

# QUANTITATIVE SHRINKING TARGET PROPERTIES FOR ROTATIONS, INTERVAL EXCHANGES AND BILLIARDS IN RATIONAL POLYGONS

JON CHAIKA AND DAVID CONSTANTINE

ABSTRACT. This paper presents quantitative shrinking target results for rotations and then extends them to flows on flat surfaces and particularly billiards in rational polygons. To do this a quantitative version of a unique ergodicity criterion of Boshernitzan is established. This is equivalent to a quantitative version of Masur's criterion.

## 1. INTRODUCTION

1.1. **Background.** Let  $\alpha \in [0, 1)$  and  $\lambda$  denote Lebesgue measure on  $[0, 1)$ .  $R_\alpha : [0, 1) \rightarrow [0, 1)$  by  $R_\alpha(x) = x + \alpha \pmod{1}$  is one of the most natural and best understood dynamical systems. For example, the following is known.

**Theorem 1.** (*Kurzweil [14]*) *For any decreasing sequence of positive real numbers  $\{b_i\}_{i=1}^\infty$  with divergent sum there exists  $\mathcal{V} \subset [0, 1)$ , a full measure set of  $\alpha$ , such that for all  $\alpha \in \mathcal{V}$  we have*

$$\lambda \left( \bigcap_{n=1}^\infty \bigcup_{i=n}^\infty B(R_\alpha^{-i}(x), b_i) \right) = 1$$

for every  $x$ .

On the other hand,

$$\lambda \left( \bigcap_{n=1}^\infty \bigcup_{i=n}^\infty B(R_\alpha^{-i}(x), b_i) \right) = 1$$

for every  $x$  and every decreasing sequence of positive real numbers  $\{b_i\}_{i=1}^\infty$  with divergent sum iff  $\alpha$  is badly approximable.

Recall that  $\alpha$  is badly approximable if all of the terms in its continued fraction expansion are uniformly bounded. Throughout this paper we will denote the continued fraction expansion of  $\alpha$  by  $[a_1, a_2, \dots]$  and by  $q_n$  the denominator of the  $n^{\text{th}}$  convergent to  $\alpha$ .

The above theorem can be stated in terms of whether points visit shrinking balls infinitely often under  $R_\alpha$ . Much of it was extended to interval exchange transformations (see Definition 4.3 and Section 4.3):

**Theorem 2.** [4]

- (1) *If  $\{b_i\}_{i=1}^\infty$  is a decreasing sequence of positive real numbers with divergent sum then for almost every IET,  $T$ , we have that  $\lambda \left( \bigcap_{n=1}^\infty \bigcup_{i=n}^\infty B(T^{-i}x, b_i) \right) = 1$  for all  $x$ .*
- (2) *On the other hand, for almost every IET there exists a decreasing sequence of positive real numbers  $\{b_i\}_{i=1}^\infty$  with divergent sum such that for almost every  $x$ ,  $\lambda \left( \bigcap_{n=1}^\infty \bigcup_{i=n}^\infty T^{-i}B(x, b_i) \right) = 0$ .*

- (3) However, for almost every IET and any decreasing sequence of positive real numbers  $\{b_i\}_{i=1}^{\infty}$  with divergent sum such that  $ib_i$  is decreasing we have that  $\lambda(\bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} T^{-i}B(x, b_i)) = 1$  for almost every  $x$ .

This statement says that, with some necessary caveats, orbits of points typically hit shrinking balls infinitely often. This paper considers the related stronger question of whether these balls are visited as often as one would expect. We now state three answers to this question for rotations. Similar results hold for any flat surface and thus for billiards in rational polygons by the unfolding construction [8] (see Section 4).

## 1.2. Statement of results in the case of rotations.

**Theorem 3.** *If  $\alpha$  is an irrational number such that  $a_n \leq n^{\frac{7}{6}}$  for all but finitely many  $n$  then*

$$\lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N \chi_{B(0, \frac{1}{i})}(R_{\alpha}^i x)}{\sum_{i=1}^N \frac{2}{i}} = 1$$

for almost every  $x$ .

It is easy to see that Theorem 3 fails to hold for a dense  $G_{\delta}$  set of rotations, including some non-Liouville numbers. Related versions Theorems 12 and 13 hold in any flat surface, and in particular billiards in rational polygons.

Theorem 3 states a very general condition which implies a result of the kind we are considering for a particular natural target. The next two results consider addressing larger families of targets.

**Theorem 4.** *If  $\alpha$  is badly approximable,  $\{b_i\}_{i=1}^{\infty}$  is non-increasing,  $y \in [0, 1)$  and  $\sum b_i = \infty$  then*

$$\lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N \chi_{B(y, b_i)}(T^i x)}{\sum_{i=1}^N \mu(B(y, b_i))} = 1$$

for almost every  $x$ .

Note that the full measure set of  $x$  depends on  $y$ .

Theorem 17 establishes a version for flows in flat surfaces (including billiards in rational polygons). The condition on directions of flows is analogous to badly approximable and in particular has Hausdorff dimension 1 in each flat surface [13]. Recently, this diophantine condition was shown to hold for an absolutely winning set of directions [5] (see [17] for a definition of absolutely winning). Theorem 1 implies that for rotations this is the mildest condition on  $\alpha$  we could hope. If we want a result that holds for almost every  $\alpha$  we need to restrict our choice of sequences further.

Consider a sequence  $\{b_i\}_{i=1}^{\infty}$  where  $ib_i$  is non-increasing and  $\sum_{i=1}^{\infty} b_i = \infty$ . We call these sequences *Khinchin sequences*. We say  $\alpha$  is *weakly bounded* if

$$\lim_{C \rightarrow \infty} \limsup_{N \rightarrow \infty} \frac{1}{N} \left( \sum_{i=1}^N \log a_i - \sum_{a_i < C} \log a_i \right) = 0.$$

**Theorem 5.** *If  $\alpha$  is weakly bounded then it has the property that for any Khinchin sequence  $\{b_i\}_{i=1}^\infty$*

$$\lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N \chi_{B(0, b_i)}(R_\alpha^i x)}{\sum_{i=1}^N 2b_i} = 1$$

for almost every  $x$ .

Once again, an analogous result holds in every flat surface (see Section 5) with similar, full measure diophantine condition. This establishes the result for billiards in rational polygons. By our methods we also obtain the following shrinking target result:

**Theorem 6.**  *$\alpha$  has the the property that for any Khinchin sequence  $\{b_i\}_{i=1}^\infty$   $\lambda(\bigcap_{n=1}^\infty \bigcup_{i=n}^\infty B(R_\alpha^i x, b_i)) = 1$  if and only if  $\limsup_{n \rightarrow \infty} \frac{\log q_n}{n} < \infty$ .*

**1.3. A related problem.** If one is concerned about a specific sequence and not concerned about a diophantine condition things are much simpler. Observe (see for example [14, Proof of Lemma 7]) that for all  $a, b, y_1, y_2 \in [0, 1)$  and  $m \neq n \in \mathbb{Z}$  we have

$$(1) \quad \lambda \times \lambda(\{(x, \theta) : |R_\theta^n(x) - y_1| \leq a \text{ and } |R_\theta^m(x) - y_2| \leq b\}) = \lambda \times \lambda(\{(x, \theta) : |R_\theta^n(x) - y_1| \leq a\}) \lambda \times \lambda(\{(x, \theta) : |R_\theta^m(x) - y_2| \leq b\}).$$

From this fact we readily get a convergence in measure statement. That is, for all  $\epsilon > 0$  we have

$$\lim_{N \rightarrow \infty} \lambda \times \lambda(\{(x, \theta) : \left| \frac{\sum_{n=1}^N \chi_{B(0, \frac{1}{n})}(R_\theta^n x)}{\sum_{n=1}^N \frac{2}{n}} - 1 \right| > \epsilon\}) = 0.$$

One can use the method of subsequences to prove the result almost everywhere. Additionally, if one wishes to consider

$$\lim_{N \rightarrow \infty} \lambda \times \lambda(\{(x, \theta) : \left| \frac{\sum_{n=1}^N \chi_{B(0, b_n)}(R_\theta^n x)}{\sum_{n=1}^N 2b_n} - 1 \right| > \epsilon\}) = 0,$$

once again one readily gets convergence in measure. One can argue that when  $\sum_{n=N}^M b_n$  is small the set of  $\alpha$  where there is an  $x$  with  $\sum_{n=N}^M \chi_{B(0, b_n)}(R_\alpha^n(x))$  large is small to prove convergence almost everywhere (at least when  $b_n$  is non-increasing with divergent sum). See for example [18] for a fine result established by a (more involved) approach like this.

This argument is a little deceptive, because in the absence of any kind of explicit condition it says nothing about how a particular sequence behaves with a particular rotation. In Section 3 we consider this problem, where the shrinking target is determined not by some predetermined analytic constraint (such as shrinking like

$\frac{1}{i}$ ) but rather arises from asking a natural question about the dynamics of  $R_\alpha$ . The proof is similar in flavor to the other results; interestingly, however, only a weaker estimate on frequency of visit times is possible. For almost every  $\alpha$  this frequency does not converge almost everywhere to the constant function.

**1.4. Outline of paper.** This paper first establishes Theorem 3 in Section 2. In Section 4 it extends Theorem 3 in two variants to any flat surface, which apply to billiards in rational polygons. The proof of Theorem 3 is a template for the other proofs (though they present some complications). This paper concludes in Section 5 by establishing Theorems 4 and 5 in a general form which also has an application to flat surfaces and billiards in rational polygons.

Most of our proofs follow the outline of the strong law of large numbers, using probabilistic methods (in the fixed dynamical system, not the parametrizing space). The strategy is to show that hitting a ball at one time and a different time are independent events. This is of course false but things like it are true. We use a quantitative version of Boshernitzan/Masur's criterion [19] [3] [15] to establish the approximate independence we can show (Theorem 15). The statement of this result requires technical definitions and is postponed to Section 4.

**1.5. Acknowledgments.** The first author would like to thank B. Fayad and D. Kleinbock for encouraging me to pursue this question. We would like to thank J. Athreya, M. Boshernitzan, A. Eskin, H. Masur, R. Vance and W. Veech for helpful conversations. The first author is partially supported by NSF grant DMS-1004372.

## CONTENTS

1. Introduction	1
1.1. Background	1
1.2. Statement of results in the case of rotations	2
1.3. A related problem	3
1.4. Outline of paper	4
1.5. Acknowledgments	4
2. Proof of Theorem 3	5
2.1. Relation to dynamics	5
2.2. Dividing	6
2.3. Basic properties	6
2.4. Describing $g_n$	7
2.5. Independence	8
2.6. The probabilistic theorem	9
3. Undetermined points as a shrinking target	11
4. Results on flat surfaces and billiards	19
4.1. Flat surfaces	19
4.2. Billiards	20
4.3. Interval exchange transformations and their symbolic coding	20
4.4. Generalizing to billiards	21
4.5. Proof of Theorem 15	23
4.6. Connection of flow to symbolic coding	24
4.7. Defining $g_j$ in this context	24
4.8. Proof of Theorem 12	26
4.9. Outline of the proof of Theorem 13	26

5. More targets	27
5.1. Proof of Theorem 4	27
5.2. Proof of Theorem 5	29
5.3. Proof of Theorem 6	31
References	32

## 2. PROOF OF THEOREM 3

### 2.1. Relation to dynamics.

**Lemma 2.1.** *Almost every  $\alpha$  satisfies the diophantine condition in Theorem 3.*

This is classical consequence of the fact that  $\sum_{n=1}^{\infty} \frac{1}{n^6} < \infty$ . See for example [12, Theorem 30].

We next prove a maximal result.

**Lemma 2.2.**  $\frac{1}{(a_{n+1}+2)q_n} < d(R^{q_n}x, x) < \frac{1}{a_{n+1}q_n}$ .

*Proof.* See, for example, [12], four lines before equation 34, which says

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

By multiplying everything by  $q_k$  and recalling that  $q_{k+1} = a_{k+1}q_k + q_{k-1}$  and so  $q_{k+1} \geq a_{k+1}q_k$  and  $q_k + q_{k+1} \leq (a_{k+1} + 2)q_k$  we obtain the lemma.  $\square$

**Lemma 2.3.**  $\sum_{i=q_{n-1}}^{q_n} \chi_{B(0, \frac{1}{i})}(R^i(x)) \leq 4\sqrt{a_n} + 1$ .

*Proof.* Observe that  $(\sqrt{a_n} + 1)d(R^{q_{n-1}}(x), x) \geq \frac{1}{\sqrt{a_n}q_{n-1}}$ . Also,  $\{i : R^i(x) \in [a, a + \frac{1}{3q_n}]\}$  has at most 1 element out of every  $q_n$  consecutive integers (Lemma 2.2).

Therefore we have  $\sum_{i=q_{n-1}}^{q_n} \chi_{B(0, \frac{1}{i})}(R^i(x)) \leq \sqrt{a_n} + \sum_{i=(\sqrt{a_n}+1)q_{n-1}}^{q_n} \chi_{B(0, \frac{1}{i})}(R^i(x)) \leq 4\sqrt{a_n} + 1$ .  $\square$

Consider a set  $S = \{x_1, \dots, x_k\}$ . Let  $P_S$  be the partition of  $[0, 1)$  into intervals bounded by elements of this set. If  $r$  is a number then  $P_r$  denotes  $P_S$  when  $S = \{0, R(0), \dots, R^{q_r-1}(0)\}$ .

**Lemma 2.4.** *There exists a function  $f$  constant on each element of  $P_n$  such that*

$$\int_0^1 |f(x) - \sum_{i=1}^N \chi_{B(0, \frac{1}{i})}(R^i x)| dx \leq \frac{2N}{q_n}.$$

*Proof.* Notice that  $\{0, R(0), \dots, R^{q_n-1}(0)\}$  is at least  $\frac{2}{q_n}$  dense and change each of the summands to be constant on the partition elements. See for example [11, Theorem 1] quoted as Theorem 7 in this paper.  $\square$

**Lemma 2.5.** *Let*

$$f(x) = \chi_{[a,b]}(x) \sum_{i=M}^N \chi_{B(0, \frac{1}{i})} R^i(x)$$

and

$$g(x) = \chi_{[R^k(a), R^k(b)]}(x) \sum_{i=M+k}^{N+k} \chi_{B(0, \frac{1}{i})}(R^i(x)).$$

Then  $\int_0^1 |f(x) - g(x)| dx \leq \sum_{i=M}^N \frac{2}{i} - \frac{2}{i+k}$  which is  $O(\frac{k}{N})$ .

The above lemma can be easily improved especially when  $[a, b]$  is small and well adapted to the dynamics of rotation. However, we do not need this.

**2.2. Dividing.** Let  $u_1 = q_{10}$  and  $u_n = \max\{q_r : \log(q_r) < \log(u_{n-1}) + \sqrt[3]{\log(u_{n-1})}\}$ ,  $u'_n = \max\{q_r : q_r < n^4 u_{n-1}\}$  and  $u''_n = \max\{q_r : q_r < n^4 u'_n\}$ . Let  $v_n$  be the number such that  $u_n = q_{v_n}$ . Let  $v'_n$  be the number such that  $u'_n = q_{v'_n}$  and  $v''_n$  be the number such that  $u''_n = q_{v''_n}$ . Let

$$g_n(x) = \sum_{i=u''_n}^{u_n} \chi_{B(0, \frac{1}{i})}(R^i x).$$

The numbers  $u_n$  are chosen so that if  $u_n = 2^t$  then  $u_{n+1}$  is roughly  $2^{t + \sqrt[3]{t}}$ .

### 2.3. Basic properties.

**Lemma 2.6.**  $\int_0^1 g_n(x) dx \leq (\sum_{i=1}^{n-1} \int_0^1 g_i(x) dx)^{\frac{1}{3}}$ .

This is by construction of  $u_n$ .

**Corollary 2.7.**  $v_n = O(n^{\frac{4}{3}})$ .

If  $s_{r+1} < s_r + \sqrt[3]{s_r}$  for all  $r$  then  $s_{r+1} = O(r^{\frac{4}{3}})$ . The result follows from the fact that there exists  $c > 1$  such that  $q_n > c^n$  for all large  $n$  and any  $\alpha, n$ .

**Lemma 2.8.** For any  $\epsilon > 0$  for all but finitely many  $n$  we have  $\int_0^1 g_n(x) dx \geq (1 - \epsilon) (\sum_{i=1}^{n-1} \int_0^1 g_i(x) dx)^{\frac{1}{3}}$ .

This is because our condition that  $a_i < i^{\frac{7}{6}}$  implies that  $\log(q_{v_n})$  is very close to  $\log(q_{v_{n-1}}) + \sqrt[3]{\log(q_{v_{n-1}})}$ .

**Corollary 2.9.**  $\liminf_{n \rightarrow \infty} \frac{\log \sum_{i=1}^{n-1} \int_0^1 g_i(x) dx}{\log n} = \frac{4}{3}$ .

This follows from the two previous lemmas.

**Lemma 2.10.**  $\sum_{i=u_{n-1}}^{u''_n} \int_0^1 \chi_{B(0, \frac{1}{i})}(R^i(x)) dx = o(\int g_n)$ .

*Proof.* By construction  $u''_n \leq \log(u_{n-1})^4 \log(u_{n-1} \log(u_{n-1})^4)^4 u_{n-1}$ . This implies that for big enough  $n$  we have  $\sum_{i=u_n}^{u''_n} \int_0^1 \chi_{B(0, \frac{1}{i})}(R^i(x)) dx \leq \log \log(u_n)^9$ .  $\square$

**Lemma 2.11.** For almost every  $x$  we have  $\sum_{n=1}^N \sum_{i=u_{n-1}}^{u''_n} \chi_{B(0, \frac{1}{i})}(R^i x) = o(\sum_{n=1}^N \int_0^1 g_n(x) dx)$ .

*Proof.* The lemma follows by the Borel-Cantelli Theorem, Lemma 2.10 and Corollary 2.9 because  $\sum_{n=1}^{\infty} \frac{\log(n)^k}{n^{\frac{4}{3}}}$  converges for any  $k$ .  $\square$

**2.4. Describing  $g_n$ .** The next proposition shows that for most points  $g_n(x)$  is not too large.

**Proposition 2.12.**  $\sum_{n=1}^{\infty} \lambda(\{x : g_n(x) > v_n^{\frac{2}{3}}\}) < \infty$ .

Define the above set to be  $B_n$ . Let  $C_k = \{x : \sum_{i=q_k}^{q_{k+1}} \chi_{[0, \frac{1}{i}]}(R^i(x)) > \sqrt[3]{a_{k+1}}\}$ , and an easier to understand set  $D_k = \bigcup_{i=0}^{q_k-1} B(R^{-i}(0), \frac{1}{\sqrt[3]{(a_{k+1})^2 q_k}})$ . To prove the proposition we need a pair of Lemmas:

**Lemma 2.13.**  $C_k \subset D_k$ .

It is a straightforward calculation (similar to Lemma 2.3) that if  $x \notin D_k$  then  $x \notin C_k$ .

**Lemma 2.14.**  $\lambda(\bigcap_{i=1}^r D_{k_i}) \leq 3^r \prod_{i=1}^r \lambda(D_{k_i})$ .

*Proof.* This follows by induction. Let  $k_1, \dots, k_r$  be an increasing sequence of natural numbers. Assume  $\bigcap_{i=1}^r D_{k_i}$  is the union of at most  $3^r \lambda(\bigcap_{i=1}^r D_{k_i}) q_{k_r}$  intervals of size at most  $\frac{2}{a_{k_r+1}^{\frac{2}{3}} q_{k_r}}$ . We intersect this set with  $D_u$  for  $u > k_r$ . Observe that  $D_u$  is the union of intervals that are at least  $\frac{1}{2q_u}$  separated. Therefore each interval in  $\bigcap_{i=1}^r D_{k_i}$  intersects at most  $\frac{2}{a_{k_r+1}^{\frac{2}{3}} q_{k_r}} 2q_u + 1$  of them. The proof follows.  $\square$

*Proof of Proposition 2.12.* Notice that by Lemma 2.3 if  $x \in B_n$  then  $x \in cv_n^{\frac{1}{12}}$  different  $D_k$  where  $a_{k+1} > v_n^{\frac{2}{3}}$ . Using Lemma 2.14 it follows that

$$\lambda(\bigcap_{i=1}^r D_{k_i}) \leq 3^r \prod_{i=1}^r \lambda(D_{k_i}) \leq 3^r 2((v_n^{\frac{2}{3}})^{-\frac{2}{3}})^r.$$

The Proposition follows by the size of  $3^{v_n^{\frac{1}{12}}} \prod_{i=1}^r \lambda(D_{k_i}) (v_n^{-v_n^{-1}})$  which is less than

$$3^{v_n^{\frac{1}{12}}} 2((v_n^{\frac{2}{3}})^{-\frac{2}{3}})^{v_n^{\frac{1}{12}}} (n^{\frac{1}{3}})^{v_n^{\frac{1}{12}}} / v_n^{\frac{1}{12}}!.$$

$\square$

The next proposition shows that large values of  $g_n$  do not contribute most of its mass.

**Proposition 2.15.**  $\lim_{n \rightarrow \infty} \frac{\int_{(0,1) \setminus B_n} g_n(x) dx}{\int_0^1 g_n(x) dx} = 1$ .

*Proof.* Observe that if  $\lambda(\bigcap_{i=1}^r D_{k_i}) = c$  and  $a_{k+1} > v_n^{\frac{2}{3}}$  then  $\lambda(\bigcap_{i=1}^{r+1} D_{k_i}) < \frac{3c}{v_n^{\frac{2}{3}}}$ . Also if  $x$  is in exactly  $r$  of the  $A_{k_i}$  and  $y$  is in the same ones and one more, then

$|\sum_{i=1}^{r+1} \left( \sum_{j=q_{k_i}}^{q_{k_i+1}} \chi_{[0, \frac{1}{j}]}(R^j(x)) - \sum_{j=q_{k_i}}^{q_{k_i+1}} \chi_{[0, \frac{1}{j}]}(R^j(y)) \right)| < v_n^{\frac{7}{12}}$ . The result follows from the fact that

$$\sum_{r=n^{\frac{1}{12}}}^{n^{\frac{1}{3}}} \binom{v_{n+1} - v_n}{r} 3^r r v_n^{\frac{7}{12}} \left( \frac{1}{v_n^{\frac{9}{4}}} \right)^r$$

is small.  $\square$

The remainder of this section is devoted to showing the next two propositions.

**Proposition 2.16.** *For almost every  $x$  we have*

$$\liminf_{N \rightarrow \infty} \frac{\sum_{i=1}^N g_i(x)}{\sum_{i=1}^N \int_0^1 g_i(x) dx} = \liminf_{M \rightarrow \infty} \frac{\sum_{i=1}^M \chi_{B(0, \frac{1}{i})}(R^i(x))}{\sum_{i=1}^M \frac{2}{i}}.$$

*Proof.*

$$\liminf_{N \rightarrow \infty} \frac{\sum_{i=1}^N g_i(x)}{\sum_{i=1}^N \int_0^1 g_i(x) dx} \leq \liminf_{N \rightarrow \infty} \frac{\sum_{i=1}^N g_i(x)}{\sum_{i=1}^N \frac{2}{i}} \leq \liminf_{M \rightarrow \infty} \frac{\sum_{i=1}^M \chi_{B(0, \frac{1}{i})}(R^i(x))}{\sum_{i=1}^M \frac{2}{i}}$$

for all  $x$ . The first inequality holds because everything is positive and because of Lemmas 2.6 and 2.10. The second inequality holds by Lemma 2.10.

The other direction follows from Lemma 2.11.  $\square$

**Proposition 2.17.** *For almost every  $x$  we have*

$$\limsup_{N \rightarrow \infty} \frac{\sum_{i=1}^N g_i(x)}{\sum_{i=1}^N \int_0^1 g_i(x) dx} = \limsup_{N \rightarrow \infty} \frac{\sum_{i=1}^N \chi_{B(0, \frac{1}{i})}(R^i(x))}{\sum_{i=1}^N \frac{2}{i}}.$$

Just as Proposition 2.16 follows from Lemmas 2.10 and 2.11, so too does Proposition 2.17.

## 2.5. Independence.

**Proposition 2.18.**  $\sum_{i \neq j}^{\infty} |\int_0^1 g_i(x) g_j(y) dx - \int_0^1 g_i(x) dx \int_0^1 g_j(x) dx| < \infty$ .

*Proof.* Consider  $P_{v'_{n+1}}$ . Observe that if  $k \leq n$  then by Lemma 2.4 the function  $g_k$  is within at most  $\frac{4u_k}{u'_{n+1}}$  of a function constant on intervals of  $P_{v'_{n+1}}$ . This is less than  $\frac{1}{n^{2+\frac{4}{9}}}$  by our choice  $u'_n$  and the Diophantine condition on  $\alpha$ . (Notice using  $u'_n a_{v_{n+1}} \geq n^4 u_{n-1}$  and  $u'_n a_{v_{n+1}} \leq u'_n a_{n^{\frac{4}{3}+1}} \leq u'_n n^{\frac{4}{3}(1+\frac{1}{6})}$  implies  $u'_n \geq u_{n-1} n^{2+\frac{4}{9}}$ .)

Consider  $P_{v'_n}$ . It is made up of intervals of two sizes – that is, there are two Rokhlin towers,  $A_{v'_{n+1}}$  and  $B_{v'_{n+1}}$ . Let us take two elements of this partition with the same length,  $I$  and  $J$ . There exists  $t$  such that  $I = J + R^t(0)$  (because they are in the same Rokhlin tower). Therefore a hit to  $J$  is followed by a hit to  $I$ . The difference of these balls is  $\frac{1}{i} - \frac{1}{i+t}$  where  $i$  is such that  $B(R^i(0), \frac{1}{i}) \cap I \neq \emptyset$ . If  $k > n$  then by Lemma 2.5 if  $I_1$  and  $I_2$  are intervals of the same size  $g_k|_{I_1}$  and  $g_k|_{I_2}$  differ

by at most  $\sum_{N=u''_n}^{u_n} \frac{1}{N} - \frac{1}{N+u'_n}$  (in  $L_1$  norm) when thought of functions of  $[0, |I_1|]$ . This is proportional to  $\frac{u'_n}{u''_n} \leq \frac{1}{n^{2+\frac{4}{9}}}$ .

**Corollary 2.19.** *If  $h$  is a positive function constant on intervals of  $P_{v'_{n+1}}$  and supported on  $A_{v'_{n+1}}$  (or  $B_{v'_{n+1}}$ ) and  $j > n$  then*

$$\left| \int_0^1 h(x)g_j(x)dx - \int_0^1 h(x)dx \int_0^1 g_j(x)dx \right| < \int_0^1 h(x) \frac{4u'_j}{u''_j} dx.$$

We make one more observation. Let  $I$  and  $J$  be intervals of two different sizes in  $P_{v'_k}$ . Then  $\left| \frac{\int_I(g_j)}{\lambda(J)} - \frac{\int_I(g_j)}{\lambda(I)} \right| < 6 \frac{2u'_k}{u''_j}$ . To prove this we use the following result.

**Theorem 7.** (Kesten [11, Theorem 1])  $\{0, \dots, R^{q_t-1}(0)\}$  has exactly one element in each  $[\frac{j}{q_t}, \frac{j+1}{q_t}]$ .

Therefore the number of hits of  $\{0, \dots, R^{q_t-1}(0)\}$  to any interval  $U$  is between  $q_t|U| - 3$  and  $q_t|U| + 3$ . Also there are at least  $\frac{u''_j}{2u'_k}$  hits to each  $I$  or  $J$ .

Assume  $i < j$  and consider two fixed partition intervals  $I, J$  of  $P_{v'_j}$ . As a consequence of the above argument

$$(2) \quad \left| \int_0^1 g_i(x)g_j(x)dx - \int_0^1 g_i(x)dx \int_0^1 g_j(x)dx \right| \leq \int_0^1 g_i(x)dx \left( \frac{1}{v_j^{2+\frac{5}{9}}} \frac{4}{v_i''} \right) + \int_I g_j(x) \frac{u_i}{u_{j-1}} dx + 6 \frac{u''_j}{2u'_j}.$$

This is done by changing  $g_i$  to be constant on intervals of  $P_{v'_j}$  appealing to Corollary 2.19 and the previous argument. It follows that

$$\sum_{i \neq j}^{\infty} \left| \int_0^1 g_i(x)g_j(x)dx - \int_0^1 g_i(x)dx \int_0^1 g_j(x)dx \right| < \sum_{i < j}^{\infty} \frac{4}{j^{2+\frac{5}{9}}} \int_0^1 g_i(x)dx \leq \sum_{j=1}^{\infty} j \frac{1}{j^{2+\frac{1}{9}}}$$

and Proposition 2.18 follows.  $\square$

## 2.6. The probabilistic theorem.

**Theorem 8.** *Let  $R_1, R_2, \dots$  be positive random variables such that*

$$\max R_n(\omega) \leq \left( \sum_{i=1}^{n-1} \int_{\Omega} R_i d\nu \right)^{\frac{2}{3}}$$

$$\sum_{i \neq j}^{\infty} \int_{\Omega} R_i(\omega) R_j(\omega) d\nu < C < \infty$$

and

$$\int_{\Omega} R_n d\nu > \frac{1}{2} \left( \sum_{i=1}^{n-1} \int_{\Omega} R_i d\nu \right)^{\frac{1}{3}}$$

then for almost every  $\omega$  we have

$$\lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N R_i(\omega)}{\sum_{i=1}^N \int_{\Omega} R_i(\nu) d\nu} = 1.$$

**Theorem 9.** (Chebyshev's inequality) Let  $R$  be a random variable with  $\int R d\mu = 0$  and finite variance then  $\mu(\{\omega : R(\omega) > c\}) \leq \frac{\int R^2 d\mu}{c^2}$ .

*Proof of Theorem 8.* This is a straightforward proof by the method of subsequences.

Consider  $(\sum_{i=1}^N R_i - \int_{\Omega} R_i d\nu)^2$ . This is a non-negative random variable with

$$\int_{\Omega} (\sum_{i=1}^N |R_i - \int_{\Omega} R_i d\nu|)^2 d\nu \leq \int_{\Omega} \sum_{i=1}^N R_i^2 + (\int_{\Omega} R_i d\nu)^2 d\nu + C \leq 2NN^{\frac{4}{3}} = 2N^{\frac{7}{3}}.$$

With the last two inequalities holding for all large enough  $N$ . We will ignore  $C$  in the following arguments. By Chebyshev's inequality  $\nu(\{\omega : |\sum_{i=1}^N R_i(\omega) - \int_{\Omega} R_i d\nu| \geq \epsilon N^{\frac{4}{3}}\}) \leq \epsilon^2 N^{-\frac{1}{3}}$ . It follows that along the subsequence  $N_i = \lfloor i^{\frac{8}{3}} \rfloor$  almost every  $\omega$  has

$$\lim_{i \rightarrow \infty} \frac{\sum_{t=1}^{N_i} R_t(\omega) - \int_{\Omega} R_t d\nu}{\sum_{i=1}^{N_i} \int_{\Omega} R_t(\omega) d\nu} = 0.$$

Next we estimate  $\nu(\{\omega : \max_{i^{\frac{8}{3}} < j < (i+1)^{\frac{8}{3}}} \sum_{t=i^{\frac{8}{3}}}^j R_t(\omega) - \int_{\Omega} R_t d\nu \geq \epsilon(i^{\frac{8}{3}})^{\frac{4}{3}}\})$  by Chebyshev's inequality. Notice that  $\sum_{t=i^{\frac{8}{3}}}^{(i+1)^{\frac{8}{3}}} (\int_{\Omega} (R_t(\omega) - \int_{\Omega} R_t(z) d\nu(z)) d\nu(\omega))^2 \leq 4i^2 (i^{\frac{8}{3}})^{\frac{4}{3}}$ .

This implies that

$$\nu(\{\omega : \sum_{t=i^{\frac{8}{3}}}^j R_t(\omega) - \int_{\Omega} R_t d\nu \geq \epsilon(i^{\frac{8}{3}})^{\frac{4}{3}}\}) \leq \frac{i^{\frac{50}{9}}}{i^{\frac{64}{9}}}$$

which has convergent sum. It follows that

$$\lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N R_i(\omega) - \int_{\Omega} R_i(\omega) d\nu}{\sum_{i=1}^N \int_{\Omega} R_i(\omega) d\nu} = 0$$

for  $\mu$ -a.e.  $\omega$  and thus

$$\lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N R_i(\omega)}{\sum_{i=1}^N \int_{\Omega} R_i(\omega) d\nu} = 1$$

for  $\nu$  a.e.  $\omega$ . □

*Proof of Theorem 1.* Propositions 2.16 and 2.17 show that it suffices to prove convergence of the sums of  $g_i$ . This follows from Theorem 8 and Proposition 2.18.  $\square$

### 3. UNDETERMINED POINTS AS A SHRINKING TARGET

In this section we consider another shrinking target problem for rotations, but one whose target arises in a very different way. Let  $\mathcal{P} = \{A_0, A_1\}$  be the partition of  $[0, 1)$  given by  $A_0 = [0, \alpha)$ ,  $A_1 = [\alpha, 1)$ . For a sequence  $c_0, c_1, \dots$  (finite or infinite) of 0's and 1's, let  $C_{c_0, \dots} = \{x \in X : T^i x \in A_{c_i} \text{ for all } i\}$  and let  $\Sigma$  be the set of finite codings  $c_0, c_1, \dots, c_n$  which actually occur, i.e. for which  $C_{c_0, \dots, c_n} \neq \emptyset$ . Let  $V_j = \{x : x \in C_{c_0, \dots, c_j} \text{ and such that } c_0, \dots, c_j, 0 \text{ and } c_0, \dots, c_j, 1 \in \Sigma\}$ . This is the set of ‘undetermined’ points at step  $j$ , that is, all points whose coding up to step  $j$  does not determine the coding at step  $j + 1$ . We want to find asymptotics on how often a point is undetermined; specifically, we will prove

**Theorem 10.** *For almost all  $x \in [0, 1)$  and almost all  $\alpha$ ,*

$$\frac{\log \sum_{j=1}^n \chi_{V_j}(x)}{\log \sum_{j=1}^n \lambda(V_j)} \rightarrow 1$$

as  $n \rightarrow \infty$ .

To understand why Theorem 10 constitutes a shrinking target problem, consider the following. Let  $\mathcal{P}_j = \bigvee_{k=0}^j R_\alpha^k \mathcal{P}$ , the partition generated by  $\mathcal{P}$  and its first  $j$  translations. For  $x \in X$ , denote by  $[[x]]_j$  the atom of  $x$  in  $\mathcal{P}_j$ . The coding  $c_0, \dots, c_j$  determines only the atom  $[[R_\alpha^j x]]_j$ . A point  $x$  will belong to  $V_j$  if and only if  $R_\alpha^j x$  is in  $[[1 - \alpha]]_j$  as the image of this atom under one more rotation contains points in both  $A_0$  and  $A_1$ . We will denote  $[[1 - \alpha]]_j$  by  $U_j$  – these are the shrinking targets which we are trying to hit. Note that  $U_j = R_\alpha^j(V_j)$ .

Before turning to the proof of Theorem 10, we give an argument as to why we there is no stronger theorem along the lines of convergence for

$$(3) \quad \frac{\sum_{j=1}^n \chi_{V_j}(x)}{\sum_{j=1}^n \lambda(V_j)}.$$

We begin with a proposition proving the existence of very large elements  $a_n$  for the continued fraction expansion and use this to show that, for very long stretches of time certain points are undetermined more often than  $\sum_{j=1}^n \lambda(V_j)$  predicts.

**Proposition 3.1.** *For any  $C \in \mathbb{R}$  and almost every  $\alpha$  there exists infinitely many  $m$  such that*

$$a_m > C \sum_{i=1}^{m-1} a_i.$$

The following lemma appears in [12, page 60].

**Lemma 3.2.** *For any  $n, b_1, \dots, b_n \in \mathbb{N}$  we have*

$$\frac{1}{3b_n^2} < \frac{\lambda(\{\alpha : a_1(\alpha) = b_1, \dots, a_n(\alpha) = b_n\})}{\lambda(\{\alpha : a_1(\alpha) = b_1, \dots, a_{n-1}(\alpha) = b_{n-1}\})} < \frac{2}{b_n^2}.$$

**Corollary 3.3.**

$$\frac{1}{10b_n} < \frac{\lambda(\{\alpha : a_1(\alpha) = b_1, \dots, a_n(\alpha) \geq b_n\})}{\lambda(\{\alpha : a_1(\alpha) = b_1, \dots, a_{n-1}(\alpha) = b_{n-1}\})} < \frac{2}{b_n}.$$

Let  $W_n = \{\alpha : \sum_{i=1}^n a_i(\alpha) < 10n \log n\}$ .

**Lemma 3.4.**  $\lambda(W_n) > \frac{1}{10}$ .

*Proof.* Let  $A_n = \{\alpha : a_i(\alpha) < n^2 \text{ for all } i\}$ . Consider  $\sum_{i=1}^n \int_{A_n} a_i(\alpha) d\lambda$ . By Lemma 3.2 this is dominated by  $\sum_{i=1}^n \sum_{i=1}^{n^2} \frac{2}{i^2} i$ . This is less than or equal to  $2n(1 + \log n^2) < 5n \log n$ . The Lemma follows from Markov's inequality and the fact that Lemma 3.2 implies that  $\lambda(A_n) > 1 - \frac{2}{n}$ .  $\square$

**Lemma 3.5.** For a set of  $\alpha$  of measure at least  $\frac{1}{40}$  we have  $\sum_{n \text{ s.t. } \alpha \in W_n} \frac{1}{Cn \log n} = \infty$ .

*Proof.* Consider  $\sum_{i=1}^n \int_{W_i} \frac{1}{Ci \log i}$ . This is at least  $\frac{\log(n)}{20}$ . Also

$$\max_{\alpha} \sum_{i \text{ s.t. } \alpha \in W_i} \frac{1}{Cn \log(n)} < 2 \log n.$$

Therefore we obtain the Lemma for a set of  $\alpha$  of measure at least  $\frac{1}{40}$ .  $\square$

Let the set of such  $\alpha$  be denoted  $S$ .

*Proof of Proposition 3.1.* Given that  $\alpha$  is in  $W_{m-1}$  the Lemma 3.2 implies that the probability that  $a_m(\alpha) \geq C \sum_{i=1}^{m-1} a_i(\alpha)$  is at least  $\frac{1}{60CN \log N}$  independent of the past outcomes. By the previous Lemma, if  $\alpha \in S$  this diverges. For any sequence of sets of finite measure  $\{B_i\}_{i=1}^{\infty}$  where there exists  $c > 0$  such that  $\lambda(B_i \cap B_j) > c\lambda(B_i)\lambda(B_j)$ , one has  $\lambda(\bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} B_i) > 0$ . Using this, we find that there is a positive measure set of  $\alpha$  such that

$$a_m > C \sum_{i=1}^{m-1} a_i$$

infinitely often. The set of such  $\alpha$  is Gauss map invariant and therefore has full measure.  $\square$

We need the following two Lemmas on the shrinking targets  $U_j$  to complete our proof of non-convergence for sums like (3). These Lemmas can be obtained using the partial fraction expansion of  $\alpha$ . We will denote by  $[y]$  the value modulo 1 of a real number  $y$  and by  $\langle\langle y \rangle\rangle$  the distance from  $y$  to the nearest integer.

**Lemma 3.6.** *Let*

$$\begin{aligned} r_j &= \max\{q_k : q_k \leq j\} \\ s_j &= \max\{q_k : q_{k+1} \leq j\} \\ t_j &= \max\{T \in \mathbb{N} : s_j + Tr_j \leq j\}. \end{aligned}$$

*Then*

$$R_{\alpha}(U_j) = [[s_j\alpha] + t_j[r_j\alpha], [r_j\alpha]]$$

*or*

$$R_{\alpha}(U_j) = [[r_j\alpha], [s_j\alpha] - t_j(1 - [r_j\alpha])],$$

and

$$\lambda(U_j) = \lambda(V_j) = \langle\langle r_j \alpha \rangle\rangle + \langle\langle s_j \alpha \rangle\rangle - t_j \langle\langle r_j \alpha \rangle\rangle.$$

*Remark 3.7.* Note that if  $r_j = q_k$ ,  $s_j = q_{k-1}$  and  $t_j = a_{k+1}$ .

*Proof.* The numbers  $r_j$  and  $s_j$  are the denominators of the best and second-best rational approximations to  $\alpha$  (respectively) with denominator less than or equal to  $j$ .

CASE 1:  $0 < [r_j \alpha] < 1/2$ . As the convergents alternate in approximating  $\alpha$  from above and below,  $1/2 < [s_j \alpha] < 1$ . The only improvement possible in  $[r_j \alpha]$  as an upper bound for  $R_\alpha(U_j)$  would come from finding some  $l$  with  $\langle\langle l \alpha \rangle\rangle < \langle\langle r_j \alpha \rangle\rangle$ . This is not possible for  $l \leq j$  as  $r_j$  is the denominator for the best approximation to  $\alpha$  with denominator  $\leq j$ . Thus the upper endpoint of  $R_\alpha(U_j)$  is  $[r_j \alpha]$  as desired.

The lower bound on  $R_\alpha(U_j)$  given by  $[s_j \alpha]$  can be improved only by adding  $[r_j \alpha]$  some number of times, as  $r_j$  is the only integer  $\leq j$  with  $\langle\langle r_j \alpha \rangle\rangle < \langle\langle s_j \alpha \rangle\rangle$ . The lower will thus be of the form  $y = [s_j \alpha] + T[r_j \alpha]$  and will be found by taking  $T$  as large as possible such that the  $s_j + Tr_j$  rotations required to produce this point do not exceed  $j$ ; this number is  $t_j$ .

We calculate  $\lambda(U_j) = \langle\langle r_j \alpha \rangle\rangle + (1 - [s_j \alpha] - t_j[r_j \alpha])$  as  $\langle\langle r_j \alpha \rangle\rangle = [r_j \alpha]$ . Since  $\langle\langle s_j \alpha \rangle\rangle = 1 - [s_j \alpha]$ , this simplifies to the desired result.

CASE 2:  $1/2 < [r_j \alpha] < 1$ . Then  $0 < [s_j \alpha] < 1/2$  and the lower endpoint of  $R_\alpha(U_j)$  is  $[r_j \alpha]$ . As before, the upper endpoint is of the form  $[s_j \alpha] - T(1 - [r_j \alpha])$ . The best such endpoint is found by taking  $T$  as large as possible, i.e. equal to  $t_j$ .

Finally, we calculate again

$$\begin{aligned} \lambda(U_j) &= \langle\langle s_j \alpha \rangle\rangle - t_j(1 - [r_j \alpha]) + (1 - [r_j \alpha]) \\ &= \langle\langle s_j \alpha \rangle\rangle - t_j \langle\langle r_j \alpha \rangle\rangle + \langle\langle r_j \alpha \rangle\rangle. \end{aligned}$$

□

**Lemma 3.8.** *For any  $i \in \mathbb{N}$ , for all  $l \in [q_i, q_{i-1} + q_i]$  or  $[q_{i-1} + (b-1)q_i, q_{i-1} + bq_i]$  where  $b \in [2, a_{i+1}]$ , the sets  $V_l = R_\alpha^{-l}U_l$  are disjoint.*

*Proof.* Let  $J$  denote an interval of the form given in the statement of the Lemma. As atoms of the sequence of partitions  $\mathcal{P}_j$ , the sets  $U_l$  change only when the orbit of 0 hits  $U_l$ . By the description of Lemma 3.6 this does not happen over the interior of any of the intervals  $J$ . Suppose that  $R_\alpha^{-l}U_l \cap R_\alpha^{-k}U_k \neq \emptyset$ ; with  $l, k \in J$ . Note that for such  $l$  and  $k$ ,  $U_l = U_k$ . Suppose  $l > k$  and we obtain  $U_l \cap R_\alpha^{l-k}U_k \neq \emptyset$ . However, the endpoints of  $U_l = U_k$  are points in the orbit of zero which are reached by step  $q_i$  at the latest so for  $R_\alpha^{l-k}U_k$  to intersect  $U_l$  would provide another point in the orbit hitting  $U_l$  before the time given by the right endpoint of the interval  $J$ . This contradicts the description of Lemma 3.6. □

Partition  $\mathbb{N}$  into the collection of all intervals  $J$  described above. Index them as  $\{J_m\}$ . Note that  $|J_m|$  is non-decreasing in  $m$  and strictly increases as we cross the integers  $q_i$ . In fact  $|J_m| = q_{i-1}$  or  $q_i$  (if  $J_m$  is part of the partition of  $I_i = [q_i, q_{i+1})$ ).

Consider the integers  $[q_{m-1}, q_m)$  for some  $m$  satisfying Proposition 3.1; this is a very large proportion of the interval  $[0, q_m)$ . We will find that there are positive measure subsets of points for which the numerators of quotients of the form (3)

over the range  $[q_{m-1}, q_m)$  differ by a factor of  $a_m$ , which is large compared to the value of  $\sum \chi_{V_j}$  up to time  $q_{m-1}$ .

Let  $j_0 = q_{m-1} + q_{m-2}$ . By the description of Lemma 3.6,  $V_{j_0}$  and  $V_{j_0+cq_{m-1}}$  intersect for all  $c < a_m$ . Points  $x$  that lie in the intersection of the  $V_{j_0+cq_{m-1}}$  for all such  $c$  satisfy  $\sum_{j=q_{m-1}}^{q_m-1} \chi_{V_j}(x) = a_m$ , whereas there is an interval of points  $y$  which lie only in  $V_{j_0}$  itself yielding  $\sum_{j=q_{m-1}}^{q_m-1} \chi_{V_j}(y) = 1$ . As  $a_m$  is much larger than  $m + \sum_{i=1}^{m-1} a_i$  – the largest possible value for  $\sum_{j=1}^{q_m-1} \chi_{V_j}(y)$  by Lemma 3.8 – convergence of the quotient (3) for almost every point fails. We sum up this failure by an arbitrary factor of  $C$ :

**Theorem 11.** *For almost all  $x \in [0, 1)$  and almost all  $\alpha$ ,*

$$\limsup_{n \rightarrow \infty} \frac{\sum_{j=1}^n \chi_{V_j}(x)}{\sum_{j=1}^n \lambda(V_j)} = \infty$$

and

$$\liminf_{n \rightarrow \infty} \frac{\sum_{j=1}^n \chi_{V_j}(x)}{\sum_{j=1}^n \lambda(V_j)} = 0.$$

Thus Theorem 10 is our best hope for this problem.

Towards Theorem 10, we claim the following set of inequalities:

$$(4) \quad C_1 n (\log n)^3 > \sum_{i=1}^n a_i \geq \sum_{j=1}^{q_n} \chi_{V_j}(x) > \frac{1}{2} n$$

for some positive constant  $C_1$  and for almost every  $\alpha$  and  $x \in [0, 1)$ .

**Lemma 3.9.** *For almost every  $\alpha$ , there exists a positive constant  $C_1$  such that  $C_1 n (\log n)^3 > \sum_{i=1}^n a_i$ .*

*Proof.* Observe that if  $G_n = \{\alpha : a_i(\alpha) \leq n^2 \ \forall i \leq n\}$  then  $\int_{G_n} \sum_{i=1}^n a_i(\alpha) d\lambda(\alpha) \leq 6n \log n$ . Also a.e.  $\alpha \in G_n$  for all but finitely many  $n$ . It follows from Markov's inequality that

$$\lambda\left(\left\{\alpha : \sum_{i=1}^n a_i(\alpha) \leq 10n(\log n)^{2+\epsilon}\right\}\right) \leq \frac{1}{\log n}^{1+\epsilon}.$$

It follows that a.e.  $\alpha$  has that  $\sum_{i=1}^{10^k} a_i(\alpha) \leq k^{2+\epsilon} 10^k$  for all but finitely many  $k$ .

This implies the lemma because for all large enough  $k$  we have  $10^{k-1} (\log 10^{k-1})^3 \geq 10^k (\log 10^k)^2$ .  $\square$

**Lemma 3.10.** *For every  $x \in [0, 1)$  and any  $\alpha$ ,  $\sum_{i=1}^n a_i \geq \sum_{j=1}^{q_n} \chi_{V_j}(x)$ .*

*Proof.* Each interval of integers  $I_i = [q_i, q_{i+1})$  is subdivided into  $a_i$  subintervals  $J_n$  as described in Lemma 3.8. As that Lemma shows, over each  $J_n$  the sets  $V_j$  are disjoint and hence can contribute at most one to  $\sum_{j=1}^{q_n} \chi_{V_j}(x)$ .  $\square$

For each  $i$ , we have the interval of integers  $I_i = [q_i, q_{i+1})$ , subdivided (as in Lemma 3.8) into subintervals  $J_n$ . Let us denote by  $J_2^i$  the second of these intervals

for each  $i$  – specifically,  $J_2^i = [q_i + q_{i-1}, 2q_i + q_{i-1})$ . We remark that when  $a_{i+1} = 1$ ,  $J_2^i$  is  $[q_{i+1}, q_{i+1} + q_i)$  and is actually a subinterval of  $I_{i+1}$ . Nonetheless, the collection  $\{J_2^i\}$  consists of pairwise disjoint intervals.

We will give a lower bound on  $\sum_{j=1}^{q_n} \chi_{V_j}(x)$  by bounding below the sum over just the  $J_2^i$ . Towards this end, let

$$h_i(x) = \sum_{j \in J_2^i} \chi_{V_j}(x).$$

**Lemma 3.11.** *For all  $i$ ,*

$$\int_{[0,1)} h_i(x) d\lambda > 1/2.$$

*Proof.* As per Lemma 3.8, over  $J_2^i$ , the  $V_j$  are disjoint, so  $h_i(x) \in \{0, 1\}$ . The length of the interval  $J_2^i$  is  $q_i$ , and for  $j \in J_2^i$ ,

$$\lambda(V_j) = \langle \langle q_{i-1} \alpha \rangle \rangle,$$

using the description of  $R_\alpha(U_j)$  provided by Lemma 3.6. By Theorem 13 in [12],  $\langle \langle q_{i-1} \alpha \rangle \rangle > \frac{1}{q_{i-1} + q_i}$ . We may then bound the interval below by

$$\int_{[0,1)} h_i(x) d\lambda > \frac{q_i}{q_i + q_{i-1}} > \frac{q_i}{2q_i} = \frac{1}{2}.$$

□

We prove with the following sequence of results that visits to the sets counted by the functions  $h_i$  are (approximately) independent events.

**Lemma 3.12.** *Let  $[c, d) \subset [0, 1)$ . Let  $f_{[c,d)}(m) = \#\{[c, d) \cap \cup_{l \in J_m} R_\alpha^{-l}(0)\}$ . Then*

$$\lambda([c, d)) |J_m| - 1 \leq f_{[c,d)}(m) \leq \lambda([c, d)) |J_m| + 1.$$

*Proof.* This follows immediately from Kesten’s Theorem 7, by counting how many times the left endpoints of intervals  $[\frac{j}{q_i}, \frac{j+1}{q_i})$  intersect  $[c, d)$ . □

**Proposition 3.13.** *For sufficiently large  $m$  (relative to  $i$ )*

$$\begin{aligned} & \left( \frac{\lambda(V_i) |J_m| - 2}{\lambda(V_i) |J_m|} \right) \lambda(V_i) \lambda(\cup_{l \in J_m} V_l) \\ & \leq \lambda(V_i \cap \bigcup_{l \in J_m} V_l) \\ & \leq \left( \frac{\lambda(V_i) |J_m| + 2}{\lambda(V_i) |J_m|} \right) \lambda(V_i) \lambda(\cup_{l \in J_m} V_l). \end{aligned}$$

*Proof.* Let  $m$  be so large that  $i \notin J_m$ . By the previous lemma, the interval  $V_i$  is hit by the left endpoints of the  $V_l$  between  $\lambda(V_i) |J_m| - 1$  and  $\lambda(V_i) |J_m| + 1$  times. As the sets  $V_l$  are disjoint over  $l \in J_m$ , this easily yields

$$(\lambda(V_i) |J_m| - 2) \lambda(V_i) \leq \lambda(V_i \cap \bigcup_{l \in J_m} V_l) \leq (\lambda(V_i) |J_m| + 2) \lambda(V_i).$$

This holds for any  $l \in J_m$  as all have the same measure. As  $|J_m| \lambda(V_i) = \lambda(\cup_{l \in J_m} V_l)$  this equation is close to asserting independence – we need only account for the errors involving the  $\pm 2$ . Translating this to an inequality with multiplicative errors yields

$$\begin{aligned}
& \left( \frac{\lambda(V_i)|J_m| - 2}{\lambda(V_i)|J_m|} \right) \lambda(V_i) \lambda(\cup_{l \in J_m} V_l) \\
& \leq \lambda(V_i \cap \bigcup_{l \in J_m} V_l) \\
& \leq \left( \frac{\lambda(V_i)|J_m| + 2}{\lambda(V_i)|J_m|} \right) \lambda(V_i) \lambda(\cup_{l \in J_m} V_l).
\end{aligned}$$

□

By using the above inequality for all  $i \in J_n$  where  $n < m$  we get the following corollary. It relates to calculating the correlation between a point being undetermined in the intervals  $J_n$  and  $J_m$ .

**Corollary 3.14.** *For any  $i \in J_n$ , and  $J_n, J_m$  disjoint,  $n < m$*

$$\begin{aligned}
& \left( \frac{\lambda(V_i)|J_m| - 2}{\lambda(V_i)|J_m|} \right) \lambda(\cup_{i \in J_n} V_i) \lambda(\cup_{l \in J_m} V_l) \\
& \leq \lambda\left(\bigcup_{i \in J_n} V_i \cap \bigcup_{l \in J_m} V_l\right) \\
& \leq \left( \frac{\lambda(V_i)|J_m| + 2}{\lambda(V_i)|J_m|} \right) \lambda(\cup_{i \in J_n} V_i) \lambda(\cup_{l \in J_m} V_l).
\end{aligned}$$

*Proof.* This follows from the previous proposition, summing the inequalities over the disjoint sets  $V_i$  for  $i \in I_n$ . (The desire to compute this sum explains our preference for the formulation in terms of multiplicative bounds above.) □

**Proposition 3.15.** *For  $j > i$*

$$\left(1 - \frac{2q_{i-1}}{q_{j-1}}\right) \int h_i d\lambda \int h_j d\lambda \leq \int h_i h_j d\lambda \leq \left(1 + \frac{2q_{i-1}}{q_{j-1}}\right) \int h_i d\lambda \int h_j d\lambda.$$

*Proof.* First,

$$\int h_i(x) h_j(x) d\lambda = \int \left( \sum_{l \in J_2^i} \chi_{V_l}(x) \right) \left( \sum_{l \in J_2^j} \chi_{V_l}(x) \right) d\lambda.$$

As over  $J_2^i$  and over  $J_2^j$  the sets  $V_l$  are disjoint, the integrand of the above has value 0 or 1 according to whether  $x \in (\cup_{l \in J_2^i} V_l) \cap (\cup_{l \in J_2^j} V_l)$ . Thus, we are calculating

$$\lambda\left(\bigcup_{l \in J_2^i} V_l \cap \bigcup_{l \in J_2^j} V_l\right).$$

By Corollary 3.14, we get

$$\begin{aligned}
& \left( \frac{\lambda(V_l)|J_2^j| - 2}{\lambda(V_l)|J_2^j|} \right) \lambda(\cup_{l \in J_2^i} V_l) \lambda(\cup_{l \in J_2^j} V_l) \\
& \leq \lambda\left(\bigcup_{l \in J_2^i} V_l \cap \bigcup_{l \in J_2^j} V_l\right) \\
& \leq \left( \frac{\lambda(V_l)|J_2^j| + 2}{\lambda(V_l)|J_2^j|} \right) \lambda(\cup_{l \in J_2^i} V_l) \lambda(\cup_{l \in J_2^j} V_l)
\end{aligned}$$

For the term  $(1 \pm \frac{2}{\lambda(V_i)|J_2^j|})$  we use any  $l \in J_2^i$ . By Kesten's Theorem 7, since  $V_l$  is an atom in the partition by the first  $q_{i+1} - 1$  points of the orbit of 0,  $\lambda(V_l)$  has size at least  $\frac{1}{q_{i+1}}$ . Likewise for  $|J_2^j|$  we want a lower bound. From the description of these intervals given in the statement of Lemma 3.8,  $|J_2^j| \geq q_{j-1}$ . Using these two bounds, the multiplicative error terms in the above become  $(1 \pm \frac{2q_{i-1}}{q_{j-1}})$ .

Returning to our inequalities for  $\int h_i h_j$ , as the  $V_l$  are disjoint over  $J_2^j$  or  $J_2^i$ , we can translate back into integrals as so:

$$\begin{aligned} \left(1 - \frac{2q_{i-1}}{q_{j-1}}\right) \int \sum_{l \in J_2^i} \chi_{V_l}(x) d\lambda &\int \sum_{l \in J_2^j} \chi_{V_l}(x) d\lambda \\ &\leq \int h_i h_j d\lambda \leq \\ &\left(1 + \frac{2q_{i-1}}{q_{j-1}}\right) \int \sum_{l \in J_2^i} \chi_{V_l}(x) d\lambda \int \sum_{l \in J_2^j} \chi_{V_l}(x) d\lambda. \end{aligned}$$

These are the desired bounds on  $\int h_i h_j d\lambda$ . □

The independence result we want is the following.

**Proposition 3.16.** *There exist constants  $C, b > 0$  such that*

$$\int_{[0,1]} h_i(x) h_j(x) d\lambda - \int_{[0,1]} h_i(x) d\lambda \int_{[0,1]} h_j(x) d\lambda < C e^{-b|i-j|}.$$

*Proof.* We need to show that the expression

$$\frac{2q_{i-1}}{q_{j-1}} \int h_i d\lambda \int h_j d\lambda$$

decays exponentially in  $|i - j|$ . A clear upper bound on  $\int h_i d\lambda, \int h_j d\lambda$  is 1. The  $q_i$  satisfy the recursion relation  $q_{i+1} = a_{i+1}q_i + q_{i-1}$ . As the  $a_i$  are positive integers, the  $q_i$  grow exponentially (by comparison with the Fibonacci sequence, e.g.). Thus, the terms  $\frac{q_{i-1}}{q_{j-1}}$  decay exponentially in  $j - i$ , finishing the proof. □

We can apply this approximate independence to prove the remaining inequality in equation (4). Let  $\tilde{h}_i(x) = h_i(x) - \int h_i(x) d\lambda$ , and note that  $\tilde{h}_i(x) \in (-1, 1)$ . Let  $\tilde{s}_n(x) = \sum_{i=1}^n \tilde{h}_i(x)$ .

**Proposition 3.17.** *For almost every  $x \in S^1$ , for sufficiently large  $n$ ,*

$$\sum_{j=1}^{q_n} \chi_{V_j}(x) > \frac{1}{2}n.$$

*Proof.* First, for all  $x \in [0, 1)$ ,  $\sum_{j=1}^{q_n} \chi_{U_j}(x) \geq \sum_{i=1}^n h_i(x)$  so we will prove the inequality for the latter sum.

Consider  $\sum_{i=1}^n \int h_i(x) d\lambda$ . By Lemma 3.11 this is bounded below by  $\frac{1}{2}n$ ; it is bounded above by  $n$  as  $h_i$  takes only 1 or 0 as a value. Applying Chebyshev's inequality to  $\tilde{s}_n$  yields (for any  $\epsilon > 0$ )

$$\begin{aligned}
\lambda(\{x : |\tilde{s}_n(x)| > \epsilon n\}) &< \frac{\int \tilde{s}_n^2(x) d\lambda}{\epsilon^2 n^2} \\
&= \frac{\sum_{i=1}^n \int \tilde{h}_i^2(x) d\lambda + 2 \sum_{i < j} \int \tilde{h}_i(x) \tilde{h}_j(x) d\lambda}{\epsilon^2 n^2} \\
&< \frac{D}{\epsilon^2 n}.
\end{aligned}$$

For the last inequality we have used the facts that  $\tilde{h}_i(x) \in (-1, 1)$  so  $\sum_{i=1}^n \int \tilde{h}_i^2(x) d\lambda < n$  and that for some positive constant  $D$ ,  $2 \sum_{i < j} \int \tilde{h}_i \tilde{h}_j d\lambda < (D-1)n$  by Proposition 3.16.

We restrict our attention to the subsequence of times  $\{n^2\}$ , obtaining

$$\lambda(\{x : |\tilde{s}_{n^2}(x)| > \epsilon n^2\}) < \frac{D}{\epsilon^2 n^2}.$$

Summing the term on the right-hand side of the above inequality over all  $n$  yields a convergent series so by the Borel-Cantelli Lemma, for almost every  $x \in [0, 1)$ ,

$$\frac{\tilde{s}_{n^2}(x)}{n^2} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Consider now the intervals  $[n^2, (n+1)^2)$ . As  $\tilde{h}_i(x) \in (-1, 1)$ , for  $k \in [n^2, (n+1)^2)$ ,

$$|\tilde{s}_{n^2}(x) - \tilde{s}_k(x)| < 2n + 1$$

so

$$\frac{|\tilde{s}_k(x)|}{k} < \frac{|\tilde{s}_{n^2}(x)| + 2n + 1}{k} \leq \frac{|\tilde{s}_{n^2}(x)| + 2n + 1}{n^2} \rightarrow 0$$

as  $k \rightarrow \infty$ .

We have now that for almost all  $x$ ,

$$\frac{\sum_{i=1}^n h_i(x) - \int h_i(x) d\lambda}{n} \rightarrow 0.$$

As  $\sum_{i=1}^n \int h_i(x) d\lambda \in (\frac{1}{2}n, n)$ , for sufficiently large  $n$ ,  $\sum_{i=1}^n h_i(x) > \frac{1}{2}n$  as desired.  $\square$

We now prove a similar series of inequalities for  $\sum_{j=1}^{q_n} \lambda(V_j)$ , namely:

$$(5) \quad 2C_1 n (\log n)^2 > 2 \sum_{i=1}^n a_i > \sum_{j=1}^{q_n} \lambda(V_j) > \frac{1}{2}n.$$

The left-most inequality is Lemma 3.9. For the right-most:

**Lemma 3.18.** *For all  $\alpha$ ,*

$$\sum_{j=1}^{q_n} \lambda(V_j) > \frac{1}{2}n.$$

*Proof.* This follows easily from Lemma 3.11 after noting that

$$\sum_{j=q_i}^{q_{i+1}} \lambda(V_j) > \sum_{j \in J_2^i} \lambda(V_j) = \int_{[0,1)} h_i(x) d\lambda.$$

$\square$

It remains only to prove

**Lemma 3.19.** *For all  $\alpha$ ,*

$$2 \sum_{i=1}^n a_i > \sum_{j=1}^{q_n} \lambda(V_j).$$

*Proof.* We will show that over the interval  $I_i = [q_i, q_{i+1})$ ,  $\sum_{j \in I_i} \lambda(V_j)$  is bounded above by  $2a_{i+1}$ . We do so by considering each subinterval  $J_n \subset I_i$  individually.  $J_1 = [q_i, q_{i+1} + q_{i-1})$  has a length of  $q_{i-1}$ . Over this interval,  $\lambda(V_j) = \langle \langle q_{i-1} \alpha \rangle \rangle + \langle \langle q_i \alpha \rangle \rangle$ . This is bounded above by  $\frac{1}{q_i} + \frac{1}{q_{i+1}}$  by [12] Thm 9. The total contribution of  $J_1$  to the sum of  $\lambda(V_j)$ 's is thus bounded above by  $\frac{q_{i-1}}{q_i} + \frac{q_{i-1}}{q_{i+1}} < 2$ .

The intervals  $J_2, \dots, J_{a_{i+1}}$  each have length  $q_i$  and over each of them  $\lambda(V_j) < \langle \langle q_{i-1} \alpha \rangle \rangle < \frac{1}{q_i}$ . They thus each provide a contribution to the relevant sum of less than one and the result follows.  $\square$

The inequalities collected above enable us to prove the main theorem:

*Proof of Thm 10.* Suppose  $n \in [q_m, q_{m+1})$ . Then we have the following:

$$\begin{aligned} \frac{1}{2}m &< \sum_{j=1}^{q_m} \chi_{V_j}(x) \leq \sum_{j=1}^n \chi_{V_j}(x) \leq \sum_{j=1}^{q_{m+1}} \chi_{V_j}(x) < C_1(m+1)(\log(m+1))^3 \\ \frac{1}{2}m &< \sum_{j=1}^{q_m} \lambda(V_j) \leq \sum_{j=1}^n \lambda(V_j) \leq \sum_{j=1}^{q_{m+1}} \lambda(V_j) < 2C_1(m+1)(\log(m+1))^3 \end{aligned}$$

Taking logs and forming the relevant quotient, we see that the  $\log(m)$  and  $\log(m+1)$  terms dominate the  $\log(\text{constant})$  and  $\log(\log(-))$  terms. As  $\frac{\log(m)}{\log(m+1)} \rightarrow 1$ , the result follows.  $\square$

## 4. RESULTS ON FLAT SURFACES AND BILLIARDS

### 4.1. Flat surfaces.

**Definition 4.1.** *A translation surface  $\omega$  is a finite union of polygons  $P_1, \dots, P_r$  such that*

- (1) *the sides of the polygons are oriented so that the interior lies to the left*
- (2) *each side is identified to exactly one parallel side of the same length. They are glued together in an opposite orientation by parallel translation.*

This definition appears in [16, Definition 4]. In flat surfaces distance and a 2-dimensional volume  $\nu_Q$  make sense because they make sense in each polygon. Direction makes sense because of the gluings. Notice that an element of  $SL_2(\mathbb{R})$  applied to a flat surface produces another one. Let us assume that there is a fixed horizontal direction.  $F_\theta^t$  denotes flow with unit speed in direction  $2\pi\theta$  of the horizontal. Straight line flows with unit speed on  $Q$  are parametrized by  $\theta \in [0, 1)$ . Fixing a direction, we can draw a line segment perpendicular to this direction and obtain an interval exchange transformation  $T_\theta$  (see [20, Section 5.1] for a discussion in a survey paper). In this way given a fixed flat surface we can obtain a one parameter family of flows and interval exchange transformations.

A specific case of a flat surface is a square with opposite sides identified. This is a torus. If we let  $\bar{v}$  denote one of the sides of the square then  $T_\theta$  is rotation by  $\cot(\theta) \bmod 1$  (or  $2\pi \cot(\theta)$  on the unit circle).

**Definition 4.2.** *A line segment in  $\omega$  is called a saddle connection if it connects two vertices of the surface and has no vertex in its interior.*

Given fixed combinatorics, let  $\mathcal{T}_\epsilon$  denote the set unit area of flat surfaces where all the saddle connections have length at least  $\epsilon$ . We will often use  $\omega$  to denote a particular flat surface.

**4.2. Billiards.** This paper also addresses a particular family of dynamical systems, billiards in rational polygons. A polygon is called *rational* if all of its angles are rational multiples of  $\pi$ . Fixing a point in the polygon and an angle  $\theta$  we can consider the trajectory of that point as it travels in a straight line in the interior and obeys the rules of elastic collision on the sides. A countable number of trajectories which hit the corners of the polygon are undefined, but they have measure zero. The trajectory of the well defined points takes angles in a finite set because the polygon is rational. Following [8] one can reflect the polygon and keep the trajectory straight. There are at most a finite number of reflections needed to return to the original orientation. Identifying sides, one obtains a straight line flow in a flat surface. A special case of this dynamical system is the billiard flow in a regular polygon.

**4.3. Interval exchange transformations and their symbolic coding.**

**Definition 4.3.** *Given  $L = (l_1, l_2, \dots, l_d)$  where  $l_i \geq 0$ , we obtain  $d$  sub-intervals of the interval  $[0, \sum_{i=1}^d l_i)$ :*

$$I_1 = [0, l_1), I_2 = [l_1, l_1 + l_2), \dots, I_d = [l_1 + \dots + l_{d-1}, l_1 + \dots + l_{d-1} + l_d).$$

*Given a permutation  $\pi$  on the set  $\{1, 2, \dots, d\}$ , we obtain a  $d$ -Interval Exchange Transformation (IET)  $T: [0, \sum_{i=1}^d l_i) \rightarrow [0, \sum_{i=1}^d l_i)$  which exchanges the intervals  $I_i$  according to  $\pi$ . That is, if  $x \in I_j$  then*

$$T(x) = x - \sum_{k < j} l_k + \sum_{\pi(k') < \pi(j)} l_{k'}.$$

We use the symbolic coding of interval exchange transformations heavily. This section also shows the well known and useful fact that IETs are basically the same as (measure conjugate to) continuous maps on compact metric spaces. For concreteness assume that  $\sum_{i=1}^d l_i = 1$ .

Let  $\tau: [0, 1) \rightarrow \{1, 2, \dots, d\}^{\mathbb{Z}}$  by  $\tau(x) = \dots, a_{-1}, a_0, a_1, \dots$  where  $T^i(x) \in I_{a_i}$ .

This map is not continuous as a map from  $[0, 1)$  with the standard topology to  $\{1, 2, \dots, d\}^{\mathbb{Z}}$  with the product topology. Observe that the left shift acts continuously on  $\tau([0, 1)) \subset \{1, 2, \dots, d\}^{\mathbb{Z}}$ . However, if the discontinuities of  $T$  have infinite and disjoint orbits (the Keane condition) then  $\tau([0, 1))$  is not closed in  $\{1, 2, \dots, d\}^{\mathbb{Z}}$  with the product topology. This is because the points immediately to the left of a discontinuity give finite blocks that do not converge to an infinite block. Let  $\hat{X}$  be the closure of  $\tau([0, 1))$  in  $\{1, 2, \dots, d\}^{\mathbb{Z}}$  with the product topology.  $\hat{X}$  corresponds to adding a countable number of points, the left hand sides of points in orbits of a

discontinuity;  $\hat{X}$  is a compact metric space. Let  $f : \hat{X} \rightarrow [0, 1)$  by  $f|_{\tau([0,1])} = \tau^{-1}$  and extend  $f$  by continuity to the rest of  $\hat{X}$ . Notice that, unlike  $\tau$ , the map  $f$  is continuous. Moreover the map is injective away from the orbit of discontinuities, where it is 2 to 1. The left shift  $S$  acts continuously on  $\hat{X}$  and if  $T$  is not in the direction of a saddle connection then the action of  $S$  on  $\hat{X}$  is measure conjugate to the action of  $T$  on  $[0, 1)$ .

Fixing a point  $x$ , that is not in the orbit of a discontinuity, let

$$w_{p,q}(x) = c_p, c_{p+1}, \dots, c_{q-1}, c_q \text{ where } \tau(x) = \dots c_{-1}, c_0, c_1, \dots$$

If  $x$  is in the orbit of a discontinuity let  $w_{p,q}(x^+) = \lim_{y \rightarrow x^+} w_{p,q}(y)$ . Let  $w_{p,q}(x^-) = \lim_{y \rightarrow x^-} w_{p,q}(y)$ . Observe that if  $T$  satisfies the Keane condition (the orbits of its discontinuities are infinite and disjoint as sets),  $p > 0$  and  $w_{1,p}(x^+) \neq w_{1,p}(x^-)$  then  $w_{-N,-1}(x^+) = w_{-N,-1}(x^-)$  for all  $N > 0$ . Let  $\mathcal{B}_l(T) = \{a_1, \dots, a_l : \bigcap_{i=1}^l T^{-i}(I_{a_i}) \neq \emptyset\}$ . This is often called the set of allowed  $l$  blocks.

Assume that there exists half open intervals  $J_1, \dots, J_r$  and natural numbers  $m_1, \dots, m_r$  such that  $T^j$  is continuous (thus an isometry) on  $J_i$  for  $0 \leq j \leq m_i$ ,  $T^j(J_i) \cap T^{j'}(J_i) = \emptyset$  for  $0 \leq j < j' \leq m_i$  and  $\bigcup_{i=1}^r \bigcup_{j=0}^{m_i} T^j(J_i) = [0, 1)$ . We say

$\bigcup_{j=1}^{m_i} T^j(J_i)$  are *Rokhlin towers*.  $m_i + 1$  is called the *height* of the Rokhlin tower.

Each  $T^j(J_i)$  is called a *level*. Every word of  $\tau([0, 1)$  is a concatenation of  $\omega_{0,m_i}(z_i)$  where  $z_i \in J_i$ . By construction,  $y_i, z_i \in J_i$  implies that  $\omega_{0,m_i}(z_i) = \omega_{0,m_i}(y_i)$ . Also  $\omega_{0,m_i-j}(T^j(y_i)) = \omega_{j,m_i}(y_i)$ . In this way a set of Rokhlin towers at a fixed stage describes to a limited extent the dynamics of a system. As one takes Rokhlin towers with more and more levels one gains a better understanding of the dynamical system.

**4.4. Generalizing to billiards.** The main results Section 4 are Theorems 12, 13 and 15:

**Theorem 12.** *In every flat surface, for almost every  $(x, \theta)$  we have*

$$\lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N \chi_{B(0, \frac{1}{i})}(T_\theta^i(x))}{\sum_{i=1}^N \frac{2}{i}} = 1.$$

**Theorem 13.** *In every flat surface, for almost every  $\theta$  we have*

$$\lim_{N \rightarrow \infty} \frac{\int_1^N \chi_{B(0, \frac{1}{\sqrt{t}})}(F_\theta^t(x)) dt}{\int_1^N \pi \left(\frac{1}{\sqrt{t}}\right)^2 dt} = 1,$$

for every  $x$ .

Because these results hold in all flat surfaces they apply to the billiard in rational polygons and in particular any regular polygon [8]. Theorem 13 is proved by methods much closer to standard quantitative ergodicity arguments than the independence results used for the rest of the paper. Notice that it is a straightforward consequence of Theorem 12 that the *flow* in almost every direction hits balls of radius  $\frac{1}{t}$  infinitely often. However, the time it spends in these balls is integrable.

This motivates the choice of target in Theorem 13, which is easier to prove than Theorem 12.

The proof of Theorem 3 can be established very similarly for almost every IET using probabilistic results of Kerckhoff [9] to obtain independence of Rokhlin towers. More is required to establish Theorem 12 for all flat surfaces. This is motivated by the desire to prove our result for billiards in any fixed rational polygon. The main step in this argument is to make Boshernitzan's criterion for unique ergodicity effective.

Let  $A \subset \{1, \dots, d\}^{\mathbb{Z}}$  be shift invariant with linear block growth. Assume that it is minimal and has invariant measure  $\mu$ . Let  $\epsilon_n := \min_{w \in A} \mu(w_1, \dots, w_n^*)$  where  $w_1, \dots, w_n^*$  denotes the cylinder defined by  $w_1, \dots, w_n$ . Boshernitzan's criterion is the following:

**Theorem 14.** (Boshernitzan [3]) *If there exists a constant  $c$  such that for infinitely many  $n$ ,  $\epsilon_n \geq \frac{c}{n}$ , then the left shift is  $\mu$  uniquely ergodic.*

This was proved for IETs by Veech [19].

Let  $n_i$  be an increasing sequence of integers such that  $\epsilon_{n_i} > \frac{c}{n_i}$  and  $n_i > 10n_{i-1}$ .

**Theorem 15.** *Let  $b$  be a block of length  $n_i$ . There exist constants  $C_1, C_2$  depending only on  $c$  such that for any words  $w, w'$  we have  $\frac{1}{n_i+L} \left| \sum_{j=1}^{n_i+L} \chi_b(S^j w) - \chi_b(S^j w') \right| < C_1 e^{-C_2 L}$*

This is a quantitative version of Boshernitzan's criterion because it tells how quickly any orbit equidistributes.

One can apply a result of Athreya which shows (via Lemma 4.11) that Boshernitzan's criterion for unique ergodicity often applies. Recall that  $\mathcal{T}_\epsilon = \{\omega : \text{all the saddle connections of } \omega \text{ have length at least } \epsilon\}$ . There are two flows on the set of flat surfaces with fixed combinatorics –  $g_t$  and  $r_\theta$ . These correspond to the actions of subgroups of  $SL(2, \mathbb{R})$ :

$$g_t = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}, \quad r_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}.$$

**Theorem 16.** (Athreya [1]) *For any  $\omega$ , all small enough  $\epsilon > 0$  there exists  $C, c > 0$  such that  $\mu(\{\theta : g_t r_\theta \omega \notin \mathcal{T}_\epsilon \text{ for all } t \in [m, m+n]\}) \leq C e^{-cn}$ . Moreover  $c$  goes to 1 as  $\epsilon$  goes to 0.*

*Remark 4.4.* Athreya's Theorem is much closer to the formulation of Masur's criterion than to the formulation of Boshernitzan's criterion. Masur's criterion is: If  $g_t r_\theta \omega$  enters  $\mathcal{T}_\epsilon$  infinitely often for some  $\epsilon > 0$  then  $F_\theta^t$  is uniquely ergodic (with respect to Lebesgue measure) [15]. By passing to the symbolic coding, Theorem 15 implies a quantitative version of Masur's criterion. The proof via Boshernitzan's criterion is clearer to us.

The proof we gave above can be established very similarly for almost every IET using probabilistic results of Kerckhoff [9] to obtain independence of Rokhlin towers. We need to work a little harder to establish the result for all flat surfaces. This is motivated by the desire to prove our result for billiards in rational polygons. The main step in this argument is to make Boshernitzan's criterion for unique ergodicity effective. Then one can apply results of Athreya which show that Boshernitzan's criterion for unique ergodicity often applies.

**Corollary 4.5.** *For all small enough  $\epsilon$  there exists  $C, \delta$  such that  $\mu(\{\theta : \epsilon_s(T_\theta) \leq \frac{\epsilon}{s} \text{ for all } s \in [m, m+r]\}) \leq \frac{C}{r^{1-\delta}}$ . Moreover  $\delta$  goes to 0 as  $\epsilon$  goes to 0.*

This corollary follows from Lemma 4.11 which is deferred to a more convenient place.

We remark that with straightforward arguments Athreya's result implies that in any fixed flat surface, for almost every  $x$  and  $\theta$

$$\limsup_{N \rightarrow \infty} \frac{\sum_{i=1}^N \chi_{B(0, \frac{1}{i})}(T_\theta^i(x))}{\sum_{i=1}^N \frac{2}{i}} < \infty.$$

**4.5. Proof of Theorem 15.** For ease of notation we treat the case where  $n_i = 1$ ; the general case is the same.

Let  $B \subset \{1, \dots, d\}$ .

Let  $a_n(w|B) = \frac{|\{i \leq n : w_i \in B\}|}{n}$ .

Let  $M_n[B] = \max_w a_n(w|B)$  and  $m_n[B] = \min_w a_n(w|B)$ . The next lemma is similar to the main result in [3].

**Lemma 4.6.** *If  $\epsilon_n > \frac{c}{n}$  then*

$$\begin{aligned} \mu\left(\left\{w : a_n(w|B) \in \left[\frac{3}{4}m_n[B] + \frac{1}{4}M_n[B], \frac{1}{4}m_n[B] + \frac{3}{4}M_n[B]\right]\right\}\right) \\ \geq c\left(\frac{1}{4}m_n[B] + \frac{3}{4}M_n[B] - \left(\frac{3}{4}m_n[B] + \frac{1}{4}M_n[B]\right)\right). \end{aligned}$$

*Proof.* Let  $u_1, \dots, u_n$  be an allowed  $n$  block with exactly  $nm_n[B]$  occurrences of a letter in  $B$  and  $v_1, \dots, v_n$  be an allowed  $n$  block with exactly  $nM_n[B]$  occurrences of a letter in  $B$ . By minimality there is  $w = \dots, u_1, \dots, u_n, \dots, v_1, \dots, v_n, \dots$ . Consider the successive blocks of length  $n$  formed by moving one place along  $w$ . At each step the change in  $a_n(\cdot|B)$  can be at most  $\frac{1}{n}$ . So there needs to be at least

$$n\left(\frac{1}{4}m_n[B] + \frac{3}{4}M_n[B] - \left(\frac{3}{4}m_n[B] + \frac{1}{4}M_n[B]\right)\right)$$

different  $n$  blocks with  $a_n(\cdot|B)$  in our desired range (these blocks are different by the fact that  $a_n(\cdot|B)$  assigns them different values). The lemma follows by our assumption on  $\epsilon_n$ .  $\square$

The next proposition is similar to results used in [19].

**Proposition 4.7.** *If  $\epsilon_{2n} > \frac{c}{2n}$  then  $[0, 1)$  is the union of at most  $3d$ -Rokhlin towers of height between  $n$  and  $2n$ , and with every level of  $\mu$ -measure at least  $\frac{c}{2n}$ .*

*Proof.* Build disjoint towers with  $n$  levels such that that their bases are intervals bounded by discontinuities of  $T^n$ . Get a maximal collection of such towers. Every point is within  $n$  forward iterates of one of these towers. Whenever one can disjointly continue a pre-existing tower by forward iterates, do so. These towers will have height at most  $2n$ . If this is not possible (that is extending the tower hits a discontinuity of  $T$  before it is exhausted) then split the levels of the tower so that it can continue. The new subintervals will be bounded by discontinuities of  $T^{2n}$  (because they hit the discontinuity in at most  $n+n$  steps).  $\square$

Given  $n_i$  let  $\mathcal{R}_i$  be a collection of towers as in Proposition 4.7.

*Remark 4.8.* Notice that by construction each level has  $\mu$ -measure  $O(\frac{1}{n_i})$ .

**Lemma 4.9.** *Let  $S_i$  be the set of towers in  $\mathcal{R}_i$  which have at least  $\frac{1}{8}^i Cc$  occurrences of the symbol 1. Then  $\mu(S_{i+1}) \geq \min\{1, \mu(S_i) + C_2\}$  where  $C_2$  is a constant.*

*Proof.* Consider the words of length  $n_{i+1}$  as being concatenations of towers from  $\mathcal{R}_i$  (i.e. words of length  $n_i$ ). By an argument similar to Lemma 4.6 a set of words of at least fixed proportion,  $C_2$ , have at least a quarter of towers in  $S_i$  and at least a quarter not in  $S_i$ . By Proposition 4.7 each tower in  $\mathcal{R}_i$  has between  $n_i$  and  $2n_i$  letters. Therefore the proportion of occurrences of the symbol 1 in these blocks is at least  $\frac{1}{8}$  proportion of occurrences of the symbol 1 in blocks in  $S_i$ . By induction this gives  $\frac{1}{8}^{i+1} Cc$  occurrences of the symbol 1.  $\square$

**Corollary 4.10.** *There exist  $r$  and  $\delta > 0$  depending only on  $c$  such that any block of length  $n_{i+r}$  contains at least  $\delta \epsilon_{n_i} n_{i+r}$  disjoint occurrences of a block of length  $n_i$ .*

*Proof.* Choose  $r$  such that  $C_2 r > 1$ . Let  $\delta = (\frac{1}{8})^r Cc$ .  $\square$

*Proof of Theorem 15.* The proof is the same as for the Perron-Frobenius Theorem with Corollary 4.10 providing the assumption. We provide a sketch. Consider the matrix  $M_{k+1}$  that charts numbers of hits of towers of  $n_k$  to the bases of towers of  $n_{k-1}$ . That is  $M_{k+1}[i, j]$  is the number of disjoint copies of the  $j^{\text{th}}$  tower of  $n_k$  lie in the  $i^{\text{th}}$  tower of  $n_{k-1}$ . By Corollary 4.10 there exist  $r, \delta$  such that all of the entries of  $M_{j+r}$  are at least  $\delta$  of what they should be. This matrix is a contraction of at least a fixed proportion on projective space. The theorem follows.  $\square$

**4.6. Connection of flow to symbolic coding.** Consider Definition 4.1. Label the sides making sure a pair sides have the same label iff they are identified. Given a fixed direction of flow we code a point by the sides its orbit hits. The symbolic coding is related to the flow similarly to Section 4.3 (though the flow is closer to the mapping torus of the coding).

**4.7. Defining  $g_j$  in this context.** Given  $\omega$  and  $\theta$  let  $F_\theta^t$  be the flow in direction  $\theta$  and  $T_\theta$  be an IET in the usual way. If  $g_t r_\theta \omega$  returns to the compact part,  $\mathcal{T}_\epsilon$ , infinitely often then let  $l_1$  be a number such that  $g_{l_1} r_\theta \omega$  is in  $\mathcal{T}_\epsilon$ . Given  $l_i$  let  $l_{i+1} = \min_{s > 10+l_i} \{s : g_s r_\theta \omega \in \mathcal{T}_\epsilon\}$ .

**Lemma 4.11.** *There exists  $M \in \mathbb{R}$  depending on  $\omega$  (and  $\theta$ ) such that if  $g_{t_0} \omega$  is in the compact part then  $\epsilon_n(T_\theta) > \frac{\epsilon}{n}$  for some  $n \in [\frac{e^{t_0}}{M}, M e^{t_0}]$ .*

*Proof.* If  $\epsilon_s(T_\theta) < \frac{\epsilon}{s}$  then direction  $\theta$  on the surface is within  $\frac{\epsilon}{s^2}$  of a saddle connection which crosses the horizontal at least  $s$  times [2, Page 750]. There exists a constant  $M$  such that this saddle connection has length in  $[\frac{s}{M}, Ms]$ . Under the action of  $g_t$  ( $t$  roughly  $\log s$ ) this is shrunk to a short saddle connection. In particular, the vertical component is shrunk and the horizontal component is  $s \frac{\epsilon}{s^2} = \frac{\epsilon}{s}$  so it is not expanded to be bigger than  $C\epsilon$ .  $\square$

**Lemma 4.12.** *In every flat surface for almost every  $\theta$  we have  $\frac{n_{i+1}}{n_i} \leq i^{1+\frac{1}{6}}$  for all but finitely many  $i$ .*

This follows from the previous lemma and Theorem 16.

Let  $i_1 = n_{10}$  and inductively define  $u_{k+1} = \max\{n_i : \log(n_i) < \log(u_k) + \sqrt[3]{\log(u_k)}\}$ . Let  $v_{k+1}$  be such that  $u_{k+1} = n_{v_{k+1}}$ . Let

$$g_k(x) = \sum_{i=u_k}^{u_{k+1}} \chi_{B(0, \frac{1}{i})} T^i(x).$$

For the remainder of the paper  $n_j$  will be denoted  $q_j$  to make a clearer connection to the previous sections.

Properties of  $g_j$ .

**Lemma 4.13.**  $\sum_{i=q_k}^{q_{k+1}} \chi_{B(0, \frac{1}{i})} T^i(x) \leq 2c^{-1} \left(\frac{q_{k+1}}{q_k}\right)^{\frac{1}{2}}$ .

*Proof.* The orbit of  $x$  can land at most once in any interval of size  $\epsilon_{q_{k+1}}$  over the time interval  $[q_k, q_{k+1}]$ . Additionally only one out of every  $q_k$  hits can land in any interval of size  $\epsilon_{q_k}$ . The orbit hits  $B(0, \frac{1}{i})$  at most  $c^{-1} \sqrt{\frac{q_{k+1}}{q_k}}$  times for the first  $\sqrt{\frac{q_{k+1}}{q_k}} q_k$  iterates and at most  $c^{-1} \sqrt{\frac{q_{k+1}}{q_k}}$  times after that because  $\epsilon_{q_{k+1}} (c^{-1} \sqrt{\frac{q_{k+1}}{q_k}} q_k) \geq \frac{1}{q_k \sqrt{\frac{q_{k+1}}{q_k}}}$ .  $\square$

**Proposition 4.14.**  $\sum_{n=1}^{\infty} \lambda(\{x : g_n(x) > n^{\frac{2}{3}}\}) < \infty$ .

This analogous to the proof of Proposition 2.12. We provide a couple of the key lemmas below. Let  $C_k = \{x : \sum_{i=q_k}^{q_{k+1}} \chi_{[0, \frac{1}{i})}(R^i(x)) > \frac{1}{c} \sqrt{\frac{q_{k+1}}{q_k}}\}$  and  $D_k = \bigcup_{i=0}^{q_k-1} B(R^{-i}(0), \frac{1}{\sqrt[3]{(\frac{q_{k+1}}{q_k})^2 q_k}})$ .

**Lemma 4.15.**  $C_k \subset D_k$ .

**Lemma 4.16.**  $\lambda(\bigcap_{i=1}^r D_{k_i}) \leq (1 + c^{-1})^r \prod_{i=1}^r \lambda(D_{k_i})$ .

*Proof.* This follows by induction. Assume  $\bigcap_{i=1}^r D_{k_i}$  is the union of less than  $(1 + \frac{1}{c})^r \lambda(\bigcap_{i=1}^r D_{k_i} q_{k_r})$  intervals of size at most

$$\frac{2}{\frac{q_{k_r+1}}{q_{k_r}}^{\frac{2}{3}} q_{k_r}}.$$

We intersect this set with  $D_u$  for  $u > k_r$ . Observe that  $D_u$  is the union of intervals that are at least  $\epsilon_{n_u}(T)$  separated. Therefore each interval in  $\bigcap_{i=1}^r D_{k_i}$  intersects at most

$$1 + \frac{2}{\frac{q_{k_r+1}}{q_{k_r}}^{\frac{2}{3}} q_{k_r}} q_u c^{-1}$$

of them. The proof follows.  $\square$

Recall that in the proof of Theorem 8 we only use control of the square of the random variable. This motivates the next proposition.

**Proposition 4.17.**  $\sum_{j=i+2}^{\infty} |\int g_i(x) g_j(x) dx - \int g_i(x) dx \int g_j(x) dx| < C$  for all  $i$ .

Lemma 2.5 establishes that if we let  $g_i = \sum \tilde{g}_{i,j}$  where each  $\tilde{g}_{i,j}$  is  $g_i$  restricted to the towers  $\mathcal{R}_{v_{i,j}}$  defined as in Section 4.5 and 0 elsewhere then if  $k > i$   $\sum_{i=1}^{\infty} \sum_{k>i+1} \sum_{j,l} |\int \tilde{g}_{i,j}(x) \tilde{g}_{k,l}(x) dx - \int_{\mathcal{R}_{u_{i+1}}} \tilde{g}_{k,l} dx \int \tilde{g}_{i,j}|$  is summable. This is because towers are successive images of the same interval (apply Lemma 2.5) and by Remark 4.8 the levels of the towers have measure at most  $\frac{2}{u_k}$  (then apply Lemma 2.4).

The proposition is completed because the different towers  $R_{k,l}$  and  $R_{k,l'}$  converge to the same distribution of hits in the tower  $R_{i,j}$  because:

**Lemma 4.18.**  $\frac{\sum_{k=1}^{n_i} \chi_{\mathcal{R}_j}(T^k x)}{\sum_{k=1}^{n_i} \chi_{\mathcal{R}_j}(T^k y)}$  converges to 1 exponentially fast in  $i$ .

This follows from Theorem 15 by letting the block be the base (or any other for that matter) level of the tower  $\mathcal{R}_j$ . This result holds pointwise and so therefore on sets (like  $\mathcal{R}_i$ ) as well.

#### 4.8. Proof of Theorem 12.

**Proposition 4.19.** *If  $R_i : \Omega \rightarrow \mathbb{R}$  are a family of random variables such that*

- (1)  $\max R_n(\omega) \leq \left( \sum_{i=1}^{n-1} \int_{\Omega} R_i d\nu \right)^{\frac{2}{3}}$
- (2)  $\int_{\Omega} R_n d\nu > \frac{1}{2} \left( \sum_{i=1}^{n-1} \int_{\Omega} R_i d\nu \right)^{\frac{1}{3}}$  and
- (3)  $\sum_{i=1}^{\infty} \sum_{j>i+1}^{\infty} \int_{[0,1)} g_i(x) g_j(x) dx < +\infty$

then for almost every  $\omega$  we have

$$\lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N R_i(\omega)}{\sum_{i=1}^N \int R_i(\nu) d\nu} = 1.$$

*Proof.* Consider  $\left( \sum_{i=1}^N R_i - \int_{\Omega} R_i d\nu \right)^2$ . This is a non-negative random variable with

$$(6) \quad \int_{\Omega} \left( \sum_{i=1}^N R_i(\omega) - \int_{\Omega} R_i(z) d\nu(z) \right)^2 d\nu(\omega) = \int_{\Omega} \sum_{i=1}^N (R_i(\omega))^2 d\nu + \left( \int_{\Omega} R_i(\omega) d\nu \right)^2 + 2 \sum_{i<j \leq N} \int_{\Omega \times \Omega} R_i(\omega) R_j(z) d\nu(\omega) d\nu(z) \leq 2NN^{\frac{4}{3}} + CN = 2N^{\frac{7}{3}} + CN$$

where  $C$  does not depend on  $N$ . The remainder of the proof is similar to Theorem 8.  $\square$

**4.9. Outline of the proof of Theorem 13.** This is a straightforward consequence of Theorem 15. To see this let  $n_i$  be defined inductively so that  $n_{i+1} = \lfloor e^{\log n_i + \sqrt[3]{\log(n_i^2)}} \rfloor$ . Let  $B_i = B(x, \frac{1}{\sqrt{n_i}}) \subset \omega$  and  $A_i = B_i \setminus B_{i+1}$ . It is straightforward to use Theorem 15 and our diophantine assumption (one also uses a version of Lemma 4.13) to show that for any  $x$  we have  $\lim_{i \rightarrow \infty} \frac{\int_{n_i}^{n_{i+1}} \chi_{A_i}(F^t x) dt}{\int_{n_i}^{n_{i+1}} \chi_{B_i}(F^t(x)) dt} = 0$ .

One then uses Theorem 15 to show that under our diophantine assumption for *every*  $x, y$  we have  $\lim_{i \rightarrow \infty} \frac{\int_{n_i}^{n_{i+1}} \chi_{B_i}(F^t x) dt}{\int_{n_i}^{n_{i+1}} \chi_{B_i}(F^t y) dt} = 1$ . Now observe that

$$\int_{n_i}^{n_{i+1}} \chi_{B(0, \frac{1}{\sqrt{i}})}(F^t x) dt \geq \int_{n_i}^{n_{i+1}} \chi_{B_i}(F^t x) dt - \int_{n_i}^{n_{i+1}} \chi_{A_i}(F^t x) dt.$$

5. MORE TARGETS

**5.1. Proof of Theorem 4.** This section establishes a more general version of Theorem 4. That is, a version for more general flat surfaces. In this paper we proved results for special targets to obtain explicit conditions. If one wants to prove results for general targets this is possible by strengthening the conditions on the dynamical system.

**Definition 5.1.**  $S : X \rightarrow X$  a minimal shift dynamical system is called linear recurrent if there exists some  $c$  such that  $\epsilon_S(n) > \frac{c}{n}$  for all  $n$ .

Let  $\omega$  be a flat surface and  $\theta$  have the property that there exists  $\epsilon > 0$  with  $g_t r_\theta \omega \in \mathcal{T}_\epsilon$  for all  $t$ . The symbolic codings of interval exchange transformations that arise from  $F_\theta^t$  are linearly recurrent (if they do not contain extraneous discontinuities). This is a set of directions of Hausdorff dimension 1 in every translation surface [13]. In particular, rotations by numbers with uniformly bounded continued fraction expansion have linearly recurrent symbolic coding. Let  $\mu$  denote the unique ergodic measure on  $S$ . Let  $d_S(\bar{x}, \bar{y}) = \frac{1}{1 + \min\{i : x_i \neq y_i\}}$ . Let  $\tau$  be as in Section 4.3.

**Theorem 17.** If  $T$  is an IET whose symbolic coding is linearly recurrent,  $\{b_i\}_{i=1}^\infty$  is non-increasing and  $\sum b_i = \infty$  then

$$\lim_{N \rightarrow \infty} \frac{\sum_{n=1}^N \chi_{B(y, b_n)}(T^n x)}{\sum_{n=1}^N \lambda(B(y, b_n))} = 1$$

for almost every  $y$ .

Note that in the case of rotations [14] implies that badly approximable/linearly recurrent rotations are the only candidates for the theorem to hold. It not hard to show (using some induced maps of badly approximable rotations or particular  $\mathbb{Z}_2$  skew products of badly approximable rotations) that this equivalence fails for general interval exchanges.

Linear recurrent systems are nice and we have the following strong quantitative ergodicity result.

**Proposition 5.2.** Let  $T$  be an IET whose symbolic coding is linearly recurrent. Then for any  $\epsilon > 0$  there exists  $C_\epsilon := C$  such that for any interval  $J$ ,  $x$  and  $N > C|J|^{-1}$  we have  $\frac{1}{N|J|} \sum_{i=1}^N \chi_J(T^i x) \in [1 - \epsilon, 1 + \epsilon]$ .

**Lemma 5.3.** Let  $S : X \rightarrow X$  be linearly recurrent. For any  $\epsilon > 0$  there exists a constant  $C = C_\epsilon$  such that for any  $\bar{y}, \bar{x} \in X$ , cylinder set  $J = *, a_1, \dots, a_k, *$  and  $N > Ck$  we have  $|\frac{1}{N} \sum_{i=1}^N \chi_J(S^i \bar{x}) - \chi_J(S^i \bar{y})| < \epsilon$ .

This follows from Theorem 15.

The following Corollary is known and can be proved directly without appealing to Theorem 15 of which it is an immediate consequence.

**Corollary 5.4.** *Let  $T$  be an IET whose symbolic coding is linearly recurrent. There exists a constant  $C$  such that for any interval  $J$ ,  $x \in [0, 1)$  we have  $\sum_{i=1}^{|J|^{-1}} \chi_J(T^i x) < C$ .*

*Proof of Proposition 5.2.* Partition  $J$  by cylinder sets of  $X$  of between  $\frac{c\epsilon}{2|J|}$  and  $\frac{\epsilon}{2|J|}$ . Let  $C_{\frac{\epsilon}{2}}$  be as in Lemma 5.3. Set  $C = C_{\frac{\epsilon}{2}}(\epsilon^{-1} + c^{-1})$ . The cylinder sets entirely in  $J$  will be hit within  $\frac{\epsilon}{2}$  of the appropriate number of times. There will be up to two small intervals which are not an entire cylinder set. These can be hit at most  $(1 + \frac{\epsilon}{2})\frac{\epsilon}{2|J|}$  times each.  $\square$

Let  $n_k$  be defined inductively so that  $\sum_{i=n_k}^{n_{k+1}} b_i = O(1)$ . Let  $g_k(x) = \sum_{i=n_k}^{n_{k+1}} \chi_{B(0, b_i)}(T^i x)$ .

By arguments similar to the previous sections,

**Proposition 5.5.** *There exist  $C_1, C_2$  depending only on the constant from the linear recurrence such that if  $b_{n_{k+1}} > \frac{1}{2} b_{n_r}$  then*

$$\left| \int_0^1 g_k(x) g_r(x) - \int_0^1 g_k(x) dx \int_0^1 g_r(x) dx \right| < C_1 e^{-C_2 j}.$$

This follows similarly to Proposition 4.17 because  $n_{r+1} - n_r > 2^j(n_{k+1} - n_k)$ . Consider the  $i \in [n_r, n_{r+1}]$  such that  $\frac{b_{i+2^j/4}}{b_i} < 1 - 2^{-\frac{j}{4}}$ . Call this set  $B$ . This is at most  $\sum_{i \in B} \frac{1}{2^j} \sum_{k=1}^{\infty} 2^{\frac{j}{4}} (1 - 2^{-\frac{j}{4}})^k = 2^{-\frac{j}{2}}$ . By Proposition 4.17 if  $\frac{b_{i+2^j/4}}{b_i} > 1 - 2^{-\frac{j}{4}}$  we have

$$\left| \int_0^1 g_k(x) \sum_{t=i}^{i+2^{\frac{j}{4}}} \chi_{B(0, b_t)}(T^t x) dx - \int_0^1 g_k(x) dx \int_0^1 \sum_{t=i}^{i+2^{\frac{j}{4}}} \chi_{B(0, b_t)}(T^t x) dx \right| < C'_1 e^{-C'_2 j}.$$

The proposition follows by considering these stretches and the fact that the other times contribute at most  $2^{-\frac{j}{4}}$  which is exponentially small in  $j$ .

**Lemma 5.6.** *There exist constants  $C_1, C_2$  where  $C_2 < 1$  such that  $\lambda(\{x : g_k(x) > l\}) < C_1 C_2^l$*

It follows from Corollary 5.4 that on any stretch  $b_k, \dots, b_l$  where  $b_k$  is  $O(\epsilon)$  that  $\sum_{i=k}^l \chi_{B(0, b_i)}(T^i x) \leq C(l-k)\epsilon$ . Therefore if  $\{x : g_k(x) > l\}$  is non empty then  $\frac{\log(b_k)}{\log(b_l)}$  is not  $o(l)$ . One can then appeal to Theorem 15 to show that the points hit at different target sizes are roughly independent.

*Proof of Theorem 17.* It suffices to show that for any  $\epsilon > 0$  we have

$$\left| \lim_{N \rightarrow \infty} \frac{\sum_{n=1}^N \chi_{B(\bar{y}, b_n)}(T^n x)}{\sum_{n=1}^N \mu(B(y, b_n))} - 1 \right| < \epsilon$$

where  $\tilde{\lim}$  is lim sup or lim inf. Let  $C$  be given by Proposition 5.2 for  $\frac{\epsilon}{2}$ . We divide our times into two separate pieces. If  $k < l$ ,  $\frac{b_{n_k}}{b_{n_l}} < 1 - \frac{\epsilon}{4}$  and  $n_l - n_k > C(b_{n_l})^{-1}$  then let  $k, k+1, \dots, l \in V$ . Let  $h_N(x) = \sum_{n \in V} g_n(x)$ . Let  $U = \mathbb{N} \setminus V$  and  $\bar{h}_N(x) = \sum_{n \in U} g_n(x)$ . By Proposition 5.2 we have

$$\left| \lim_{N \rightarrow \infty} \frac{h_N(x)}{\int_0^1 h_N(x) dx} - 1 \right| < \epsilon$$

for all  $x$ . By Propositions 5.5 and 4.19

$$\left| \lim_{N \rightarrow \infty} \frac{\bar{h}_N(x)}{\int_0^1 \bar{h}_N(x) dx} - 1 \right| = 0$$

for almost every  $x$  (if  $U$  is infinite).  $\square$

**5.2. Proof of Theorem 5.** This section presents results that hold for almost every  $\alpha$ . Additionally, in every surface almost every direction is weakly bounded. This follows from [1, Theorem 1.1 (3)] because being weakly bounded is implied by the fact that for all  $\epsilon$  we have that  $g_t r_\theta \omega$  is in some fixed compact part for all but an  $\epsilon$  proportion of the time. (The compact part can be allowed to depend on  $\theta$ .)

Given a direction of flow  $\alpha$  let  $f(k) = \max\{i : q_i < 2^k\}$ . Let  $a_{k+1} = \frac{q_{k+1}}{q_k}$ . Recall  $q_k = n_k$ . Given  $C \in \mathbb{R}$  and  $\alpha$  let  $S_C(\alpha) = \{k : a_{f(k)} < C\}$ .

**Lemma 5.7.** *If  $\alpha$  is weakly bounded then  $S_C(\alpha)$  has density that goes to 1 as  $C$  goes to infinity.*

Recall that the density of a sequence of natural numbers  $A$  is  $\liminf_{N \rightarrow \infty} \frac{|A \cap [0, N]|}{N}$ .

*Proof.* By definition  $\log(q_k) = \log(q_1) + \sum_{i=2}^k \log(a_i)$ .  $|S_C(\alpha) \cap [0, r]| = \sum_{i: a_{f(i)} < C}^r 1$ .

If  $2^r$  is roughly  $q_k$  then this proportional to  $\sum_{i: a_i < C}^k \log(a_i)$ .  $\square$

**Lemma 5.8.** *If  $\{b_i\}_{i=1}^\infty$  is a Khinchin sequence then  $c_k = \sum_{i=2^k}^{2^{k+1}} b_i$  is a decreasing sequence with divergent sum.*

It satisfies  $2^{k+1} b_{2^{k+1}} \log 2 \leq c_k \leq 2^k b_{2^k} \log 2$ . Note that  $\sum_{k=1}^\infty 2^k b_{2^k}$  diverges because  $\{b_i\}_{i=1}^\infty$  is non-increasing with divergent sum.

Let  $g_k(x) = \sum_{i=2^k}^{2^{k+1}} \chi_{B(0, b_i)}(R^i x)$ .

**Proposition 5.9.** For any  $C > 0$  we have  $\lim_{N \rightarrow \infty} \frac{\sum_{k \in S_C}^N g_k(x)}{\sum_{k \in S_C}^N \int g_k(x) dx} = 1$  for almost every  $x$ .

*Proof.* Let  $S_C$  be enumerated in an increasing sequence  $\{k_i\}_{i=1}^\infty$ . There exists  $r$  depending only on  $C$  such that if  $n_i < k_j$  then  $n_{i+1} < k_{j+r}$ . The proof follows by establishing Propositions 4.17 and 4.19.  $\square$

**Lemma 5.10.** If  $\{b_i\}_{i=1}^\infty$  is a Khinchin sequence then there exists  $C$  such that  $\sum_{i=q_k}^{q_{k+1}} \chi_{B(0, b_i)}(R^i x) \leq C a_{k+1}$ .

This follows from the fact that  $b_i < \frac{C}{i}$  for all  $i$  and Lemma 2.3 (easily adapted to targets of size  $\frac{C}{i}$ ).

**Lemma 5.11.** For any  $K > 0$  there exists  $c', C_1$  depending only on  $K$  such that  $|\int_0^1 g_k(x) g_j(x) dx - \int_0^1 g_k(x) dx \int_0^1 g_j(x) dx| < C_1 e^{-c' |\{i \in S_K : k < i < j\}|}$ .

This is a consequence of Theorem 15.

**Corollary 5.12.** For any weakly bounded  $\alpha$  we have

$$\liminf_{N \rightarrow \infty} \frac{\sum_{n=1}^N \chi_{B(0, b_n)}(R_\alpha x)}{\sum_{n=1}^N 2b_n} = 1$$

for almost every  $x$ .

*Proof.* Proposition 4.19 and Lemma 5.11 implies that

$$\liminf_{N \rightarrow \infty} \frac{\sum_{k \in S_C}^N g_k(x)}{\sum_{k \in S_C}^N \int_0^1 g_k(x) dx} = 1.$$

For any decreasing sequence  $\{d_i\}_{i=1}^\infty$  we have that if  $A_1 \subset A_2 \subset \dots$  is a sequence of nested sequences whose density converges to 1 then

$$\lim_{C \rightarrow \infty} \lim_{N \rightarrow \infty} \frac{\sum_{i \in A_C}^N d_i}{\sum_{i=1}^N d_i} = 1.$$

The proof follows by Lemma 5.8 and Lemma 5.7.  $\square$

**Lemma 5.13.** For any weakly bounded  $\alpha$  we have

$$\lim_{C \rightarrow \infty} \lim_{N \rightarrow \infty} \frac{\sum_{k \notin S_C}^N g_k(x)}{\sum_{n=1}^{2^N} 2b_n} = 0$$

for almost every  $x$ .

This follows similarly to Lemma 2.14.

**Corollary 5.14.** *For any weakly bounded  $\alpha$  we have*

$$\limsup_{N \rightarrow \infty} \frac{\sum_{n=1}^N \chi_{B(0, b_n)}(R_\alpha x)}{\sum_{n=1}^N 2b_n} = 1$$

for almost every  $x$ .

**5.3. Proof of Theorem 6.** This section only applies to rotations, so  $q_n$  denotes the denominator of the  $n^{\text{th}}$  convergent to  $\alpha$ . Other than this the terminology is the same as in the previous subsection. The sufficiency is true for IETs with basically the same proof. The necessity is not true even for 3-IETs.

**Lemma 5.15.** *If  $\limsup_{n \rightarrow \infty} \frac{\log q_n}{n} < \infty$  then  $S_C(\alpha)$  has positive lower density for large enough  $C$ .*

*Proof.* Because  $q_{i+1} = a_i q_i + q_{i-1}$ , we have  $\log(q_1) + \sum_{i=2}^k \log(a_i + 1) > \log(q_k) > \log(q_1) + \sum_{i=2}^k \log(a_i)$ .  $|S_C(\alpha) \cap [0, r]| = \sum_{i: a_{f(i)} < C}^r 1$ . Notice  $q_{i+2} > 2q_i$ . It follows if  $2^r > q_k$  then  $|S_C(\alpha) \cap [0, r]| > k - 2|\{i \leq r : a_i > C\}|$ . Under our assumption for large enough  $C$  and  $k$  we have  $k - 2|\{i \leq r : a_i > C\}|$  is proportional to  $k$ . (Otherwise  $\limsup_{k \rightarrow \infty} \frac{1}{k}(\log(q_1) + \sum_{i=2}^k \log(a_i))$  would be  $\infty$ .) The lemma follows with the trivial observation that if  $2^r < q_k$  then  $|[0, r] \setminus S_C| < \log(q_k)$ .  $\square$

**Proposition 5.16.** *If  $\limsup_{n \rightarrow \infty} \frac{\log q_n}{n} < \infty$  then for large enough  $C$ ,*

$$\liminf_{k \rightarrow \infty} \frac{\sum_{i \in S_C}^k g_i(x)}{\sum_{i=1}^k \int_0^1 g_i(y) dy} > 0$$

for almost every  $x$ .

*Proof.* By Proposition 5.9 the numerator is roughly  $\sum_{i \in S_C}^k \int_0^1 g_i(y) dy$ . It follows from the previous lemma that

$$\liminf_{k \rightarrow \infty} \frac{\sum_{i \in S_C}^k \int_0^1 g_i(y) dy}{\sum_{i=1}^k \int_0^1 g_i(y) dy} > 0.$$

$\square$

This establishes the “if” part of Theorem 6.

*Proof of Theorem 6: “Only if” for rotations.* Observe that for rotations if  $q_j \leq i < q_{j+1}$  then  $d(R^i x, R^{i+q_j} x) \leq \frac{1}{a_{j+1} q_j}$  by Lemma 2.2 and we have  $\lambda(B(R_\alpha^{i+q_j}(x, \epsilon) \setminus B(R_\alpha^i(x, \epsilon))) \leq \min\{2\epsilon, \frac{1}{a_{j+1} q_j}\}$ . Assume  $\alpha$  has  $\limsup_{n \rightarrow \infty} \frac{\log(q_n)}{n} = \infty$ .

Then there exists  $k_1, \dots$  such that  $\frac{\log q_{k_i}}{k_i} > 10^i$ . Let  $b_i = \frac{1}{i \log q_{k_j}}$  for all  $q_{k_{j-1}} \leq i < q_{k_j}$ . Notice  $b_i$  is a Khinchin sequence. (In particular  $\sum_{i=N}^{q_{k_j}} b_i$  is roughly  $\frac{\log(q_{k_j}) - \log(N)}{\log(q_{k_j})}$ .)

By the fact that  $b_i$  is non-increasing we have

$$\lambda \left( \bigcup_{i=cq_n}^{(c+1)q_n} B(R^i x, b_i) \setminus \bigcup_{i=(c-1)q_n}^{cq_n} B(R^i x, b_i) \right) \leq q_n \min \left\{ b_{cq_n}, \frac{1}{a_{n+1}q_n} \right\}.$$

It is not hard to see that under our assumptions  $\sum_{k=1}^{\infty} 2^k \min \left\{ b_{2^k}, \frac{1}{a_{f(k)+1}q_{f(k)}} \right\}$  converges. In particular, consider  $\sum_{k=f^{-1}(q_{n_j})}^{f^{-1}(q_{n_{j+1}})} \min \left\{ \frac{\log(2)}{\log(q_{n_{j+1}})}, \frac{1}{a_{f(k)+1}q_{f(k)}} \right\}$  and split the sum into the few  $k \in S_{2^j}$  (where one uses the estimate  $\frac{\log(2)}{\log(q_{n_{j+1}})}$ ) and the most  $k$  that are not. This implies that  $\bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} B(R^{-i}x, b_i)$  has measure zero.  $\square$

#### REFERENCES

- [1] Athreya, J: *Quantitative recurrence and large deviations for Teichmüller geodesic flow*. Geom. Dedicata 119 (2006), 121-140.
- [2] Boshernitzan, M: *A condition for minimal interval exchange maps to be uniquely ergodic*. Duke Math. J. 52 (1985), no. 3, 723-752.
- [3] Boshernitzan: *A condition for unique ergodicity of minimal symbolic flows*. Erg. Th. & Dynam. Sys. 12 (1992), no. 3, 425-428.
- [4] Chaika, J: *Shrinking targets for IETs: Extending a theorem of Kurzweil*. Geom. Func. Anal. to appear.
- [5] Chaika, J; Cheung, Y; Masur, H: *Winning games for bounded geodesics on Teichmüller discs*. arxiv:1109.5976
- [6] Chernov, N; Kleinbock, D: *Dynamical Borel-Cantelli lemmas for Gibbs measures*. Israel J. Math. 122 (2001).
- [7] Eskin, A; Masur, H: *Asymptotic formulas on flat surfaces*. Erg. Th. & Dynam. Sys. 21 (2001), no. 2, 443-478.
- [8] Katok, A; Zemljakov, A N: *Topological transitivity of billiards in polygons*. Mat. Zametki 18 (1975), no. 2, 291-300.
- [9] Kerckhoff, S. P: *Simplicial systems for interval exchange maps and measured foliation*. Ergod. Th. & Dynam. Sys. 5 (1985), 257-271.
- [10] Kerckhoff, S., Masur, H., Smillie, J: *Ergodicity of billiard flows and quadratic differentials*. Annals of Math 1986 124 293-311.
- [11] Kesten, H: *On a conjecture of Erdős and Szűz related to uniform distribution mod 1*. Acta Arith. 12 (1966), 193-212
- [12] Khinchin: *Continued fractions*. Dover
- [13] Kleinbock, D, Weiss, B: *Bounded geodesics in moduli space*. Int. Math. Res. Not. 2004, no. 30, 1551-1560.
- [14] Kurzweil, J: *On the metric theory of inhomogeneous diophantine approximation*. Studia. Math. 15 (1955) 84-112.
- [15] Masur, H: *Hausdorff dimension of the set of nonergodic foliations of a quadratic differential*. Duke Math. J. 66 (1992), no. 3, 387-442.
- [16] Masur, H: *Ergodic theory of translation surfaces*. Handbook of dynamical systems. Vol. 1B, 527-547, Elsevier B. V., Amsterdam, 2006.
- [17] McMullen, C: *Winning sets, quasiconformal maps and Diophantine approximation*. Geom. Funct. Anal. 2010, 20 726-740,
- [18] Schmidt, W.: *A metrical theorem in diophantine approximation*. Canad. J. Math. 12 1960 619-631.
- [19] Veech, W: *Boshernitzan's criterion for unique ergodicity of an interval exchange transformation*. Erg. Th. & Dynam. Sys. 7 (1987), no. 1, 149-153.

- [20] Zorich, A: *Flat surfaces*. Frontiers in number theory, physics, and geometry. I, 437–583, Springer, Berlin, 2006.

*E-mail address:* `jonchaika@math.uchicago.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CHICAGO, 5734 S. UNIVERSITY AVENUE,  
CHICAGO, IL 60637, USA

*E-mail address:* `dc@math.uchicago.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CHICAGO, 5734 S. UNIVERSITY AVENUE,  
CHICAGO, IL 60637, USA