Let us estimate the first term on the right-hand side of (78) by

$$\frac{h^{(n_{l+1})}\log\|A^{(n_l,n_{l+1})}\|}{r^2} \leq \frac{h^{(n_{l+1})}\log\|A^{(n_l,n_{l+1})}\|}{h^{(n_l)2}} \leq \frac{\|A^{(n_l,n_{l+1})}\|\log\|A^{(n_l,n_{l+1})}\|}{\text{const}\, d^l}$$

where we have used Corollary 3.1 in the last bound. The limit of this ratio as $l \to \infty$ is zero by equation (26) (property (4) of Proposition 3.2). □

3.3. *Growth of Birkhoff sums of the derivatives.* Let $h_{\mathcal{I}}^{(n)}$ and $\lambda_{\mathcal{I}}^{(n)}$ be the sequences of heights and lengths of towers for $T^{\mathcal{I}}$. For $T^{\mathcal{I}} \in \mathcal{M}^+$, let $\{n'_{l'}\}_{l'\in\mathbb{N}}$ be the sequence of balanced times for $T^{\mathcal{I}}$ given by Proposition 3.2. Let $\sigma_{l'} = \sigma_{l'}(T^{\mathcal{I}})$. We define

$$\Sigma_{l'}^-(T) \doteqdot \bigcup_{i=0}^{[\sigma_{l'} h_{\mathcal{I}}^{(n'_{l'+1})}]} T^{-i}[1-\sigma_{l'}\lambda_{\mathcal{I}}^{(n'_{l'})}, 1). \tag{80}$$

COROLLARY 3.3. (Growth of $S_r(v)$) *Let $T \in \mathcal{M}^-$. For any $\epsilon > 0$ there exists $l'_o > 0$ such that, if $l' \geq l'_o$, for any $r \in \mathbb{N}$ and $x \in I^{(0)}$ such that*

$$h_{\mathcal{I}}^{(n'_{l'})} < r < h_{\mathcal{I}}^{(n'_{l'+1})} \quad \text{and} \quad x \notin \Sigma_{l'}^-(T),$$

denoting by $x_M \doteqdot \max_{0\leq i<r} T^i x$, we obtain

$$(1-\epsilon)r\log r \leq S_r(v)(x) \leq (1+\epsilon)r\log r + \frac{\kappa'+1}{1-x_M},$$

where κ' is the same as that given by Proposition 3.2.

Proof. Corollary 3.3 is simply obtained by restating Proposition 3.4 for u and $T^{\mathcal{I}}$ and using the relation with v given by (22). Note that $\min_i (T^{\mathcal{I}})^i(1-x) = \min_i \mathcal{I}T^i(x) = \min_i(1-T^i x) = 1-x_M$ and that $(1-x) \in \Sigma_{l'}^+(T^{\mathcal{I}})$ if and only if $x \in \Sigma_{l'}^-(T)$. □

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

COROLLARY 3.4. (Growth $S_r(f')$) *For each $T \in \mathcal{M}$ and $C^+ \neq C^-$ there exist $C_1, C_1', C_2, C_3 > 0$ and r_o such that, for $r \geq r_o$, if*

$$\begin{array}{l} h^{(n_l)} \leq r < h^{(n_{l+1})}, \\ h_{\mathcal{I}}^{(n'_{l'})} \leq r < h_{\mathcal{I}}^{(n'_{l'+1})} \end{array} \quad \text{and} \quad x \notin \Sigma_l^+(T) \cup \Sigma_{l'}^-(T), \tag{81}$$

and x is not a singularity of $S_r(f)$,

$$S_r(f')(x) \leq -C_1 r \log r + \frac{C^+(\kappa'+1)}{1-x_M} \quad \text{if } C^+ > C^-, \tag{82}$$

$$S_r(f')(x) \geq C_1 r \log r - \frac{C^-(\kappa+1)}{x_m} \quad \text{if } C^+ < C^-, \tag{83}$$

$$|S_r(f')(x)| \leq C_2 r \log r + \frac{C_2(\kappa+1)}{x_m} + \frac{C_2(\kappa'+1)}{1-x_M}, \tag{84}$$

where x_m, x_M, κ and κ' are as in Corollary 3.3 and Proposition 3.4.

Moreover, for all $x \in I^{(0)}$ different from singularities of $S_r(f)$,

$$|S_r(f')(x)| \leq C_3 r^2 + \frac{C_3(\kappa+1)}{x_m} + \frac{C_3(\kappa'+1)}{1-x_M}. \tag{85}$$

Proof. Assume $C^+ > C^-$. Consider the sequence $\alpha_r \to 0$ in Proposition 3.1. One can choose r_1 so that, for each $r \geq r_1$, we have $(C^+ - \alpha_r) > (C^- + \alpha_r)$. Hence it is also possible to choose $\epsilon > 0$ so that, for $r \geq r_1$, $C_1 \doteqdot (C^+ - \alpha_r)(1-\epsilon) - (C^- + \alpha_r)(1+\epsilon) > 0$. By Proposition 3.1, Proposition 3.4 and Corollary 3.3, which can be applied by the assumptions (81) when $r \geq r_o$, $r_o \doteqdot \max\{r_1, h^{(n_{l_o})}, h_{\mathcal{I}}^{(n'_{l'_o})}\}$,

$$\begin{aligned} S_r(f')(x) &\leq -(C^+ - \alpha_r) S_r(u)(x) + (C^- + \alpha_r) S_r(v)(x) \\ &\leq -(C^+ - \alpha_r)(1-\epsilon) r \log r + (C^- + \alpha_r)(1+\epsilon) r \log r + \frac{(C^- + \alpha_r)(\kappa'+1)}{1-x_M} \\ &\leq -C_1 r \log r + \frac{C^+(\kappa'+1)}{1-x_M}. \end{aligned}$$

The case $C^+ < C^-$ can be treated analogously.

Also (84) follows similarly: enlarging r_1 so that if $r \geq r_1$, $\alpha_r \leq \min(C^+, C^-)$, by Proposition 3.1,

$$|S_r(f')(x)| \leq (C^+ + C^-)(S_r(u)(x) + S_r(v)(x)) \tag{86}$$

and $S_r(u)$, $S_r(v)$ can again be estimated by Proposition 3.4 and Corollary 3.3.

For (85), apply to (86) the rough estimate on $S_x(u)$ for all points given by Corollary 3.2 and the analogous one for $S_x(v)$ which follows from $T \in \mathcal{M}^-$. □

COROLLARY 3.5. (Growth $S_r(f'')$) *For each $T \in \mathcal{M}$ and $C^+ \neq C^-$, there exist $C_4 > 0$ and r_o such that for $r \geq r_o$, if (81) holds,*

$$|S_r(f'')(x)| \leq C_4 \max\left\{\frac{1}{x_m}, \frac{1}{1-x_M}\right\}\left(r \log r + \frac{\kappa+1}{x_m} + \frac{\kappa'+1}{1-x_M}\right).$$

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

Proof. By definition of logarithmic singularity, there exists $\delta > 0$ such that $0 \le f''(y) \le 2C^+/y^2$ if $y < \delta$ and $0 \le f''(y) \le 2C^-/(1-y)^2$ if $y > 1-\delta$. Let $M_{f''}$ be the maximum of $|f''|$ on $[\delta, 1-\delta]$. Hence, for each x that is not a singularity of $S_r(f)$,

$$\begin{aligned}|S_r(f'')(x)| &\le 2C^+ S_r(1/x^2)(x) + 2C^- S_r(1/(1-x)^2)(x) + rM_{f''} \\ &\le 2C^+ \frac{1}{x_m} S_r(u)(x) + 2C^- \frac{1}{1-x_M} S_r(v)(x) + rM_{f''}. \qquad (87)\end{aligned}$$

Applying Proposition 3.4 and Corollary 3.3 one obtains the desired estimate. □

4. *Construction of the mixing partitions*

In this section we construct partitions $\eta_m(t)$ that verify the mixing criterion (see Lemma 1.1). The construction is carried out in three main steps, formulated in §4.1 as Propositions 4.1, 4.2 and 4.3. Their proofs are given in §§4.3, 4.4 and 4.5, respectively. We anticipate in §4.2 the final area estimates, which conclude the proof of Theorem 1.1.

4.1. *Partitions properties.*

Denoting by $[\cdot]$ the integer part, let

$$R_M(t) \doteqdot [t/m_f] + 2. \qquad (88)$$

PROPOSITION 4.1. (Preliminary partitions) *For each $0 < \delta < 1$ and $M > 1$, there exist $t_0 > 0$ and partial partitions $\eta_p(t)$ for $t \ge t_0$, such that* $\mathrm{Leb}(\eta_p(t)) > 1-\delta$ *and the following properties hold for each $I = [a, b) \in \eta_p(t)$.*

(1) *Continuity intervals:*

$$T^j \text{ is continuous on } \bar{I} = [a, b] \quad \text{for each } 0 \le j \le R_M(t).$$

(2) *Control of interval sizes:*

$$\frac{1}{t \log\log t} \le \mathrm{Leb}(I) \le \frac{2}{t \log\log t}.$$

(3) *Control of the distance from singularities:*

$$\mathrm{dist}(T^j I, 0) \ge \frac{M}{t \log\log t}, \quad \mathrm{dist}(T^j I, 1) \ge \frac{M}{t \log\log t}, \quad 0 \le j \le R_M(t).$$

(4) *Control of the number of discrete iterations:*

$$\frac{t}{3} \le r(x, t) \le R_M(t) \le \frac{2}{m_f} t \quad \text{for all } x \in I.$$

We note that the function $t \log\log t$ in Properties (2) and (3) of Proposition 4.1 is chosen for convenience, but could be replaced by any expression $t\alpha(t)$ where $\alpha(t)$ is a positive function such that $\lim_{t\to+\infty} \alpha(t) = +\infty$ and which we choose so that $\lim_{t\to+\infty} \alpha(t)/(\log t)^\nu = 0$ for some $0 < \nu < 1$.

Assume now that $T \in \mathcal{M}$. For definiteness, assume also that the asymmetry constants of the roof function satisfy $C^+ > C^-$. Using the estimates on the growth of Birkhoff sums obtained in §3, we can refine the partitions $\eta_p(t)$ to obtain the following.

PROPOSITION 4.2. (Stretching partitions) *For $T \in \mathcal{M}$ and $C^+ > C^-$, there exist $C_1', C_2', C'' > 0$ such that for each $0 < \delta < 1$, $M > 1$, if $\eta_p(t)$ are the corresponding partitions in Proposition 4.1, there exists a one-parameter family of refined partitions $\eta_s(t) \subset \eta_p(t)$ with $\mathrm{Leb}(\eta_s) > \mathrm{Leb}(\eta_p) - \delta$ and there exists $t_1 > t_0$ such that, when $t \geq t_1$, for any $x \in \eta_s(t)$ and integer r with $t/3 \leq r \leq 2t/m_f$,*

$$S_r(f')(x) \leq -C_1' r \log r, \tag{89}$$

$$|S_r(f')(x)| \leq C_2' r \log r, \tag{90}$$

$$S_r(f'')(x) \leq \frac{C''}{M} r^2 (\log r)(\log\log r). \tag{91}$$

Let us show that Proposition 4.2 implies in particular that $r(\cdot, t)$ is an increasing function on each interval $I \in \eta_s(t)$ for $t \geq t_1$. Assume $x < y$ are points of I. Since in particular $S_{r(x,t)}(f') < 0$, the function $S_{r(x,t)}(f)$ is strictly decreasing. Hence $S_{r(x,t)}(f)(y) < S_{r(x,t)}(f)(x) \leq t$. Using again the definition of $r(\cdot, t)$, we obtain $r(y, t) \geq r(x, t)$.

4.1.1. *Geometric description of the dynamics of partition elements.* Let $I = [a, b)$ be an element of the partition $\eta_s(t)$. Consider $\varphi_t(I)$ and let us give first a geometric description of $\varphi_t(I)$ for $t \gg 1$. For $C^+ > C^-$, as just proved, $r(\cdot, t)$ is an increasing function on each $[a, b) \in \eta_p(t)$. Hence, let

$$r(a) = r(a, t) = \min_{x \in [a,b)} r(x, t), \qquad r(b) = r(b, t) = \max_{x \in [a,b)} r(x, t),$$

$$J = J([a, b), t) \doteqdot r(b, t) - r(a, t) + 1. \tag{92}$$

The dependence on t will be omitted when t is clear from the context.

The image $\varphi_t(I)$ splits into several curves and J gives exactly their cardinality. More precisely, consider the equation $S_r(f) = t$ on I, which has a solution exactly for $r = r(a) + 1, \ldots, r(b)$, unique by monotonicity. Denote by y_j the solution of

$$S_{r(a)+j}(f)(y_j) = t, \quad j = 1, \ldots, J - 1, \tag{93}$$

so that by equation (3), $r(y_j, t) = r(a) + j$ and $\varphi_t(y_j, 0) = (T^{r(y_j,t)}(y_j), 0)$. The points $a \doteqdot y_0 \leq y_1 \leq \cdots \leq y_j \leq y_{j+1} \leq y_{J-1} \leq y_J \doteqdot b$ are splitting points, meaning that the image $\varphi_t(I)$ consists of J curves, which are the graphs of $t - S_{r(a)+j}(f)$ restricted to $I_j \doteqdot [y_j, y_{j+1})$, $j = 0, \ldots, J-1$. Each curve projects to $T^{r(a)+j}(I_j)$, by (3). In particular, they project to the orbit $T^{r(a)+j}(I)$, for $j = 0, \ldots, J - 1$.

4.1.2. *Properties of the mixing partitions.* Given $\delta > 0$, if $[b_1, b_2]$ is the base of the rectangle R, denote by χ the indicator of $[b_1 + \delta, b_2 - \delta]$. Choose $t_2 \geq t_1$ so that $2/(t_2 \log\log t_2) < \delta$ and the mesh of the partitions $\eta_p(t)$ for $t \geq t_2$ is bounded by δ by property (2) in Proposition 4.1. Denoting by $h(R)$ the height of the rectangle R, for $j = 0, \ldots, J - 1$, let

$$I_j^{h(R)} \doteqdot \{x \mid t - h(R) \leq S_{r(a)+j}(f)(x) \leq t\}. \tag{94}$$

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

Points in $I_j^{h(R)}$ are those that reach the correct height to intersect R, i.e. if $x \in I_j^{h(R)}$, then $\varphi_t(x, 0)$ is contained in the horizontal strip $I^{(0)} \times h(R)$ (as shown in the proof of Lemma 4.1). For $j = 0, \dots, J - 1$, denote

$$\Delta f^j = \Delta f^j([a, b), t) \doteqdot S_{r(a)+j}(f)(a) - S_{r(a)+j}(f)(b), \quad \Delta f \doteqdot \Delta f^0.$$

Note that $\Delta f^j \geq 0$. The quantity Δf expresses the delay accumulated between the endpoints in time t. Also, the quantity Δf^j gives the vertical stretch of the graph of $t - S_{r(a)+j}(f)|_{[a,b)}$.

PROPOSITION 4.3. (Mixing partitions) *Let $T \in \mathcal{M}^+$, $C^+ > C^-$. Given $\epsilon > 0$ and $0 < \delta < 1$, there exist $M(\epsilon)$, $\bar{t} > t_2$ and refined partial partitions $\eta_m(t) \subset \eta_s(t)$, where $\eta_s(t)$ are the partitions given by Proposition 4.2, such that* $\mathrm{Leb}(\eta_m(t)) > \mathrm{Leb}(\eta_s(t)) - 2\delta$ *and for each $I = [a, b) \in \eta_m(t)$, $J(I, t) \to +\infty$ as $t \to +\infty$ and for $t \geq \bar{t}$ the following properties hold.*

(1) *Uniform vertical distribution:*

$$\left| \frac{h(R)(b-a)}{\Delta f^j(I,t)} - \mathrm{Leb}(I_j^{h(R)}) \right| \leq \frac{h(R)(b-a)}{\Delta f^j(I,t)}\epsilon, \quad j = 1, \dots, J(I,t) - 2. \tag{95}$$

(2) *Variation of slopes:*

$$\left| \frac{\Delta f(I,t)}{\Delta f^j(I,t)} - 1 \right| \leq \epsilon, \quad j = 0, \dots, J(I,t) - 2. \tag{96}$$

(3) *Asymptotic number of curves:*

$$\left| \frac{J([a,b),t) - 1}{\Delta f(I,t)} - 1 \right| \leq \epsilon. \tag{97}$$

(4) *Equidistribution on the base: for some $\bar{x} \in I$,*

$$\left| \frac{1}{J([a,b),t) - 1} \left(\sum_{j=0}^{J([a,b),t)-2} \chi(T^{r(a)+j}(\bar{x})) \right) - (b_2 - b_1 - 2\delta) \right| \leq \epsilon. \tag{98}$$

4.2. *Area estimates.* Let us show that the properties in Proposition 4.3 are enough to deduce the estimate (8) of the mixing criterion (Lemma 1.1) and hence conclude the proof of Theorem 1.1.

LEMMA 4.1. *For each $I = [a, b) \in \eta_m(t)$, $t \geq \bar{t}$ and $x \in I$,*

$$\mathrm{Leb}([a,b) \cap \varphi_{-t}(R)) \geq \sum_{j=1}^{J([a,b),t)-2} \chi(T^{r(a)+j}x)\, \mathrm{Leb}(I_j^{h(R)}). \tag{99}$$

Proof. Let us first show that $I_j^{h(R)}$, $j \in \mathbb{N}$, are all disjoint. If $y \in I_j^{h(R)}$, then for each integer $s > 0$, $S_{r(a)+j+s}(f)(y) = S_{r(a)+j}(f)(y) + S_s(f)(T^{r(a)+j}y) > S_{r(a)+j}(f)(y) + h(R)$ since $h(R) < m_f$. Hence, by (94), $S_{r(a)+j+s}(f)(y) > t$ and $y \notin I_{j+s}^{h(R)}$. A similar argument works for $s < 0$.

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

Assume that j_0 is such that $1 \le j_0 \le J([a,b),t) - 2$ and $\chi(T^{r(a)+j_0}x) = 1$, i.e. $T^{r(a)+j_0}x \in [b_1+\delta, b_2-\delta]$. Recalling that t_2 was chosen so that $\sup_{I \in \eta_m(t)} \mathrm{Leb}(I) < \delta$ for $t \ge t_2$, when $t \ge \bar{t} \ge t_2$, we have $T^{r(a)+j_0}I \subset [b_1, b_2]$.

It is sufficient to show that $I^{h(R)}_{j_0} \subset [a,b) \cap \varphi_{-t}(R)$ to reach our conclusion. If $y \in I^{h(R)}_{j_0}$, by definition $t - h(R) \le S_{r(a)+j_0}(f)(y) \le t$. Also, recalling that $h(R) < m_f$,

$$t \le t - h(R) + f(T^{r(a)+j_0}y) < S_{r(a)+j_0+1}(f)(y),$$

which shows that $r(y,t) = r(a) + j_0$ and also that $y \in [y_{j_0}, y_{j_0+1}] \subset [a,b)$, by monotonicity and definition of splitting points (93).

It follows, by definition (94) of $I^{h(R)}_{j_0}$ and (3) of the flow action, that

$$\varphi_t(y,0) = (T^{r(a)+j_0}y, t - S_{r(a)+j_0}(f)(y)) \in [b_1,b_2] \times [0, h(R)] = R.$$

This shows that $I^{h(R)}_{j_0} \subset [a,b) \cap \varphi_{-t}(R)$. □

Let us estimate the right-hand side of (99). For $t \ge \bar{t}$,

$$\sum_{j=1}^{J([a,b))-2} \chi(T^{r(a)+j}(x))\,\mathrm{Leb}(I^{h(R)}_j)$$

$$\ge (1-\epsilon)h(R)(b-a) \sum_{j=1}^{J([a,b),t)-2} \chi(T^{r(a)+j}(x)) \frac{1}{\Delta f^j} \tag{100}$$

$$\ge (1-\epsilon)^2 h(R)(b-a) \sum_{j=1}^{J([a,b))-2} \chi(T^{r(a)+j}(x)) \frac{1}{\Delta f} \tag{101}$$

$$\ge (1-\epsilon)^3 h(R)(b-a) \sum_{j=1}^{J([a,b))-2} \frac{\chi(T^{r(a)+j}(x))}{J([a,b),t)-1} \tag{102}$$

$$\ge (1-\epsilon)^3 h(R)(b-a)(b_2 - b_1 - 2\delta - 2\epsilon) \xrightarrow{\epsilon,\delta \to 0} \mu(R)(b-a). \tag{103}$$

We used, in order, the following properties of Proposition 4.3: property (1) to obtain (100), property (2) to obtain (101), property (3) to obtain (102) and, finally, to obtain (103) we combined property (4) with $\chi(T^{r(a)}x)/(J(I,t)-1) \le \epsilon$ for $t \ge \bar{t}$ if $\bar{t}$ is enlarged if necessary, since $J(I,t)$ tends to infinity.

When ϵ and δ are chosen sufficiently small, together with the Lemma 4.1, this concludes the proof of (8). From Lemma 1.1, we obtain Theorem 1.1.

4.3. *Preliminary partitions.* Let us prove Proposition 4.1. Consider a fixed continuous time t. The maximum number of discrete iterations of T when flowing by t, i.e. $r_M(t) \doteqdot \sup_{x \in I^{(0)}} r(x,t)$, can be bounded from above for each x by using that $f \ge m_f > 0$ and the definition of $r(x,t)$. We obtain

$$r(x,t)m_f \le S_{r(x,t)}(f)(x) \le t.$$

Recalling the definition (88), we obtain $r_M(t) + 1 \le R_M(t)$.

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

4.3.1. *Continuity intervals of controlled size.* It is easy to see that any iterate $T^n \doteqdot T\cdot\cdots\cdot T$ obtained iterating T n times is again an IET: denoting by $\beta_0 = 0 < \beta_1 < \cdots < \beta_{d-1} < 1$ the discontinuities of T, the discontinuities of T^N are

$$\{T^{-j}\beta_i \mid i = 0, \ldots, d-1;\ 0 \le j < N\}. \tag{104}$$

Note that T^N is an exchange of at most $Nd + 1$ intervals.

Let $\eta_0(t)$ be the *partition* of $I^{(0)}$ *into continuity intervals* for $T^{R_M(t)}$, i.e. the partition into semi-open intervals whose endpoints coincide with the set (104) where $N = R_M(t)$. By construction, for each $0 \le j \le R_M(t)$, T^j restricted to any $[a, b) \in \eta_0(t)$ is continuous.

Given $M > 1$, consider the set

$$U_1 \doteqdot \bigcup_{\substack{0\le i\le d\\ 0\le j\le R_M(t)}} \overline{\mathrm{Ball}}\left(T^{-j}\beta_i, \frac{2M}{t\log\log t}\right),$$

which consists of closed balls of radius $2M/t\log\log t$ centered at the endpoints of $\eta_0(t)$. Let $\eta_1(t)$ be the partial partition obtained from $\eta_0(t)$ by throwing away all intervals completely contained in U_1. Since, using (88),

$$\mathrm{Leb}(U_1) \le \frac{4M}{t\log\log t} d\left(\frac{t}{m_f} + 3\right) \xrightarrow{t\to+\infty} 0, \tag{105}$$

it follows that $\mathrm{Leb}(\eta_1(t)) \ge 1 - \mathrm{Leb}(U_1)$ converges to one. Moreover, by construction, each $I \in \eta_1(t)$ contains at least one $y \notin U_1$. Hence, since the endpoints of I are centers of the balls in U_1, $\mathrm{Leb}(I) \ge 4M/t\log\log t$.

4.3.2. *Distance from singularities.* Let

$$U_2 \doteqdot \bigcup_{0\le j\le R_M(t)} T^{-j}\left[0, \frac{M}{t\log\log t}\right) \cup \bigcup_{0\le j\le R_M(t)} T^{-j}\left[1 - \frac{M}{t\log\log t}, 1\right).$$

Let $\eta_2(t) = \eta_1(t)\backslash U_2$. By construction, if $x \in \eta_2(t)$,

$$\mathrm{dist}(T^s x, 0) \ge \frac{M}{t\log\log t}, \quad \mathrm{dist}(T^s x, 1) \ge \frac{M}{t\log\log t}, \quad 0 \le s \le R_M(t),$$

which is property (2) of Proposition 4.1. Similarly to (105), also $\mathrm{Leb}(U_2) \xrightarrow{t\to\infty} 0$. Given $\delta > 0$, choose t_0 so that $\mathrm{Leb}(\eta_2(t)) \ge \mathrm{Leb}(\eta_1(t)) - \mathrm{Leb}(U_2) > 1 - \delta/2$ for $t \ge t_0$. Intervals $I \in \eta_2(t)$ are either intervals of $\eta_1(t)$ or are obtained by some $I' \in \eta_1(t)$ by cutting an interval of length at most $M/(t\log\log t)$ on one or both sides of I'. Hence, $\mathrm{Leb}(I') > 2M/t\log\log t$.

Let $\widetilde{\eta}_2(t)$ be the collection of intervals of the form $[a, b') \subset [a, b)$ associated to each $[a, b) \in \eta_2(t)$. Choosing each b' close enough to each b, one still has $\mathrm{Leb}([a, b')) > 2M/t\log\log t$ and $\mathrm{Leb}(\widetilde{\eta}_2(t)) > 1 - \delta/2$. Since T^j, for $0 \le j \le R_\mathrm{M}(t)$, is continuous on $[a, b']$, property (1) of Proposition 4.1 holds for $[a, b') \in \widetilde{\eta}_2(t)$.

Construct $\eta_3(t)$ from $\widetilde{\eta}_2(t)$ by cutting each of the intervals $I \in \widetilde{\eta}_2(t)$ in pieces which satisfy the length control property (property (2)) of Proposition 4.1. For example, cut first $[\mathrm{Leb}(I)/(1/t\log\log t)] - 1$ intervals of length exactly $1/(t\log\log t)$ starting from the left, so that the last remaining interval has length at most $2/t\log\log t$.

Properties (1) and (3) still hold and $\mathrm{Leb}(\eta_3(t)) = \mathrm{Leb}(\widetilde{\eta}_2(t)) > 1 - \delta/2$ for $t \ge t_0$.

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

4.3.3. *Control of the number of discrete iterations.* Let us bound $r(x,t)$ from below when $x \in \eta_3(t)$. As a consequence of property (3),

$$f(T^j x) \leq \mathrm{const} \log(t \log\log t), \quad 0 \leq j \leq R_M(t). \tag{106}$$

Hence $S_{r(x,t)+1}(f)(x) \leq (r(x,t)+1) \mathrm{const} \log(t \log\log t)$ and since, by definition of $r(x,t)$, we have $S_{r(x,t)+1}(f)(x) > t$,

$$r(x,t) \geq \frac{t}{\mathrm{const} \log(t \log\log t)} - 1 \xrightarrow{t \to +\infty} +\infty, \tag{107}$$

uniformly for all $x \in \eta_3(t)$. Since $f \in L^1$ and T is ergodic, by the Birkhoff ergodic theorem, for each $\delta > 0$ there exist a measurable set E_δ and $N_\delta > 0$ such that $\mathrm{Leb}(E_\delta) < \delta/2$ and

$$\left| \frac{1}{r} S_r(f)(x) - \int f(s)\, ds \right| < 1 \quad \text{for all } x \notin E_\delta,\ r \geq N_\delta.$$

Define a refined partial partition $\eta_4(t) \doteqdot \eta_3(t) \backslash \{I \in \eta_3(t) \mid I \subset E_\delta\}$. For $t \geq t_0$, $\mathrm{Leb}(\eta_4(t)) \geq \mathrm{Leb}(\eta_3(t)) - \delta/2 \geq 1-\delta$. By construction for each $I \in \eta_4(t)$ there is at least one $x_I \in I$ such that $|(1/r)S_r(f)(x_I) - 1| < 1$ for all $r \geq N_\delta$. Enlarging t_0 if necessary, by (107) we can assure $r(x_I,t) > N_\delta$ for each x_I, $I \in \eta_4(t)$. Hence $S_{r(x_I,t)+1}(f)(x_I) < 2(r(x_I,t)+1)$, which, together with $S_{r(x_I,t)+1}(f)(x_I) > t$, gives

$$r(x_I,t) > t/2 - 1 \quad \text{for all } x_I,\ I \in \eta_4(t). \tag{108}$$

To control all other $r(x,t)$, $x \in \eta_4(t)$, let us estimate the variation $r(x,t) - r(x_I,t)$ when $x \in I \in \eta_4(t)$. Assume $r(x,t) < r(x_I,t)$, otherwise we already have the lower bound. By Properties (1) and (3), $S_r(f)$ is continuous on I and $(T^r)' = 1$ for $0 \leq r \leq R_M$, so $(S_r(f)(x))' = S_r(f')(x)$. By the mean value theorem there exists z between x_I and x such that

$$|S_{r(x,t)}(f)(x_I) - S_{r(x,t)}(f)(x)| \leq |S_{r(x,t)}(f')(z)||x - x_I|.$$

Apply the rough bound on $S_r(f')$ in Corollary 3.2, enlarging t_0 again by (107) so that $r(x,t) \geq r_0$ for $t \geq t_0$. Combining it with property (3) already proved, which gives $1/x_m \leq t \log\log t/M$ and $1/(1-x_M) \leq t \log\log t/M$, we find that $|S_{r(x,t)}(f')(z)| \leq \mathrm{const}\, t^2$ for $t \geq t_0$. Since $\mathrm{Leb}(I) \leq 2/(t \log\log t)$,

$$|S_{r(x,t)}(f)(x_I) - S_{r(x,t)}(f)(x)| \leq \frac{\mathrm{const}\, t}{\log\log t}. \tag{109}$$

Hence, using (109) and $S_{r(x_I,t)}(f)(x_I) \leq t$ and then $S_{r(x,t)}(f)(x) > t - f(T^{r(x,t)}x)$ and (106),

$$\begin{aligned}(r(x_I,t) - r(x,t))m_f &\leq S_{r(x_I,t)}(f)(x_I) - S_{r(x,t)}(f)(x_I) \\ &\leq t - S_{r(x,t)}(f)(x) + \frac{\mathrm{const}\, t}{\log\log t} \leq \mathrm{const} \log(t \log t \log t) + \frac{\mathrm{const}\, t}{\log\log t} = o(t).\end{aligned}$$

Re-arranging and using the control for x_I given by (108), $r(x,t) \geq r(x_I,t) - o(t) \geq t/2 - 1 - o(t)$. Hence, recalling also $r(x,t) \leq R_{\mathrm{M}}(t) \leq t/m_f + 2$, if t_0 is sufficiently large, when $t \geq t_0$, for each $x \in I \in \eta_4(t)$, $t/3 \leq r(x,t) \leq R_{\mathrm{M}}(t) \leq 2t/m_f$, which is property (4). Since the other properties still hold, setting $\eta_p(t) \doteqdot \eta_4(t)$ proves Proposition 4.1.

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

4.4. *Stretching partitions.* Let us prove Proposition 4.2.

Let $T \in \mathcal{M}$. For each t, let $l(t)$ and $l'(t)$ be uniquely determined by

$$h^{(n_{l(t)})} \le R_M(t) < h^{(n_{l(t)+1})}, \quad h_{\mathcal{I}}^{(n'_{l'(t)})} \le R_M(t) < h_{\mathcal{I}}^{(n'_{l'(t)+1})}, \tag{110}$$

where $\{n_l\}_{l\in\mathbb{N}}$ and $\{n'_{l'}\}_{l'\in\mathbb{N}}$ are the sequences of balanced times given by Proposition 3.2 for T and $T^{\mathcal{I}}$, respectively.

LEMMA 4.2. *There exists $L \in \mathbb{N}$ independent of t such that if $t/3 \le r \le R_M(t)$ then*

$$h^{(n_{l(t)-L})} \le r < h^{(n_{l(t)+1})}, \quad h_{\mathcal{I}}^{(n'_{l'(t)-L})} \le r < h_{\mathcal{I}}^{(n'_{l'(t)+1})}. \tag{111}$$

Proof. Let $l \in \mathbb{N}$ be such that $d^l \ge \max\{6\kappa/m_f, 6\kappa'/m_f\}$ (recall that κ and $\kappa'$are given by Proposition 3.2 and Corollary 3.3).

By property (25) in Proposition 3.2, we can apply Lemma 3.1 considering products of positive matrices that appear every $\bar{l}$ balanced steps and obtain, recalling the choice of l, by balance of the induction steps and (110), $h^{(n_{l(t)-l\bar{l}})} \le (\kappa/d^l)\min_j h_j^{(n_{l(t)})} \le (\kappa/d^l)(2t/m_f) \le t/3 \le r$. Analogous expressions can also be obtained for $h_{\mathcal{I}}^{(n'_{l'(t)-l\bar{l}})}$ and show that setting $L \doteqdot l\bar{l}+1$ we obtain (111). □

Define the set $\Sigma_t = \Sigma_t(T)$ as

$$\Sigma_t \doteqdot \bigcup_{l=l(t)-L}^{l(t)-1} \Sigma_l^+(T) \cup \bigcup_{l=l'(t)-L}^{l'(t)-1} \Sigma_{l'}^-(T) \cup \bar{\Sigma}_{l(t)}^+(T) \cup \bar{\Sigma}_{l'(t)}^-(T), \tag{112}$$

where the sets $\Sigma_l^+(T)$ and $\Sigma_{l'}^-(T)$ were defined in (59) and (80) and where

$$\bar{\Sigma}_{l(t)}^+(T) \doteqdot \bigcup_{i=0}^{\min\{R_M(t),[\sigma_{l(t)}h^{(n_{l(t)+1})}]\}} T^{-i}[0, \sigma_{l(t)}\lambda^{(n_{l(t)})}]$$

and $\bar{\Sigma}_{l'(t)}^-$ is the analogous truncation of $\Sigma_{l'(t)}^-$. Note also that we have

$$\mathrm{Leb}(\Sigma_t) \xrightarrow{t\to+\infty} 0. \tag{113}$$

Equation (113) follows from (60) and from the analogous property for each $\mathrm{Leb}(\Sigma_{l'}^-)$ (note that $l(t)$ and $l'(t) \to +\infty$ as $t \to +\infty$), and from the fact that Σ_t is contained in a union of at most $2(L+1)$ sets which are either Σ_l^+ with $l \ge l(t)-L$ or $\Sigma_{l'}^-$ with $l' \ge l'(t)-L$.

Proof of Proposition 4.2. Fix T, $C^+ > C^-$, $\delta > 0$ and $M > 1$ and let $\eta_p(t)$ be the preliminary partitions given by Proposition 4.1 for $t \ge t_0$. Consider the set $\Sigma_t(T)$ defined in (112) and, by (113), choose $t_1 \ge t_0$ so that $\mathrm{Leb}(\Sigma_t) < \delta/2$ for $t \ge t_1$. Define $\eta_s(t)$ as the partition obtained from $\eta_p(t)$ throwing away all the intervals which intersect Σ_t. If $I \in \eta_p(t)$ and $I \cap \Sigma_t \neq \emptyset$, then, from property (1) in Proposition 4.1, either $I \subset \Sigma_t$ or, for some $0 \le j \le R_M(t)$, T^jI contains either some points $\sigma_l\lambda^{(n_l)}$ with $l(t)-L \le l \le l(t)$

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

or some $1-\sigma_{l'}\lambda_{\mathcal{I}}^{(n_{l'})}$ with $l'(t)-L\le l'\le l'(t)$. Hence, using (113), property (2) in Proposition 4.1 and bounding the number of such points, we obtain

$$\mathrm{Leb}(\eta_s(t))\ge \mathrm{Leb}(\eta_p(t))-\mathrm{Leb}(\Sigma_t)-\frac{2}{t\log\log t}2(L+1)\frac{2t}{m_f}$$

and, enlarging t_1 if necessary, both the last two terms in the previous equation are less than $\delta/2$.

Let $t/3\le r\le 2/m_f$ and $x\in\eta_s(t)$. Let us show that the assumptions of Corollary 3.4 and Corollary 3.5 on the growth of $S_r(f')$ and $S_r(f'')$ hold. By Lemma 4.2, there exist l, l', with $l(t)-L\le l\le l(t)$ and $l'(t)-L\le l'\le l'(t)$ such that $h^{(n_l)}\le r<h^{(n_{l+1})}$ and $h_{\mathcal{I}}^{(n'_{l'})}\le r<h_{\mathcal{I}}^{(n'_{l'+1})}$. Since by construction of $\eta_s(t)$, $x\notin\Sigma_t$, in particular, if $l<l(t)$ and $l'<l'(t)$, $x\notin\Sigma_l^+$, $x\notin\Sigma_{l'}^-$. Hence in this case the assumptions (81) of Corollaries 3.4 and 3.5 hold. In the case where $l=l(t)$ or $l'=l'(t)$, we only have $x\notin\bar\Sigma_l^+$ or $x\notin\bar\Sigma_{l'}^-$, but also in this case the corollaries hold since the only property needed in their proof is that $T^ix\notin[0,\sigma_{l(t)}\lambda^{(n_{l(t)})}]$ or $[1-\sigma_{l'(t)}\lambda_{\mathcal{I}}^{(n'_{l'(t)})},1)$ for $0\le i\le r$ and $r\le R_{\mathrm{M}}(t)$.

Since $r\ge t/3\ge t_1/3$, enlarging t_1, one can assure that $r\ge r_o$, for the r_o given by Corollaries 3.4 and 3.5 and since $x\in I\in\eta_p(t)$, x is not a singularity of $S_r(f)$ by properties (1) and (3) of Proposition 4.1. Hence one can apply Corollaries 3.4 and 3.5. Moreover, by property (3),

$$x_m\doteqdot\min_{0\le i<r}T^ix\ge\frac{M}{t\log\log t}\quad\text{and}\quad 1-x_{\mathrm{M}}\doteqdot\min_{0\le i<r}(1-T^ix)\ge\frac{M}{t\log\log t}.$$

Thus, we obtain, respectively,

$$S_r(f')(x)\le -C_1r\log r\left(1-\frac{C^+(\kappa'+1)t\log\log t}{C_1Mr\log r}\right),\tag{114}$$

$$|S_r(f')(x)|\le C_2r\log r\left(1+\frac{(\kappa+\kappa'+2)t\log\log t}{Mr\log r}\right),\tag{115}$$

$$|S_r(f'')(x)|\le C_4\frac{t\log\log t}{M}r\log r\left(1+\frac{(\kappa'+\kappa+2)t\log\log t}{Mr\log r}\right).\tag{116}$$

Recalling that $t\le 3r$ and again by enlarging t_1 if necessary, one can ensure that the last terms in (114), (115) and (116), involving $t\log\log t/r\log r$, are less than $1/2$. Hence we obtain (89), (90) and (91), respectively. □

COROLLARY 4.1. *For each $I=[a,b)\in\eta_s(t)$, $x\in I$ and $r(a,t)\le r\le r(b,t)$,*

$$\mathrm{const}(t\log t)\le|S_r(f')(x)|\le\mathrm{const}'(t\log t),\tag{117}$$

$$\Delta f(I,t)\ge\mathrm{const}\left(\frac{\log t}{\log\log t}\right)\xrightarrow{t\to+\infty}+\infty,\tag{118}$$

$$|\Delta f(I,t)|=o(\log t),\quad |\Delta f^1(I,t)|=o(\log t).\tag{119}$$

Proof. Equation (117) follows from (89) and (90), since $t/3\le r(a,t), r(b,t)\le 2t/m_f$. Since $S_{r(a,t)}(f)$ is continuous with its derivative on $[a,b]$ (properties (1) and (3) in Proposition 4.1), by the mean value theorem there exists $z\in I$ such that

$$\Delta f=S_{r(a,t)}(f)(a)-S_{r(a,t)}(f)(b)=-S_{r(a,t)}(f')(z)(b-a)\ge\frac{\mathrm{const}(t\log t)}{2t\log\log t},$$

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

where we applied (89) and the control on the interval sizes (property (2) in Proposition 4.1). The proof of (119) is obtained similarly by using (117). □

4.5. *Mixing partitions.* Let us prove Proposition 4.3.

4.5.1. *Uniform vertical distribution.* Let us show that given $\epsilon > 0$, choosing $M > M_0(\epsilon)$, each $I \in \eta_s(t)$ satisfies property (1) of Proposition 4.3. Let us recall the following definition used in [**5**] (see also [**13, 16, 17**]).

Definition 4.1. Given $\epsilon > 0$, the function g on the interval $[a, b]$ is *ϵ-uniformly distributed* if, for any c, d such that $\inf_{[a,b)} g \le c \le d \le \sup_{[a,b)} g$, the measure of the set $I_{c,d} = \{x \in [a, b) \mid c \le g(x) \le d\}$ satisfies

$$(1-\epsilon)\frac{d-c}{\sup_{[a,b)} g - \inf_{[a,b)} g} \le \frac{\mathrm{Leb}(I_{c,d})}{b-a} \le (1+\epsilon)\frac{d-c}{\sup_{[a,b)} g - \inf_{[a,b)} g}. \tag{120}$$

In [**5**], Fayad proves the following criterion to get uniform distribution.

LEMMA 4.3. (Fayad) *If g is monotonic and*

$$\sup_{[a,b)} |g''(x)||b-a| \le \epsilon \inf_{[a,b)} |g'(x)|, \tag{121}$$

then g is ϵ-uniformly distributed on $[a, b]$.

For each $[a, b) \in \eta_s(t)$, consider $S_{r(a)+j}(f)$, for $j = 1, \ldots, J([a, b), t) - 2$. From property (2) of Proposition 4.1, (89) and (91) and $r \le 2t/m_f$,

$$\begin{aligned}\frac{\sup_{[a,b)} |S_{r(a)+j}(f'')(x)||b-a|}{\inf_{[a,b)} |S_{r(a)+j}(f')(x)|} &\le \frac{C''r^2(\log r)(\log\log r)(b-a)}{MC_1' r \log r} \\ &\le \frac{C''((2t/m_f)\log\log(2t/m_f))(2/t \log\log t)}{MC_1'} \\ &\xrightarrow{t\to+\infty} \frac{4C''}{MC_1' m_f}.\end{aligned}$$

Choosing $M > M_0(\epsilon) \doteqdot 8C''/C_1' m_f \epsilon$ and then $t_3 \ge t_2$ large enough, the last expression is less than ϵ for $t \ge t_3$. Hence, each $S_{r(a)+j}(f)$, for $j = 1, \ldots, J-2$, being also decreasing, is ϵ-uniformly distributed on $[a, b) \in \eta_s(t)$ by Lemma 4.3. The set $I_j^{h(R)}$ defined in (94) is of the form $I_{c,d}$ for $g = S_{r(a)+j}(f)$, $c = t - h(R)$ and $d = t$; by the definition (93) of splitting points, one can check that $d = g(y_j) \le \sup_{[a,b)} g$ and $c \ge g(y_{j+1}) \ge \inf_{[a,b)} g$. Hence, by (120), its measure is bounded by

$$(1-\epsilon)\frac{h(R)}{\Delta f^j([a,b),t)}(b-a) \le \mathrm{Leb}(I_j^{h(R)}) \le (1+\epsilon)\frac{h(R)}{\Delta f^j([a,b),t)}(b-a). \tag{122}$$

This proves the uniform vertical distribution property (95).

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

4.5.2. *Rough upper and lower bound on the number of curves.* Recall that the number of curves generated from each $\varphi_t(I)$ is given by $J(I,t) = r(b,t) - r(a,t) + 1$.

LEMMA 4.4. *For each $I = [a, b)$,*

$$r(T^{r(a)}b, \Delta f) + 1 \leq J(I,t) \leq r(T^{r(a)+1}b, \Delta f^1) + 3. \tag{123}$$

Lemma 4.4 shows that the number of strips is related to the number of fibers that the point $T^{r(a)}b$ still has to cover in time Δf when a stops, because of the delay accumulated through the stretching of $S_{r(a)}(f)$.

Proof. Applying the relation $S_{r_1+r_2}(f)(x) = S_{r_1}(f)(x) + S_{r_2}(f)(T^{r_1}x)$ and the definition (2) of $r(\cdot,\cdot)$,

$$\begin{aligned} S_{r(T^{r(a)}b,\Delta f)+r(a)}(f)(b) &= S_{r(a)}(f)(b) + S_{r(T^{r(a)}b,\Delta f)}(f)(T^{r(a)}b) \\ &\leq S_{r(a)}(f)(b) + \Delta f = S_{r(a)}(f)(a) \leq t. \end{aligned}$$

Hence, $r(b,t) \geq r(T^{r(a)}b, \Delta f) + r(a,t)$, which is the first inequality in (123).

For the second inequality,

$$\begin{aligned} S_{r(T^{r(a)+1}b,\Delta f^1)+r(a)+2}(f)(b) &= S_{r(a)+1}(f)(b) + S_{r(T^{r(a)+1}b,\Delta f^1)+1}(f)(T^{r(a)+1}b) \\ &> S_{r(a)+1}(f)(b) + \Delta f^1 = S_{r(a)+1}(f)(a) > t, \end{aligned}$$

which implies that $r(b) < r(T^{r(a)+1}b, \Delta f^1) + r(a) + 2$. □

COROLLARY 4.2. *Let $\bar{J}(t) \doteqdot \sup_{I \in \eta_s(t)} J(I,t)$. Then*

$$\bar{J}(t) = o(\log t). \tag{124}$$

Proof. By Lemma 4.4, $J(I,t) \leq r(T^{r(a)+1}b, \Delta f^1) + 3 \leq \Delta f^1(I,t)/m_f + 3$. Recalling that $\Delta f^1(I,t) = o(\log t)$ by (119) of Corollary 4.1 we obtain the bound. □

Consider the map $R_t : I^{(0)} \to I^{(0)}$ given by $R_t(x) = T^{r(x,t)}x$, which is the projection of $\varphi_t(x,0) \in X_f$ to the base $I^{(0)}$. Note that R_t in general is not one-to-one. The following lemma is used by Kochergin in [**13**] (Lemma 1.3).

LEMMA 4.5. (Kochergin) *For any measurable set $S \subset I^{(0)}$,*

$$\mathrm{Leb}(R_t^{-1}S) \leq \int_S \left(\frac{f(x)}{m_f} + 1 \right) dx. \tag{125}$$

Since $f \in L^1$, by absolute continuity of the integral, for any $\delta > 0$ it is possible to choose δ_1 such that the right-hand side of (125) is bounded by δ as long as $\mathrm{Leb}(S) < \delta_1$. Hence we obtain the following corollary.

COROLLARY 4.3. *For each $\delta > 0$, there exists $\delta_1 > 0$ such that for any measurable $S \subset I^{(0)}$, if $\mathrm{Leb}(S) < \delta_1$, then $\mathrm{Leb}(R_t^{-1}S) < \delta$.*

LEMMA 4.6. *There exist partitions $\eta_5(t) \subset \eta_s(t)$ and $t_3 \geq t_2$ such that, for $t \geq t_3$, $\mathrm{Leb}(\eta_5(t)) \geq \mathrm{Leb}(\eta_s(t)) - \delta$ and for each $x \in I \in \eta_5(t)$,*

$$|T^r x| \geq \frac{1}{(\log t)^2}, \quad |1 - T^r x| \geq \frac{1}{(\log t)^2}, \tag{126}$$

for each $r(a,t) \leq r \leq r(a,t) + J(I,t)$.

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

Proof. Define

$$U_3(t) \doteqdot \bigcup_{i=-[\bar{J}(t)]}^{[\bar{J}(t)]} T^i\left(\left[0, \frac{1}{(\log t)^2}\right]\right) \cup \bigcup_{i=-[\bar{J}(t)]}^{[\bar{J}(t)]} T^i\left(\left[1-\frac{1}{(\log t)^2}, 1\right]\right). \tag{127}$$

Since the continuity intervals for $T^{[\bar{J}(t)]}$ and $T^{-[\bar{J}(t)]}$ are at most $d([\bar{J}(t)]+1)$ (see §4.3.1), the set $U_3(t)$ consists of at most $O(\bar{J}(t)^2)$ disjoint intervals. Consider the $(2/t \log\log t)$-neighborhood of $U_3(t)$, i.e. let $U_4(t) \doteqdot \{x \in I^{(0)} \mid d(x, U_3(t)) \leq 2/(t \log\log t)\}$. Hence, using Corollary 4.2,

$$\mathrm{Leb}(U_4(t)) \leq \frac{4\bar{J}(t)+4}{(\log t)^2} + \mathrm{const}\frac{\bar{J}(t)^2}{t \log\log t} \xrightarrow{t\to+\infty} 0.$$

Choosing $t_3 > t_2$ so that for $t \geq t_3$, $\mathrm{Leb}(U_4) < \delta_1$ where δ_1 is given by Corollary 4.3, we obtain $\mathrm{Leb}(R_t^{-1}(U_4)) < \delta$. Define a refined partition $\eta_5(t) \subset \eta_s(t)$ by

$$\eta_5(t) \doteqdot \eta_s(t)\backslash\{I=[a,b) \in \eta_s(t) \mid \bar{I}=[a,b] \subset R_t^{-1}U_4(t)\}.$$

Clearly $\mathrm{Leb}(\eta_5(t)) \geq \mathrm{Leb}(\eta_s(t)) - \delta$. Let us show that, for each $I \in \eta_5(t)$, we obtain (126). By construction there exists $x \in I$ such that $R_t(x) = T^{r(x,t)}x \notin U_4(t)$. Hence, by Proposition 4.1, $T^{r(x,t)}y \notin U_3(t)$ for each $y \in \bar{I}$. For each $r = r(a), \ldots, r(a)+J$, the point $T^r y$ satisfies the inequalities (126) by definition of $U_3(t)$, because $T^{r(x,t)}y \notin U_3(t)$, as shown above, and $|r(x,t) - r| \leq J$. □

LEMMA 4.7. (Rough lower bound on J) *Let* $\underline{J}(t) \doteqdot \inf_{I\in\eta_5(t)} J(I,t)$, *so*

$$\underline{J}(t) \geq \mathrm{const}\frac{\log t}{(\log\log t)^2} \xrightarrow{t\to+\infty} +\infty. \tag{128}$$

Proof. For each $I = [a,b) \in \eta_5(t)$, by Lemma 4.6, $f(T^r b) \leq \mathrm{const}\log(\log t)$ for each $r(a,t) \leq r \leq r(a,t) + J(I,t)$. Hence, since by Lemma 4.4, $J(I,t) \geq r(T^{r(a)}b, \Delta f) + 1$,

$$J(I,t) \geq \frac{S_{r(T^{r(a)}b,\Delta f)+1}(f)(T^{r(a)}b)}{\max_{0\leq i<r(T^{r(a)}b,\Delta f)+1} f(T^{r(a)+i}b)} \geq \mathrm{const}\frac{\Delta f}{\log(\log t)},$$

which gives (128) by using the bound (118) on Δf. □

4.5.3. *Variation of slopes.* Given $I = [a,b) \in \eta_5(t)$ the variation of the average slope of the curves $(t - S_{r(a)+j}(f))|_I$, for $0 \leq j \leq J(I,t)$ can be written as

$$\begin{aligned}|\Delta f^j - \Delta f| &= |S_{r(a)+j}(f)(a) - S_{r(a)+j}(f)(b) - S_{r(a)}(f)(a) + S_{r(a)}(f)(b)| \\ &= \left|\sum_{i=0}^{j-1} f(T^{r(a)+i}a) - \sum_{i=0}^{j-1} f(T^{r(a)+i}b)\right| \leq \sum_{i=0}^{j-1} |f'(T^{r(a)+i}c)|(b-a),\end{aligned}$$

where in the last estimate $a \leq c \leq b$ by the mean value theorem. Using Lemma 4.6, $|f'(T^{r(a)+i}c)| \leq \mathrm{const}(\log t)^2$ for each $0 \leq i < j \leq \bar{J}(t)$. Hence, applying also the bound on $\bar{J}(t)$ given by Corollary 4.2, the growth estimate (118) for Δf and the size control of $(b-a)$ (property (2) of Proposition 4.1),

$$\left|\frac{\Delta f^j - \Delta f}{\Delta f}\right| \leq \frac{\bar{J}(t)\sup_{i=0}^{J-1}|f'(T^{r(a)+i}c)|(b-a)}{\Delta f} \leq \mathrm{const}\frac{\log t(\log t)^2}{(\log t/\log\log t)t \log\log t},$$

which converges to zero as $t \to +\infty$. Enlarge $t_3 > 0$ so that the right-hand side is less than ϵ for $t \geq t_3$ to obtain (96).

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

4.5.4. *Equidistribution on the base and asymptotic number of curves.* Both the equidistribution on the base and the exact asymptotic number of curves follow by proving uniform convergence on a large set for the Birkhoff sums of χ and f, respectively. More precisely, one seeks uniform control for the points of the form $T^{r(a)}b$ where b are the endpoints of the partition intervals $[a, b)$.

Let $\widetilde{\eta}_5(t)$ be a narrowing of $\eta_5(t)$ obtained by keeping only the central third of each interval:

$$\widetilde{\eta}_5(t) \doteqdot \left\{ \left[a + \frac{(b-a)}{3}, b - \frac{(b-a)}{3} \right) \,\middle|\, [a, b) \in \eta_5(t) \right\}.$$

For each $\varepsilon > 0$ and $\delta_1 > 0$, by ergodicity of T and T^{-1} one can find U_5 and $N > 0$ such that $\mathrm{Leb}(U_5) < \delta$ and for each $x \notin U_5$ and $n \geq N$,

$$\left| \frac{S_n(f, T^i)(x)}{n} - 1 \right| < \varepsilon, \quad \left| \frac{S_n(\chi, T^i)(x)}{n} - (b_2 - b_1 - 2\delta) \right| < \varepsilon, \quad i = 1, -1. \quad (129)$$

If δ_1 is given by Corollary 4.3 in correspondence of $\delta/3 > 0$, we obtain $\mathrm{Leb}(R_{-t}^{-1}(U_5)) < \delta/3$. Define $\eta_6(t) \subset \eta_5(t)$ by throwing away all intervals $I \in \eta_5(t)$ such that the corresponding $\widetilde{I}$ is completely contained in $R^{-t}(U_5)$. Hence, $\mathrm{Leb}(\eta_6(t)) \geq \mathrm{Leb}(\eta_5(t)) - \delta \geq \mathrm{Leb}(\eta_s(t)) - 2\delta$ by Lemma 4.6.

By construction, for each $I \in \eta_6(t)$, there exists $\bar{x}$ such that $|\bar{x} - a|, |\bar{x} - b| > (b-a)/3$ and $T^{r(\bar{x},t)}\bar{x} \notin U_5$ and hence equation (129) holds for $x = \bar{x}$. Arguing as in Corollary 4.1 to prove (118), both $\Delta f([a, \bar{x}), t)$ and $\Delta f([\bar{x}, b), t)$, as $t \to \infty$, are bounded from below by $\mathrm{const}(\log t/\log\log t)$. As in Lemma 4.4,

$$r(\bar{x}) - r(a) \geq r(T^{r(a)}\bar{x}, \Delta f([a, \bar{x}), t)), \quad r(b) - r(\bar{x}) \geq r(T^{r(\bar{x})}b, \Delta f([\bar{x}, b), t)).$$

Hence, by the same proof as in Lemma 4.7, both $r(\bar{x}) - r(a)$ and $r(b) - r(\bar{x})$ tend to infinity uniformly as t increases. Choose t_4 so that, for $t \geq t_4$, both $r(\bar{x}) - r(a) > N$ and $r(b) - r(\bar{x}) > N$. Hence, the estimates in (129) hold when $n = r(\bar{x}) - r(a)$ or $r(b) - r(\bar{x})$ and $x = T^{r(\bar{x},t)}\bar{x}$. Moreover, they also hold for $i = -1$ and $x = T^{r(\bar{x},t)-1}\bar{x}$. To see this, in the case of f, use the fact that

$$S_{r(\bar{x})-r(a)}(f, T^{-1})(T^{r(\bar{x},t)-1}\bar{x}) = S_{r(\bar{x})-r(a)+1}(f, T^{-1})(T^{r(\bar{x},t)}\bar{x}) - f(T^{r(\bar{x},t)}\bar{x})$$

and from Lemma 4.6 and the analogous result of Lemma 4.7 for $\Delta f([a, \bar{x}), t)$,

$$\frac{f(T^{r(\bar{x},t)}\bar{x})}{r(\bar{x}) - r(a)} \leq \mathrm{const} \log(\log t)^2 \frac{(\log\log t)^2}{\log t},$$

which can be made arbitrarily small by enlarging t_4 if necessary. In the case of χ, just use the fact that $\chi \leq 1$ and $r(\bar{x}) - r(a)$ tends to infinity.

Let us combine these estimates decomposing the Birkhoff sums as

$$S_{r(b)-r(a)}(\chi, T)(T^{r(a)}\bar{x}) = S_{r(\bar{x})-r(a)}(\chi, T^{-1})(T^{r(\bar{x})-1}\bar{x}) + S_{r(b)-r(\bar{x})}(\chi, T)(T^{r(\bar{x})}\bar{x})$$

and using the fact that

$$\frac{r(b) - r(\bar{x})}{r(b) - r(a)} + \frac{r(\bar{x}) - r(a)}{r(b) - r(a)} = 1.$$

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

We obtain

$$\left|\frac{S_{r(b)-r(a)}(\chi, T)(T^{r(a)}\bar{x})}{r(b)-r(a)} - (b_2 - b_1 - 2\delta)\right| \leq 2\epsilon, \tag{130}$$

which proves equidistribution on the base (98) for $\bar{x} \in I$.

LEMMA 4.8. *Enlarging t_4 if necessary, for each $[a, b) \in \eta_6(t)$, $t \geq t_4$,*

$$\left|\frac{1}{r(b)-r(a)} S_{r(b)-r(a)}(f)(T^{r(a)}b) - 1\right| \leq 2\varepsilon.$$

Proof. By the mean value theorem, there exists $z \in [\bar{x}, b)$ such that

$$\begin{aligned}
&|S_{r(b)-r(\bar{x})}(f)(T^{r(\bar{x},t)}\bar{x}) - S_{r(b)-r(\bar{x})}(f)(T^{r(\bar{x},t)}b)| \\
&\quad \leq |S_{r(b)-r(\bar{x})}(f')(T^{r(\bar{x},t)}z)| \cdot (b-\bar{x}) \leq \bar{J}(t) \sup_{r(\bar{x}) \leq i \leq r(b)} |f'(T^i z)|(b-\bar{x}) \\
&\quad \leq \mathrm{const}\frac{\log t (\log t)^2}{t \log\log t},
\end{aligned}$$

where we have used Corollary 4.2 to bound $\bar{J}(t)$, Lemma 4.6 to bound $|f'(T^i z)|$ and property (2) in Proposition 4.1 to control the size $(b-a)$. Hence, enlarging t_4, from the analogous estimate for $T^{r(\bar{x},t)}\bar{x}$, we obtain for $t \geq t_4$,

$$\left|\frac{1}{r(b)-r(\bar{x})} S_{r(b)-r(\bar{x})}(f)(T^{r(\bar{x},t)}b) - 1\right| \leq 2\varepsilon. \tag{131}$$

In a similar way, from the analogous estimate for $T^{r(\bar{x},t)-1}\bar{x}$, we obtain

$$\left|\frac{1}{r(a)-r(\bar{x})} S_{r(a)-r(\bar{x})}(f, T^{-1})(T^{r(\bar{x},t)-1}b) - 1\right| < 2\varepsilon. \tag{132}$$

Combining (131) and (132) and decomposing the Birkhoff sums as

$$S_{r(b)-r(a)}(f, T)(T^{r(a)}b) = S_{r(\bar{x})-r(a)}(f, T^{-1})(T^{r(\bar{x})-1}b) + S_{r(b)-r(\bar{x})}(f, T)(T^{r(\bar{x})}b)$$

we obtain the following lemma. □

LEMMA 4.9. *Enlarging t_4 if necessary, for each $[a, b) \in \eta_6(t)$, if $t \geq t_4$,*

$$\left|\frac{S_{r(b)-r(a)}(f)(T^{r(a)}b)}{\Delta f} - 1\right| \leq \varepsilon.$$

Proof. Since we can rewrite

$$\begin{aligned}
S_{r(b)-r(a)}(f)(T^{r(a)}b) &= S_{r(b)}(f)(b) - S_{r(a)}(f)(b) + S_{r(a)}(f)(a) - S_{r(a)}(f)(a) \\
&= \Delta f + S_{r(b)}(f)(b) - S_{r(a)}(f)(a),
\end{aligned}$$

from $t - f(T^{r(a)}a) < S_{r(a)}(f)(a) \leq t$ and $t - f(T^{r(b)}b) < S_{r(b)}(f)(b) \leq t$, we obtain

$$\left|\frac{S_{r(b)-r(a)}(f)(T^{r(a)}b)}{\Delta f} - 1\right| \leq \frac{\max\{f(T^{r(b)}b), f(T^{r(a)}a)\}}{\Delta f} \leq \mathrm{const}\frac{(\log\log t)^2}{\log t} \to 0,$$

by using Lemma 4.6 and (118). □

https://doi.org/10.1017/S0143385706000836 Published online by Cambridge University Press

From Lemmas 4.8 and 4.9, for each $\epsilon > 0$, and by choosing ε appropriately, we obtain

$$\left| \frac{r(b,t) - r(a,t)}{S_{r(b)-r(a)}(f)(T^{r(a)}b)} \frac{S_{r(b)-r(a)}(f)(T^{r(a)}b)}{\Delta f} - 1 \right| \le \epsilon$$

for $t \ge t_4$. Recalling that $J([a,b),t) - 1 = r(b,t) - r(a,t)$, this concludes the proof of the asymptotic number of curves (97).

Setting $\eta_m(t) = \eta_6(t)$, this completes the verification that the partitions $\eta_m(t)$, for an appropriate choice of δ and $t \ge \bar{t} \doteqdot \max\{t_3, t_4\}$, satisfy all the properties listed in Proposition 4.3.

Acknowledgements. I would like to thank my advisor Professor Ya. G. Sinai for proposing the problem and constantly guiding me with suggestions, patience and encouragement. I would also like to thank both K. Khanin and A. Avila for useful discussions, P. Batchourine for listening to some parts of this work and the referee for carefully reading the manuscript and pointing out misprints.

References

[1] V. I. Arnold. Topological and ergodic properties of closed 1-forms with incommensurable periods. *Funktsional. Anal. i Prilozhen.* **25**(2) (1991), 1–12. (Engl. Transl. *Funct. Anal. Appl.* **25**(2) (1991), 81–90.)

[2] A. Avila, S. Gouezel and J.-C. Yoccoz. Exponential mixing for the Teichmüller flow. arXiv: math.DS/0511614, *Pub. Math. IHES* to appear.

[3] A. I. Bufetov. Decay of correlations for the Rauzy–Veech–Zorich induction map on the space of Interval exchange transformation and the central limit theorem for the Teichmüller flow on the moduli space of abelian differentials. *J. Amer. Math. Soc.* **19** (2006), 579–623.

[4] I. P. Cornfeld, S. V. Fomin and Ya. G. Sinai. *Ergodic Theory*. Springer, New York, 1980.

[5] B. R. Fayad. Analytic mixing reparametrizations of irrational flows. *Ergod. Th. & Dynam. Sys.* **22**(2) (2002), 437–468.

[6] K. Frączek and M. Lemańczyk. A class of special flows over irrational rotations which is disjoint from mixing flows. *Ergod. Th. & Dynam. Sys.* **24** (2004), 1083–1095.

[7] K. Frączek and M. Lemańczyk. On disjointness properties of some smooth flows. *Fund. Math.* **185**(2) (2005), 117–142.

[8] A. B. Katok. Interval exchange transformations and some special flows are not mixing. *Israel J. Math.* **35**(4) (1980), 301–310.

[9] M. Keane. Interval exchange transformations. *Math. Z.* **141** (1975), 25–31.

[10] S. P. Kerckhoff. Simplicial systems for interval exchange maps and measured foliations. *Ergod. Th. & Dynam. Sys.* **5** (1985), 257–271.

[11] K. M. Khanin. Mixing for area-preserving flows on the two-dimensional torus. *Internat. J. Modern Phys.* B **10**(18–19) (1996), 2167–2188.

[12] A. V. Kochergin. 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), 512–518. (Engl. Transl. *Soviet Math. Dokl.* **13** (1972), 949–952.)

[13] A. V. Kochergin. Mixing in special flows over a shifting of segments and in smooth flows on surfaces. *Mat. Sb.* **96**(138) (1975), 471–502.

[14] A. V. Kochergin. Nonsingular saddle points and the absence of mixing. *Mat. Zametki* **19**(3) (1976), 453–468. (Engl. Transl. *Math. Notes* **19**(3), 277–286.)

[15] A. V. Kochergin. A mixing special flow over a rotation of the circle with an almost Lipschitz function. *Mat. Sb.* **193**(3) (2003), 51–78.

[16] A. V. Kochergin. Non-degenerate fixed points and mixing in flows on a 2-torus. *Mat. Sb.* **194**(8) (2003), 83–112. (Engl. Transl. *Sb. Math.* **194**(8) (2003), 1195–1224.)

[17] A. V. Kochergin. Non-degenerate fixed points and mixing in flows on a 2-torus.II. *Mat. Sb.* **195**(3) (2004), 83–112. (Engl. Transl. *Sb. Math.* **195**(3) (2004), 317–346.)

[18] A. V. Kochergin. Some generalizations of theorems on mixing flows with nondegenerate saddles on a two-dimensional torus. *Mat. Sb.* **195**(9) (2004), 19–36.

[19] A. V. 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.

[20] M. Lemańczyk. Sur l'absence de mélange pour des flots spéciaux au-dessus d'une rotation irrationnelle. *Colloq. Math.* **84–85** (2000), 29–41.

[21] S. Marmi, P. Moussa and J.-C. Yoccoz. The cohomological equation for Roth-type interval exchange maps. *J. Amer. Math. Soc.* **18**(4) (2005), 823–872.

[22] G. Rauzy. Échanges d'intervalles et transformations induites. *Acta Arith.* **XXXIV** (1979), 315–328.

[23] Y. G. Sinai and K. M. Khanin. Mixing for some classes of special flows over rotations of the circle. *Funktsional. Anal. i Prilozhen.* **26**(3) (1992), 1–21. (Engl. Transl. *Funct. Anal. Appl.* **26**(3) (1992), 155–169.)

[24] W. A. Veech. Projective Swiss cheeses and uniquely ergodic interval exchange transformations. *Ergod. Th. & Dynam. Sys.* Birkhäuser, Berlin, 1981, pp. 113–193.

[25] W. A. Veech. Gauss measures for transformations on the space of interval exchange maps. *Ann. of Math.* (2) **115** (1982), 201–242.

[26] A. Zorich. Finite Gauss measure on the space of interval exchange transformation. Lyapunov exponents. *Ann. Inst. Fourier, Grenoble* **46** (1996), 325–370.

[27] A. Zorich. How do the leaves of a closed 1-form wind around a surface? *Trans. Amer. Math. Soc.* **197** (1999), 135–178.