QUANTITATIVE SHRINKING TARGET PROPERTIES FOR ROTATIONS AND INTERVAL EXCHANGES

JON CHAIKA AND DAVID CONSTANTINE

1. Introduction

Let $\alpha \in [0,1)$ and $\lambda$ denote Lebesgue measure on $[0,1)$. $R_\alpha : [0,1) \to [0,1)$ by $R_\alpha(x) = x + \alpha \mod 1$ is one of the most natural and best understood dynamical systems. For example, Herman Weyl proved the following:

Theorem. Let $\alpha \notin \mathbb{Q}$. Then for any $\epsilon > 0$ we have

$$\lim_{N \to \infty} \frac{\sum_{i=1}^{N} \chi_{B(i\epsilon, x)}(R_i^\alpha x)}{N2\epsilon} = 1$$

This paper concerns the following question: What if the ball’s radius is allowed to shrink? The focus of this paper is on treating families of sequences $\{r_i\}$ simultaneously and obtaining explicit conditions on $\alpha$. The following is the main result of this paper:

Theorem 1. There exists an explicit full measure diophantine condition on $\alpha$ so that if $\alpha$ satisfies this condition then for any sequence $\{r_i\}$ so that $ir_i$ is non-increasing and $\sum_{i=1}^{\infty} r_i = \infty$ we have

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

for almost every $x$.

If $\alpha$ is badly approximable (a measure zero full Hausdorff dimension set) then we can relax the condition on the sequences further:

Theorem 2. Let $y \in [0,1)$. If $\alpha$ is badly approximable, $\{r_i\}_{i=1}^\infty$ is non-increasing and $\sum r_i = \infty$, then

$$\lim_{N \to \infty} \frac{\sum_{i=1}^{N} \chi_{B(y, r_i)}(R_i^\alpha x)}{\sum_{i=1}^{N} 2r_i} = 1$$

for almost every $x$.

Kurzweil showed that the conclusion of Theorem 2 can hold at most for badly approximable $\alpha$:
Theorem. (Kurzweil [20]) For any decreasing sequence of positive real numbers \( \{r_i\}_{i=1}^{\infty} \) with divergent sum there exists \( V \subset [0,1) \), a full measure set of \( \alpha \), such that for all \( \alpha \in V \) we have

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

for every \( x \).

On the other hand,

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

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

Let us make a few remarks to make the statements of Theorems 1 and 2 precise. We call a sequence \( \{r_i\} \) where \( ir_i \) is non-increasing and \( \sum r_i = \infty \) a Khinchin sequence. Let \( [a_1,...] \) be the continued fraction expansion of \( \alpha \). The number \( \alpha \) is badly approximable if \( \limsup_{n \to \infty} a_n < \infty \). The diophantine condition in Theorem 1 is as follows:

- \( a_n < \frac{2}{3} n \) for all but finitely many \( n \) and
- \( \lim \limsup_{C \to \infty} \limsup_{N \to \infty} \frac{1}{N} \left( \sum_{i=1}^{N} \log a_i - \sum_{a_i < C} \log a_i \right) = 0. \)

We will prove our results not just for rotations, but also for interval exchange transformations (Definition 1) satisfying similar diophantine assumptions. We mention D. Kim and S. Marmi [18], S. Galatolo [12], L. Marchese [22], M. Boshernitzan and J. Chaika [6], M. Marmi, S. Mousa and J-C Yoccoz [23] where a variety of Diophantine results for interval exchanges and rotations are proven. A key tool in extending our work to IET's is a quantitative version of Boshernitzan’s criterion for unique ergodicity (see Section 1 for terminology, historical discussion and proof):

**Theorem 3.** Let \( T \) be a minimal interval exchange transformation. Let \( e_T(n) \) denote the minimum length of an \( n \)-block of \( T^n \). Let \( c > 0 \). Assume \( n_j \in \mathbb{N} \) have the following two properties:

1. \( \frac{n_{j+1}}{n_j} > 2 \)
2. \( e_T(n) > \frac{1}{n_j} \).

Let \( J \) be an \( n_j \)-block of \( T \). There exist constants \( C_1, C_2 \) depending only on \( c \) such that for any points \( x, x' \) we have

\[
\left| \sum_{j=1}^{n_{j+1} \wedge L} \chi_{J}(S^j x) - \chi_{J}(S^j x') \right| < C_1 e^{-C_2 L}.
\]

Quantitative equidistribution results for interval exchanges have also been proven in [28], [11] and [2].

1.1. **A related problem.** To motivate our interest in collections of sequences and diophantine conditions we discuss a similar problem whose answer has been known since at least the 50’s.

If one is concerned about a specific sequence and not concerned about a diophantine condition things are much simpler. Observe (see for example [20], 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
Quantitative Shrinking Targets

(1) \( \lambda \times \lambda(\{(x, \theta) : |R^m_\theta(x) - y_1| \leq a \text{ and } |R^n_\theta(x) - y_2| \leq b\}) = \lambda \times \lambda(\{(x, \theta) : |R^m_\theta(x) - y_1| \leq a\}) \lambda \times \lambda(\{(x, \theta) : |R^n_\theta(x) - y_2| \leq b\}). \)

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

\[
\lim_{N \to \infty} \lambda \times \lambda(\{(x, \theta) : \frac{1}{N} \sum_{n=1}^{N} \chi_{B(0, \frac{1}{N})}(R^n_\theta x) - 1| > \epsilon\}) = 0.
\]

In fact, if one wishes to consider

\[
\lim_{N \to \infty} \lambda \times \lambda(\{(x, \theta) : \frac{1}{N} \sum_{n=1}^{N} \chi_{B(0,b_n)}(R^n_\theta x) - 1| > \epsilon\}) = 0,
\]

once again one readily gets convergence in measure.

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 Appendix B 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, Theorem 5. For almost every \( \alpha \) this frequency does not converge almost everywhere to the constant function (Theorem 6).

1.2. Outline of paper. We prove our results following the outline of the strong law of large numbers. We first prove Theorem 1. In Section 2.2 we prove Proposition 1. This says, in the presence of the diophantine assumption, a large part in the sum in the conclusion of Theorem 1 is made up of approximately independent quantities. The independence comes via Lemma 7 from effective equidistribution (Theorem 3) and approximate \( T \) invariance (Lemma 3). Section 2.3 shows via independence that Theorem 1 is true if we ignore some of the terms in the sum. Section 2.4 treats the terms ignored in Section 2.3. We then prove Theorem 2 in two parts. In Section 3.1 we treat \( r_i \) where \( \sup \{i r_i \} < \infty \). In Section 3.2 we treat the general case. Section 4 proves Theorem 4 which is used in the earlier sections. There are two appendices. Appendix A provides a treatment of the symbolic coding of an IET. This is well known material included for completeness. Appendix B has a complementary result for rotations, Theorem 6.

1.3. Acknowledgments. The first author named 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.
2. Proof of Theorem 1

2.1. Setup.

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

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

Given a permutation \( \pi \) on the set \( \{1, 2, ..., d\} \), 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

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

The points \( \{\sum_{i=1}^{j} l_i\} \) are the discontinuities of \( T \).

Recall the symbolic coding of an IET (Appendix A). Given an IET \( T \), let \( e_T: \mathbb{N} \rightarrow \mathbb{R} \) be defined as follows: \( e_T(n) \) is the minimum distance between \( 2 \) discontinuities of \( T^n \). If two discontinuities orbit into each other then \( e_T(n) \) is defined to be \( 0 \). Since \( T^{-1}(\{0, 1\}) \) are contained in the set of discontinuities we have that \( e_T(n) \) is at most the measure of the smallest \( (n-1) \)-block (see Appendix A). Notice that \( e_T \) is a non-increasing function.

Fix \( \xi > 0 \). Let \( n_i \) be defined inductively by \( n_{i+1} = \min\{2^k > n_i : e_T(2n_{i+1}) > \xi\} \). Let \( g_i(x) = \sum_{j=n_i}^{2n_i} \chi_{T^{-j}B(\frac{i}{n_{i-1}})}(x) \). Let \( a_i = \frac{n_i}{n_{i-1}} \). This section proves that if \( T \) is an IET so that for every \( \epsilon > 0 \) there exists \( \xi > 0, C \) so that \( a_i \leq \epsilon^4 \) and

\[
\limsup_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \log a_i - \frac{\sum_{i=1}^{N} \log a_i}{a_i < C} < \epsilon
\]

then for any Khinchin sequence \( r_i \) we have

\[
\lim_{N \to \infty} \frac{\sum_{i=1}^{N} \chi_{B(0, r_i)}(T^i x)}{\sum_{i=1}^{N} 2r_i} = 1
\]

for almost every \( x \).

Observe that if \( \{b_i\} \) is a Khinchin sequence then \( \int g_i = O(\int g_j) \) for any \( j < i \).

Lemma 1. \( g_i(x) \leq 1 + \frac{2n_i}{\xi} 2r_{n_i} \) for all \( i \).

Lemma 2. (Boshernitzan [1], Lemma 4.4) If the orbits of the discontinuities of \( T \) are infinite and distinct then for any interval \( J \) of size \( e_T(n) \) there exist integers \( p \leq 0 \leq q \) (which depend on \( J \)) such that

1. \( q - p \geq n \)
2. \( T^i \) acts continuously on \( J \) for \( p \leq i < q \)
3. \( T^i(J) \cap T^j(J) = \emptyset \) for \( p \leq i < j < q \).

Proof of Lemma 4. By Lemma 2’s conclusion 3, if \( T^{j} x, T^{j+r} x \in J \) then \( |J| \leq e_T(r) \). Partition \( B(\frac{1}{2}, r_{n_i}) \) into intervals of size \( e_T(n_i) \geq \frac{\xi}{2n_i} \).
2.2. Estimate on \( \int g_j(x)g_i(x) \). Let \( i > j \).

Proposition 1. There exists \( C \) so that for all \( j \)

\[
\sum_{i=j+1}^{\infty} \int g_i g_j - \|g_j\|_1 \|g_i\|_1 < C\|g_j\|_1.
\]

\( C \) depends only on \( \xi \).

Lemma 3. There exists \( C \) so that for every \( j \)

\[
\sum_{k=j+1}^{\infty} \max\{\|g_k - g_k \circ T^s\|_1 : 0 \leq s < n_{i+1}\} < C\|g_j\|_1.
\]

Proof.

(2) \[
\sum_{i=M}^{N} \chi_{T^{-i}B(\frac{1}{2}r_i)}(x) - \sum_{i=M}^{N} \chi_{T^{-i}B(\frac{1}{2}r_i)}(T^s x) =
\]

\[
\sum_{i=M}^{M+s-1} \chi_{T^{-i}B(\frac{1}{2}r_i)}(x) - \sum_{i=M}^{N} \chi_{T^{-i}B(\frac{1}{2}r_i)}(x) + \sum_{i=M+s}^{N} \chi_{T^{-i}B(\frac{1}{2}r_i)}(x) - \sum_{i=M+s}^{N} \chi_{T^{-i}B(\frac{1}{2}r_{i-s})}(x)
\]

Since we assume that \( r_i \) is non-increasing the \( L_1 \) norm of this is at most

\[
2sr_M + 2 \sum_{i=M+s}^{N} r_{i-s} - r_i \leq 2(s+1)r_M.
\]

Because our sequences are Khinchin sequences we obtain \( \sum Cn_{i+1}r_{n_k} \leq n_ir_{2n_i}2(1-\frac{1}{\sqrt{2}})^{-1} \). The lemma follows since \( \|g_i\|_1 > n_ir_{2n_i} \).

Given a finite set \( S \subset [0,1] \) let \( P_S \) be the finite partition of \([0,1]\) defined by connected components of \([0,1]\) \( \setminus S \).

Lemma 4. If \( S \) is \( \epsilon \) dense then there exists a function \( h \) which is constant on each element of \( P_S \) and whose \( L_1 \) difference from \( g_i \) is at most \( 2n_i \epsilon \).

It is straightforward to check that the characteristic function for any interval is \( 2\epsilon \) away from a function constant on elements of \( P_S \). The lemma follows because \( g_i \) is the sum of \( n_i \) characteristic functions of intervals.

Lemma 5. Let \( J \) be an \( m \)-block. Recall that \( d \) is the number of subintervals in our IET. Then there are at most \( d(2 - \log(\epsilon))n_i \) between \( \frac{1}{|J|} \) and \( m \).

Proof. We show that there exist \( d \) numbers \( k_1, ..., k_d \) between \( \frac{1}{|J|} \) and \( m \) so that if \( k_i < i < k_{i+1} \) then \( e_T(i) < \frac{1}{k_{i+1}} \). Consider \( J \), an \( m \)-block. \( J = [T^{-k\delta}, T^{-L\delta}] \)

where \( \delta, \delta' \) are either 0, 1 or discontinuities and where \( T^{-r}\delta'' \in J \) implies \( r \geq m \). Observe that to each interval of \( T_j \) there is a corresponding interval of \( T^{-1}_j \) with the same length and return times: \( r_1 \leq r_2 \leq ... \leq r_d \) (these are the \( k_i \) mentioned in the first sentence). There exists a return time, \( r_1 \) of size at most \( \frac{1}{|J|} \). If \( r_1 < m \) then the boundary point of this interval has to be in the orbit of \( \delta, \delta' \). So it is either \( T^{-s-k}\delta \) or \( T^{-s-L}(\delta') \) where \( s \leq r_i \). Pushing \( J \) forward by \( k, L < n \) respectively we obtain two \( s \)-blocks, one of which returns to \( J \) at \( r_1 \) and one that is still outside.
The part that is still outside will have no points return before \( r_2 \) and so its length is at most \( \frac{1}{r_2} \). Inductively we have \( k + 1 \) disjoint \( r_k + 1 \) blocks contained in \( J \), one of which does not have any points that return to \( J \) before \( r_{k+1} \) and so has length smaller than \( \frac{1}{r_{k+1}} \). \( \square \)

**Lemma 6.** If \( \|f_1 - f_2\| < \epsilon_1 \) and \( \|g_1 - g_2\| < \epsilon_2 \) then
\[
\int f_1(x)g_1(x) \leq \int f_2(x)g_2(x) + \|f\|_\infty \epsilon_2 + \|g\|_\infty \epsilon_1 + \epsilon_1 \epsilon_2.
\]

**Lemma 7.** Assume \( h \) is a function satisfying \( \|h - h \circ T^i\|_1 < \delta \) for \( i \leq n \) and
\[
|n|J| - \{0 < i < n : T^i(x) \in J\}| < n\delta'.
\]
Then
\[
\int h\chi_J - |J| \int h \leq (|J| + \delta') \int h + \delta.
\]

**Proof.** Let \( e_i(x) = h(x) - h \circ T^i(x) \).

\[
\int h(x)\chi_J(x) = \int \frac{1}{n} \sum_{i=1}^{n} (h \circ T^i(x) + e_i(x))\chi_J(x) \leq \int \frac{1}{n} \sum_{i=1}^{n} h \circ T^i(x)\chi_J(x) + \delta = \int h(x)\frac{1}{n} \sum_{i=1}^{n} \{1 \leq i \leq n : T^{-i}(x) \in J\} + \delta \leq (|J| + \delta')(\int h(x)) + \delta. \quad \square
\]

Notice in the first inequality above we only use that \( ||\chi_J||_\infty \leq 1 \) so by the same proof we obtain:

**Corollary 1.** Let \( J_1, ..., J_k \) be intervals such that
\[
|n|J_i| - \{0 < i < n : T^i(x) \in J_i\}| < n\delta'
\]
for all \( i \leq k \) and let \( h \) be a function so that \( ||h - h \circ T^i||_1 < \delta \) for all \( 0 \leq i \leq n \). Then
\[
| \int h(x) \sum_{i=1}^{k} \chi_{J_i}(x) - \int h \sum_{i=1}^{k} \chi_{J_i} | < (\sum_{i=1}^{k} \chi_{J_i} + \delta') \int h + \delta \sum_{i=1}^{k} \chi_{J_i}. \]

**Proof of Proposition**

\[
\sum_{i=j+1}^{\infty} | \int g_i g_j - \int g_j g_i | = \sum_{i > j \text{ such that } n_i < r_{2n_j}^{-1}} | \int g_i g_j | + \sum_{i > j \text{ such that } n_i \geq r_{2n_j}^{-1}} | \int g_i g_j - \int g_j g_i |.
\]

We first estimate the first term. Let \( t = 2n_j r_{2n_j} \), so by the Khinchin condition \( r_k \leq \frac{t}{4} \) for all \( k \geq 2n_j \). Notice that if \( N > t^{-1}2n_j \) then \( \frac{1}{N} < r_{2n_j} \). So the first term has at most \( j < i \leq j + \log_2(t^{-1}) + 1 \) summands. Moreover, for each such \( i \), \( g_i \) is at most \( t \log(2) + \frac{t^2}{2} \). We now show that there exists \( C \) so that \( \int |g_i g_j - \int g_i \int g_j | \) is at most \( Ct \log(2) + \frac{t^2}{2} \). By Lemma \( ||g_i||_\infty \leq 1 + \frac{t}{4} \). Since \( \int g_i g_j \leq ||g_j||_1 ||g_i||_\infty \) it
follows that there exists $\hat{C}$ so that the first summation is at most $\hat{C} t (\log(t^{-1}) + 1)$. Note: Because $r_k \leq \frac{n_k}{2}$, $\hat{C}$ can be chosen uniformly over all $j$.

Now we examine the second summand. We will use Lemma 3 to show that $g_i$ has little correlation with $f_{i,j}$, a function that is close to $g_j$. We will then apply Lemma 4 to show that $g_i$ and $g_j$ have little correlation.

If $J$ is an $\frac{n_{3j+i}}{2}$ block then by Theorem 3 we have

$$\left| \left| \{0 \leq i < n_{3j+i} : T^i x \in J \} - n_{3j+i} |J| \right| \right| < C_1 C_2^{\frac{1}{4j+2}}$$

for any $x$. Let $c_{j,i} = \max \{||g_i - g_i \circ T^r|| : 0 \leq r \leq n_{3j+i} \}$. If $f_i$ is the sum of $k$ disjoint characteristic functions of $\frac{n_{3j+i}}{2}$ blocks then

$$\left| \int g_i(x) f_i(x) - \int g_i(x) \int f_i(x) \right| \leq C_1 C_2^{\frac{1}{4j+2}} \int g_i(x) + c_{j,i}.$$ 

By Lemma 3 we have that $\sum_{i=j+1}^{\infty} c_{j,i} < C||g_j||_1$ and since by the Khinchin condition $||g_i||_1 \leq ||g_j||_1$ we have

$$\sum_{i=j}^{\infty} \left| \int g_i(x) f_i(x) - \int g_i(x) \int f_i(x) \right| \leq \hat{C}||g_j||_1. \quad (5)$$

Let $S_{i,j}$ be the set of discontinuities of $T^i \frac{n_{3j+i}}{2}$. By Lemma 4 we have that $S_{i,j}$ is $\frac{1}{n_{3j+i}}$ dense. Because $\{r_q \}$ is a Khinchin sequence, by our choice of $k_j$ and $n_k$ this is at most $2^{-\frac{i-j}{4}+r} r_{2n_j}$ for all $i > 4r$. By Lemma 5 for each $i > k_j + 4r$ there exists $h_i$ such that

$$||h_i - g_j||_1 \leq n_j r_{2n_j} 2 \cdot 2^{-\frac{i-j}{4}+r} < 2 \cdot 2^{-\frac{i-j}{4}+r}||g_j||_1. \quad (6)$$

So by Lemma 6 we have

$$\int g_j g_i - \int g_j \int g_i \leq \int h_i g_i - \int h_i \int g_i + 2||g_i||_\infty 2 \cdot 2^{-\frac{i-j}{4}+r}||g_j||_1.$$ 

By Lemma 7 $||g_i||_\infty$ is bounded.

Combining Equations 5 and 6 the proposition follows. \qed

2.3. Abstract setting.

**Proposition 2.** Let $g_i : [0, 1] \to \mathbb{R}^+$ so that for all $i$ there exists $C_1, C_2$: 

1. $||g_i||_\infty < C_2$
2. $\sum_{i=1}^{\infty} \int g_i = +\infty$
3. $\sum_{j=i+1}^{\infty} \int g_j(x) g_i(x) - \int g_i(x) \int g_j(x) < C_1 ||g_i(x)||_1$.

Then

$$\lim_{N \to \infty} \frac{\sum_{i=1}^{N} g_i(x)}{\sum_{i=1}^{N} \int g_i(x)} = 1$$

for a.e. $x$.

We prove this in two steps, Lemmas 10 and 11 below. Let $m_0 = 0$ and $m_k$ be defined inductively by $m_{k+1} = \min \{1 : \sum_{m_1}^{m_k+1} \int |g_i| \geq 1 \}$. Let $F_i = g_i - \int g_i$. Observe that $F_i$ satisfies the following:
(1) $\int F_i = 0$
(2) $\|F_i\|_{\infty} < C_2$
(3) $\sum_{j=1}^{\infty} \int F_j(x)F_i(x) < C'_2 \|F_i(x)\|_1$

To prove Proposition 2 we use the following classical results:

**Lemma 8.** (Chebyshev’s inequality) Let $R$ be a random variable with $\int R d\mu = 0$ and finite variance then $\mu(\{\omega : R(\omega) > c\}) \leq \frac{\int R^2 d\mu}{c^2}$.

**Lemma 9.** (Borel-Cantelli) If $A_1, \ldots$ are measurable sets and $\sum_{i=1}^{\infty} m(A_i) < \infty$ then $m(\{x : x \in A_i \text{ for infinitely many } i\}) = 0$.

**Lemma 10.**
$$\lim_{N \to \infty} \frac{\sum_{i=1}^{m_N^2} F_i(x)}{N^2} = 0$$

for a.e. $x$.

**Proof.** We first show that there exists $\tilde{C} > 0$ so that
$$\int (\sum_{i=1}^{m_M} F_i(x))^2 < \tilde{C} M.$$

The left hand side is $\int \sum_{i=1}^{m_M} F_i(x)^2 + 2 \sum_{j=1}^{m_M} F_i(x)F_j(x)$. By assumptions (2) and (3) the absolute value of the left hand side is at most $\sum_{i=1}^{m_M} (C_2^2 + 2C'_1) \int |F_i(x)|$.

Also by assumption (2), $\sum_{i=1}^{m_M} \int |F_i(x)|^2 \leq m_M \max\{C_2^2, 2\}$.

Now by Chebyshev’s inequality for each $N$ we have
$$\lambda(\{x : \sum_{i=1}^{m_N^2} F_i(x) > \delta N^2\}) < \frac{\tilde{C}}{\delta^2 N^2}$$

for each $\delta$. These sums converge and so for any $\delta > 0$ we have that for almost every $x$
$$\limsup_{N \to \infty} \left| \frac{\sum_{i=1}^{m_N^2} F_i(x)}{N^2} \right| \leq \delta.$$

□

Let $\delta > 0$ be given and
$$A_N = \{x : \max_{r \leq m_N^2} \left| \sum_{i=m_N^2}^{r} g_i(x) \right| > \delta N^2\}.$$

**Lemma 11.** $\sum_{i=1}^{\infty} \lambda(A_n) < \infty$.

**Proof.** Observe that since
$$\sum_{i=m_N^2+1}^{m_N^2} \|g_i\|_1 < 2N + 1 + C_2$$

and $F_i(x) = g_i(x) - \|g_i\|_1$, if
We next prove a maximal result. Let \( s \) the Khinchin condition. After an interval has decayed to at most \( N \) we have
\[
\max_{r \leq m^{(N+1)/2}} \left| \sum_{i=m_N^2}^{r} g_i(x) \right| > \delta N^2 \text{ then } \sum_{i=m_N^2}^{m^{(N+1)/2}} F_i(x) > \delta N^2 - 2N + 1 + C_2.
\]

Analogously to the first step of Lemma 10 we have that \( \int (\sum_{i=m+1}^{m_N^2} |F_i(x)|)^2 \leq C'' \). It follows that \( \int \max_{r \leq m^{(N+1)/2}} \left| \sum_{i=m_N^2}^{r} g_i(x) \right| < C'' N \) and by Chebyshev’s inequality \( \lambda(A_N) < C'' N^{-3} \).

Proof of Proposition 3. By Lemma 10 it suffices to show
\[
\limsup_{N \to \infty} \frac{\max_{r \leq m^{(N+1)/2}} \sum_{i=m_N^2+1}^{r} F_i(x)}{N^2} = 0
\]
for almost every \( x \). This follows by Lemma 11. \( \square \)

2.4. Controlling the \( \limsup \). We restrict our attention to \( \sum_{j \notin \cup_{n_i,2n_i}} \chi_{B(\frac{t}{2},r_j)}(T^j x) \).

Let \( \beta_i = \sum_{j=2n_i+1}^{n_i+1} \chi_{B(\frac{t}{2},r_j)}(T^j x) \). If \( \xi \) is small enough then for most \( i \), \( \beta_i \) is the zero function. We next prove a maximal result. Let \( s_{i+1} = \frac{n_{i+1}}{n_i} \).

Lemma 12. \( \sum_{j=2n_i}^{n_i+1} \chi_{B(0,r_j)}(T^j(x)) \leq 6\xi^{-1} \sqrt{2n_ir_{2n_i}} \sqrt{s_{i+1}} \).

Proof. We prove the lemma by the following trivial estimate:
\[
\max_{x} \sum_{j=2n_i}^{n_i+1} \chi_{B(0,r_j)}(T^j(x)) \leq \max_{x} \sum_{j=2n_i}^{u} \chi_{B(0,r_j)}(T^j(x)) + \max_{x} \sum_{j=u}^{n_i+1} \chi_{B(0,r_j)}(T^j(x)).
\]
By Lemma 10 there are at most \( 2\xi^{-1} 2n_i r_{2n_i} \) hits to an interval of size \( 2r_{2n_i} \) on an orbit of length \( 2n_i \). Let \( t = 2n_i r_{2n_i} \), so \( r_{2n_i} = \frac{t}{2n_i} \) and \( r_j \leq \frac{t}{j} \) for all \( j > 2n_i \) by the Khinchin condition. After \( \sqrt{t/\sqrt{s_{i+1}}} \) sets of \( 2r_{2n_i} \) by the Khinchin condition the interval has decayed to at most \( \frac{t}{\sqrt{s_{i+1}}} \). Since the first \( n_{i+1} \) elements of the orbit are \( \frac{\xi}{s_{i+1}} = \frac{\xi}{r_{2n_i}} \) separated there can only be \( \sqrt{\frac{t}{\sqrt{s_{i+1}}}} \) of them in this interval. Let \( u = \max \{ \sqrt{7}, 1 \} \sqrt{s_{i+1}} 2r_{2n_i} \), and the lemma follows. \( \square \)

Proposition 3. For any \( \epsilon > 0 \) and almost every \( x \) we have \( \sum_{i=1}^{N} \beta_i(x) < \epsilon \sum_{i=1}^{n_N} 2r_i \) for all large enough \( k \).

We prove this via the following probabilistic result:

Lemma 13. Let \( H_i : (0,1) \to \mathbb{R}^+ \) be a family of functions so that for every \( \epsilon > 0 \) there exists \( M \) so that
\[
(1) \sum_{i=1}^{N} ||H_i||_1 < \epsilon NC_N \text{ for all } M > N. \\
(2) \max_{i < N} \{ H_i(x) \} < C_0 C_N^\frac{2}{N} \\
(3) \sum_{N > j > i} \int H_i(x) H_j(x) < C_2 C_N^\frac{2}{N} ||H_i||_1.
\]

Then for almost every \( x \) we have \( \limsup_{N \to \infty} \frac{\sum_{i=1}^{N} H_i(x)}{C_N} = 0 \).
Let $R_i = H_i - \int H_i$.

**Proof.** As before we compute the variance. $\int (\sum_{i=1}^{N} R_i(x))^2 \leq \epsilon C_N^{\frac{3}{2}} + C_3 C_N^{\frac{3}{4}}$. This follows because $\|h\|_2^2 \leq \|h\|_1 \|h\|_\infty$ and Condition 3. By Chebyshev’s inequality

$$\lambda(\{x : \sum_{i} R_i(x) > \epsilon C_N\}) \leq C_6 C_N^{-\frac{3}{2}}.$$ 

Let $k_r = \min\{M : C_N > r \text{ for all } N > M\}$. By the Borel-Cantelli Lemma it follows that

$$\limsup_{N \to \infty} \frac{1}{N^4} \left| \sum_{i=1}^{k_r(N+1)^4} R_i(x) \right| \leq \epsilon.$$

Now consider $\sum_{i=k_n^4}^{k_{(N+1)^4}} R_i(x)$. By the definition of $R_i$ we have

$$\max_{L<k_{(N+1)^4}} \sum_{i=k_n^4}^{L} R_i(x) \leq \sum_{i=k_n^4}^{k_{(N+1)^4}} R_i(x) + \|H_i\|_1.$$

So the square of the $L_2$ norm is at most $C_7 N^3 (N^4)^{\frac{3}{2}}$. By Chebyshev’s inequality

$$\lambda(\{x : \max_{L<k_{(N+1)^4}} \left| \sum_{i=k_n^4}^{L} R_i(x) \right| > \epsilon N^4\}) \leq C_8 N^{-2}$$

where $C_8$ depends on $\epsilon$. By the Borel-Cantelli Lemma almost every $x$ can have $\sum_{j=k_n^4}^{k_{(N+1)^4}} R_i(x) < \epsilon n^4$ where $r < k_{(N+1)^4}$ only finitely many times. So considering $N$ as $m^4 + i$ where $m$ is the largest positive integer such that $m^4 \leq N$ we obtain

$$\limsup_{N \to \infty} \frac{1}{C_N} \left| \sum_{i=1}^{N} R_i(x) \right| \leq 2\epsilon$$

for almost every $x$. □

**Proof of Proposition 3.** It suffices to show that $\beta_i$ satisfy the assumption of Lemma 13 with $C_N = \sum_{i=1}^{nN} 2r_i$. Conditions (1) and (2) follows from our diophantine assumption on $\alpha$, the definition of $\beta_i$ and Lemma 12. Condition (3) follows analogously to Proposition 1. □

**Proof of Theorem 1.** By Proposition 2 we have

$$\lim_{N \to \infty} \frac{\sum_{i=1}^{N} g_i(x)}{\sum_{i=1}^{N} \int g_i} = 1.$$ 

Choose $\delta > 0$. There exists $\xi > 0$ so that with $g_i$ defined for this $\xi$ we have

$$\liminf_{N \to \infty} \frac{\sum_{i=1}^{N} g_i}{\sum_{i=1}^{N} \int g_i} > 1 - \delta.$$
From this it follows that
\[
\liminf_{N \to \infty} \sum_{i=1}^{N} \chi_{B(0,b_i)}(R_i^i x) = 1.
\]

Now
\[
\limsup_{N \to \infty} \sum_{i=1}^{\infty} \chi_{B(y,r_i)}(R_i^i x) \leq \limsup_{N \to \infty} \sum_{i=1}^{N} \sum_{j=1}^{n_N} 2r_i.
\]

By Proposition \(3\) the second summand contributes at most \(\varepsilon\) and the first summand follows by Propositions \(1\) and \(2\) so the theorem follows. □

3. Proof of Theorem 2

This section proves: If \(T\) is an IET so that there exists \(\sigma > 0\) with \(e_T(n) > \frac{\sigma}{n}\) for all \(n\) then for any decreasing sequence \(r_i\) with divergent sum we have:

\[
\lim_{N \to \infty} \sum_{i=1}^{N} \chi_{B(y,r_i)}(T^i x) = 1
\]

for almost every \(x\).

Let \(\sigma\) be a constant so that \(e_T(n) \geq \frac{\sigma}{n}\) for all \(n\). Let \(g_i = \sum\sum_{j=2}^{2r_i-1} \chi_{B(y,r_i)}(T^j x)\). We will prove Theorem 2 by splitting the \(r_i\) into two parts and showing that, for any \(\epsilon\), we have convergence within \(\epsilon\). In Section 3.1 we handle those times when \(i r_i\) is small – specifically when \(i r_i\) is less than a certain parameter \(M\). The proofs in this section work regardless of the value of \(M\). The case when \(i r_i < M\) is handled in a manner similar to the proof of Theorem 1. In Section 3.2 we deal with the case when \(i r_i \geq M\). For the proofs in this section we will have to choose a sufficiently large value of \(M\). This case is handled directly via Theorem 3, which will be proved in Section 4 below.

3.1. \(i r_i \to \infty\). For this subsection we assume that there exists \(M\) so that \(i r_i \leq M\) for all \(i\).

Lemma 14. There exists \(C_1, C_2\) that depends only on \(\sup e_T(n)\) and \(\sup \sigma r_n\) such that if \(i < j\) then
\[
\max_{k < 2^{i+j}} \|g_j - g_j \circ T^k\| < C_1 C_2^{i-j}.
\]

Proof. This is similar to Lemma 3. The estimate we have is \(C_2^{i+j} r_{2j}\). By our assumption on \(r_i\) this is at most \(C_2^{j-i} r_j\). □

Lemma 15. Let \(g_i : [0,1] \to \mathbb{R}^+\) be such that for all \(i\) there exists \(C_1, C_2\) such that:

1. \(\|g_i\|_{\infty} < C_1\)
2. \(\sum_{i=1}^{\infty} \|g_i\| = +\infty\)
(3) \( \sum_{j=i+1}^{\infty} \int g_j(x)g_i(x) - \int g_i(x) \int g_j(x) < C_2 \|g_i\|_1. \)

Then
\[
\lim_{N \to \infty} \frac{\sum_{i=1}^{N} g_i(x)}{\sum_{i=1}^{N} |g_i(x)|} = 0
\]
for a.e. \( x \).

This is a weaker result than Lemma 13.

By our assumption on \( r_i \) and Lemma 1 we have \( \|g_i\|_\infty < 2M\sigma^{-1} \) and so condition 1 is satisfied. The divergence of \( \sum r_i \) implies condition 2. Condition 3 follows analogously to Proposition 1. Indeed if \( 2^{-k} < r_{2l} \) then
\[
\int g_i g_j \leq C_1 C_3^{j-k} \int g_i \int g_j \leq C'_1 C_3^{j-k} M \int g_i.
\]

3.2. \( ir_i \) big. When \( ir_i \geq M \) we want to use the next lemma, which requires \( M \) sufficiently large:

**Lemma 16.** Let \( T \) be of constant type and \( C > 1 \).

\[
\lim_{s \to \infty} \limsup_j \frac{1}{C^{j+1} - C^j} \sum_{i \in C^j} \chi_{B(\frac{1}{2M C^j}, T^i x)}(T^i x) - 1 = 0.
\]

*Proof.* Because \( T \) is of constant type, fixing \( C, k \), by Lemma 5 for large enough \( M \), \( B(\frac{1}{2M C^j}, T^i x) \) can be approximated up to an \( \epsilon \) proportion by \( C^{j-k} \)-blocks (of \( T \)). We can apply Theorem 3 with \( L = k \). By choosing \( k \) large enough (given \( C, \sigma \)) we can treat the \( C^{j-k} \)-block. The rest of \( B(\frac{1}{2M C^j}, T^i x) \) is hit at most \( (2\sigma^{-1} + 1)\epsilon M M^{C^j+1} - C^j \) times by Lemma 5. The lemma follows.

The next lemma lets us split up the natural numbers into chunks where we appeal to Section 3.1 and chunks where we can apply Lemma 16. Throughout the remainder of this section \( C > 1 \) is probably very close to 1. In an abuse of notation \( r_{C^L} \) denotes \( r_{[C^L]} \)

**Lemma 17.** Define

\[ G_{C, \rho, M} = \{ j \in \mathbb{N} : r_{C^{j+1}} \geq \frac{M}{C^j+1} \text{ and } r_{C^j} \leq \rho r_{C^{j+1}} \} \]

and

\[ B_{C, \rho, M} = \{ j \in \mathbb{N} \setminus G_{C, \rho, M} : r_{C^{j+1}} \geq \frac{M}{C^j+1} \}. \]

For any \( \epsilon > 0, \rho < 1 < M \) there exists \( C > 1 \) so that for any \( \{r_i\} \subset \mathbb{R}^+ \) where \( r_i \geq r_{i+1} \) and \( \limsup_{i \to \infty} r_i = \infty \) we have

\[
\limsup_{N \to \infty} \frac{\sum_{j \in \mathbb{C}^j C^{j+1]} \sum_{i \in B_{C, \rho, M}} r_i}{\sum_{i=1}^{N} r_i} < \epsilon.
\]

Moreover, if \( p > q \) and \( pr_p > 2qr_q \) then
Lemma 18. Let $T$ be an IET of constant type, \( \{r_i\} \) nonincreasing. Then there exists $C'$ depending only on the constant type so that

\[
\sum_{j=C^i}^{C^{i+1}} \chi_B(\frac{1}{j} r_j) (T^i x) < C'(\sum_{j=C^i}^{C^{i+1}} 2r_j + \log(\frac{r_C}{r_{C^{i+1}}}) + 1).
\]

Proof. By Lemma 2 if $m > n$ then $T^m x \in B(a, r_m)$ and $T^n x \in B(a, r_m)$ then $e_T(m-n) \geq 2r_m$. Let $t_1 < ... < t_k \in [2^i, 2^{i+1}]$ be numbers so that $T^{t_j}(x) \in B(a, r_{t_j})$. If $2r_{t_{j+1}} > r_{t_j}$ then $\sum_{j=t_j}^{t_{j+1}} 2r_j \geq \frac{2}{3}$. So $1 + \sum_{i=t_1}^{t_k} 2r_i \leq (k - \log_2(\frac{r_C}{r_{C^{i+1}}})) \frac{2}{3}$. \hfill \( \square \)

With Equation (7) we obtain:

Corollary 2. For every $\epsilon$ there exists $C, \rho, M$ so that for all $x$

\[
\frac{\sum_{i=0}^{C^j+1} \chi_B(\frac{1}{j} r_j) (T^i x)}{\sum_{i=C^j}^{C^{j+1}} 2r_i} < \epsilon.
\]

Lemma 19. For any $\epsilon > 0$, $C > 1$ there exists $M_0 > 1 > \rho$ so that if $M > M_0$ for any $j \in G_{C, \rho, M}$

\[
\frac{\sum_{i=C^j}^{C^{j+1}} \chi_B(\frac{1}{j} r_j) (T^i x)}{\sum_{i=C^j}^{C^{j+1}} 2r_i} \in [1 - \epsilon, 1 + \epsilon].
\]

The lemma follows by using Lemmas 11 and 16 to show that

\[
\frac{\max \left\{|0 \leq i < C^{j+1} : T^i x \in B(\frac{1}{2}, r_j)\} \right\}}{\min \left\{|0 \leq i < C^{j+1} : T^i x \in B(\frac{1}{2}, r_{C^{j+1}})\} \right\}} < 1 + \epsilon.
\]

Proof. Let $\rho = 1 - \frac{1}{8} \sigma \epsilon$. Following Lemma 16 choose $L$ so that

\[
\limsup_j \frac{1}{C^{j+1} - C^j} \sum_{i=C^j}^{C^{j+1}} \chi_B(\frac{1}{j} r_j) (T^i x) - 1 < \epsilon^2.
\]

Let $M_0 = 2L$. By Lemma 16 we have
\[
\min \sum_{i=C}^{C+j+1} \chi_{B(\frac{1}{2}, r_{C+j+1})}(T^i x) \geq (1 - \epsilon^2)2r_{C_j} \rho(C - 1).
\]

Let us consider \(B(\frac{1}{2}, r_i) \setminus B(\frac{1}{2}, r_{C+j+1})\) when \(i \geq C_j\). This is contained in two fixed intervals of size at most \((1 - \rho)r_{C_j}\). By Lemma 1 we have

\[
\max_{x} \{|C_j \leq i < C_{j+1}: T^i x \in B(\frac{1}{2}, r_i) \setminus B(\frac{1}{2}, r_{C_{j+1}})\} \leq 2\sigma^{-1}2(1 - \rho)(C - 1)r_{C_j}(C_{j+1} - C_j).
\]

It follows that

\[
\frac{\max_{x} \{|C_j \leq i < C_{j+1}: T^i x \in B(\frac{1}{2}, r_i) \setminus B(\frac{1}{2}, r_{C_{j+1}})\}}{\min_{x} \{|C_j \leq i < C_{j+1}: T^i x \in B(\frac{1}{2}, r_{C_{j+1}})\}} < \frac{\epsilon}{2}. \quad \square
\]

Proof of Theorem 3. It suffices to show that for all \(\epsilon > 0\) there exists \(C > 1\) so that

\[
\liminf_{N \to \infty} \frac{\sum_{j=1}^{N} \sum_{i = C_j}^{C_{j+1}} \chi_{B(\frac{1}{2}, r_i)}(T^i x)}{\sum_{j=1}^{N} \sum_{i = C_j}^{C_{j+1}} 2r_i} > 1 - \epsilon
\]

and

\[
\limsup_{N \to \infty} \frac{\sum_{j=1}^{N} \sum_{i = C_j}^{C_{j+1}} \chi_{B(\frac{1}{2}, r_i)}(T^i x)}{\sum_{j=1}^{N} \sum_{i = C_j}^{C_{j+1}} 2r_i} < 1 + \epsilon.
\]

The proof follows by choosing \(\rho, C, M\) partitioning the \(j\) into 3 parts, \(G_{C, \rho, M}, B_{C, \rho, 2M}\) and the rest. By Section 3 for any \(M, C, \rho\) the indices not in \(G_{C, \rho, M}, B_{C, \rho, 2M}\) give a limit of 1 for almost every \(x\). By Lemma 19 we can choose \(M\) large enough and \(\rho\) close enough to 1 so that \(G_{C, \rho, M}\) has \(\liminf > 1 - \frac{\epsilon}{2}\) and \(\limsup < 1 + \frac{\epsilon}{2}\). By Corollary 2 we can simultaneously choose \(C, \rho, M\) (perhaps increasing \(M\)) so that we do not need to worry about \(B_{C, \rho, 2M}\) disturbing the \(\limsup\). By Equation 7 \(B_{C, \rho, M} \supset B_{C, \rho, 2M}\) does not disturb the \(\liminf\). \(\square\)

4. Quantitative Boshernitzan’s criterion

Theorem 4. (Boshernitzan [5]) Let \(S : X \to X\) be the left shift acting minimally on a symbolic dynamical system. Let \(\mu\) be an \(S\)-invariant measure. Let \(\epsilon_n\) be the \(\mu\) measure of the smallest cylinder set of length \(n\). 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 [27], in which case the invariant/ergodic measure is Lebesgue. Masur [21] established the analogous, in fact stronger, result for flows on flat surfaces.

Let \(n_i\) be an increasing sequence of integers such that \(\epsilon_{n_i} > \frac{c}{n_i}\) and \(n_i > 10n_{i-1}\). Let us recall Theorem 4

Theorem. Let \(S, \mu, \epsilon_n\) be as in Theorem 4. 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

\[
\frac{1}{n_{i+L}} \sum_{j=1}^{n_{i+L}} \chi_b(S^j w) - \chi_b(S^j w') < C_1 e^{-C_2 L}.
\]
This is a quantitative version of Boshernitzan’s criterion because it tells how quickly any orbit equidistributes. Quantitative ergodicity statements for IETs and flows have been profitably studied with deep results in \cite{11, 28} and \cite{2}.

For ease of notation we treat the case where $n_i = 1$; the general case is the same. Let $B \subset \{1, \ldots, d\}$.

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

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 a key step in \cite{5}.

**Lemma 20.** If $\epsilon_n > \frac{c}{n}$ then

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

**Proof.** Let $u_1, \ldots, u_n$ be an allowed $n$ block with exactly $nm_n[B]$ occurrences of a letter in $B$ and $v_1, \ldots, v_n$ be an allowed $n$ block with exactly $nM_n[B]$ occurrences of a letter in $B$. By minimality there is $w = \ldots, u_1, \ldots, u_n, \ldots, v_1, \ldots, v_n, \ldots$. Consider the successive blocks of length $n$ formed by moving one place along $\omega$. 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 $\epsilon_n$. \hfill $\Box$

The next proposition is similar to results used in \cite{27}.

**Proposition 4.** 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). \hfill $\Box$

Given $n_i$ let $\mathcal{R}_i$ be a collection of towers as in Proposition 4.

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

**Lemma 21.** Let $S_i$ be the set of towers in $\mathcal{R}_i$ which have at least $\frac{1}{4} Cc$ 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.** Assume $\mu(S_i) < 1$. 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 20 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 each tower in $R_i$ has between $n_i$ and $2n_i$ letters. Therefore the proportion of occurrences of the symbol 1 in these blocks is at least $\frac{1}{4}$ proportion of occurrences of the symbol 1 in blocks in $S_i$. By induction this gives $\frac{1}{8}i+1Cc$ occurrences of the symbol 1.

**Corollary 3.** There exist $r$ and $\delta > 0$ depending only on $c$ such that any block of length $n_{i+r}$ contains at least $\delta n_i n_{i+r}$ disjoint occurrences of a block of length $n_i$.

**Proof.** Choose $r$ such that $C_2r > 1$. Let $\delta = (\frac{1}{8})^rCc$. \hfill \blacksquare

**Proof of Theorem 3.** We prove this by induction assuming it is true for $L = kr$ and proving it for $L = (k + 1)r$. To each Rokhlin tower given by Proposition 4 for $n_{i+kr}$ give a symbol. Given an $n_{i+(k+1)r}$ block write it as a concatenation of these symbols (plus a prefix and suffix of length at most $n_{i+kr}$). Consider the symbols that correspond to the $n_{i+kr}$ towers that have the maximal and minimal frequency of a given letter. Denote these frequencies by $\Xi$ and $\xi$, respectively. By corollary 3 each $n_{i+(k+1)r}$ block has $\delta$ proportion of its letters coming from each of these towers. So the frequency of each symbol is between $\delta \Xi + (1 - \delta)\xi$ and $(1 - \delta)\Xi + \delta\xi$. The theorem follows by induction. \hfill \blacksquare

### Appendix A. Symbolic coding

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

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

$$w_{p,q}(x) = c_p, c_{p+1}, ..., c_{q-1}, c_q$$

where $\tau(x) = ... c_{-1}, c_0, c_1, ...$

This is a block of length $q - p$.

The map $\tau$ is not continuous as a map from $[0, 1)$ with the standard topology to $\{1, 2, ..., d\}^\mathbb{Z}$ with the product topology. Observe that the left shift acts continuously on $\tau([0, 1)) \subset \{1, 2, ..., d\}^\mathbb{Z}$. However, if the discontinuities of $T$ have infinite and disjoint orbits (the Keane condition) then $\tau([0, 1))$ is not closed in $\{1, 2, ..., d\}^\mathbb{Z}$ 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 $\{1, 2, ..., d\}^\mathbb{Z}$ with the product topology. $\hat{X}$ results from 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 $f : \hat{X} \to [0, 1)$ 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 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 $[0, 1)$. 


If $x$ is in the orbit of a discontinuity let $w_{p,q}(x^+)=\lim_{y\to x^+} w_{p,q}(y)$. Let $w_{p,q}(x^-)=\lim_{y\to x^-} w_{p,q}(y)$. 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 $B_l(T) = \{ a_1, ..., a_l : \bigcap_{i=1}^l T^{-i}(J_{a_i}) \neq \emptyset \}$. This is often called the set of allowed $l$ blocks. Observe that the preimages under $\tau$ of allowed $l$ blocks are bounded by discontinuities of $T^l$, 0, and 1. Note that $|B_{l+1}(T)| \leq |B_l(T)| + d - 1$ for all $l \geq 1$. That is all but $d-1$ $l$-blocks have a unique continuation to an $l+1$ block.

Assume that there exist half open intervals $J_1, ..., J_r$ and natural numbers $m_1, ..., m_r$ 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 $\omega_{0,m_i}(y_i) = \omega_{0,m_i}(y_i)$. Also $\omega_{0,m_i-1}(T^j(y_i)) = \omega_{j,m_i}(y_i)$. 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.

**Appendix B. Undetermined points as a shrinking target**

**B.1. Statement of the problem.** 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 $c_0, c_1, \ldots$ (finite or infinite) of 0’s and 1’s, let 

$C_{c_0, c_1, \ldots} = \{ x \in X : T^i x \in A_{c_i} \text{ for all } i \}$

and let $\Sigma$ be the set of finite codings $c_0, c_1, \ldots, c_n$ which actually occur, i.e. for which $C_{c_0, c_1, \ldots} \neq \emptyset$. Let $V_j = \{ x : x \in C_{c_0, c_1, \ldots, c_j} \text{ and such that } c_0, \ldots, c_j, 0, \text{ and } c_0, \ldots, c_j, 1 \in \Sigma \}$. This is the set of ‘undetermined’ points at step $j$, that is, 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 5.** For almost all $x \in [0, 1)$ and almost all $\alpha$,

$$\lim_{n \to \infty} \frac{\log \sum_{j=1}^n \chi_{V_j}(x)}{\log \sum_{j=1}^n \lambda(V_j)} = 1.$$ 

To understand why Theorem 5 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 $j$ translates. For $x \in X$, denote by $[[x]]_j$ the atom of $x$ in $\mathcal{P}_j$. The coding $c_0, \ldots, c_j$ 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 $U_j = R_\alpha^j(V_j)$.
B.2. Failure of a stronger convergence. Before turning to the proof of Theorem 5, we give an argument as to why we there is no stronger theorem along the lines of convergence for

\[
\frac{\sum_{j=1}^{n} \chi_{v}(x)}{\sum_{j=1}^{n} \lambda(v)}.
\]

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 \(\sum_{j=1}^{n} \lambda(v)\) predicts.

**Proposition 5.** 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 [17, page 60].

**Lemma 22.** For any \(n, b_1, ..., b_n \in \mathbb{N}\) we have

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

**Corollary 4.**

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

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

**Lemma 23.** \(\lambda(W_n) > \frac{1}{10}\).

**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 22 this is dominated by \(\sum_{i=1}^{n} \sum_{i=1}^{n} \frac{2}{i^2}\). 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 22 implies that \(\lambda(A_n) > 1 - \frac{2}{n}\). \(\square\)

**Lemma 24.** 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 \(\sum_{i=1}^{n} \int_{W_i} \frac{1}{C \log i}\). 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 \[\text{2}\] Given that $\alpha$ is in $W_{m-1}$ the Lemma \[\text{22}\] implies that the probability that $a_m(\alpha) \geq C \sum_{i=1}^{m-1} a_i(\alpha)$ is at least $\frac{1}{\log N \log N}$ 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 $\lambda(\bigcap_{i=1}^{\infty} B_i) > 0$. 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 \[\text{2}\]. These Lemmas can be obtained using the partial fraction expansion of $\alpha$. We will denote by $\lfloor y \rfloor$ 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 25. Let

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

Then

$$R_\alpha(U_j) = \lfloor s_j \alpha \rfloor + t_j \langle\langle r_j \alpha \rangle\rangle$$

or

$$R_\alpha(U_j) = \langle\langle r_j \alpha \rangle\rangle + (1 - \langle\langle s_j \alpha \rangle\rangle) - t_j \langle\langle r_j \alpha \rangle\rangle.$$

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 $j$.

Remark 2. 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: $0 < \langle\langle r_j \alpha \rangle\rangle < 1/2$. As the convergents alternate in approximating $\alpha$ from above and below, $1/2 < \langle\langle s_j \alpha \rangle\rangle < 1$. The only improvement possible in $\langle\langle r_j \alpha \rangle\rangle$ 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 $\langle\langle r_j \alpha \rangle\rangle$ as desired.

The lower bound on $R_\alpha(U_j)$ given by $\langle\langle s_j \alpha \rangle\rangle$ can be improved only by adding $\langle\langle r_j \alpha \rangle\rangle$ 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 $y = \lfloor s_j \alpha \rfloor + Tr_j \langle\langle r_j \alpha \rangle\rangle$ 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 - \langle\langle s_j \alpha \rangle\rangle - t_j \langle\langle r_j \alpha \rangle\rangle) as $\langle\langle r_j \alpha \rangle\rangle = \lfloor r_j \alpha \rfloor$. Since $\langle\langle s_j \alpha \rangle\rangle = 1 - \langle\langle s_j \alpha \rangle\rangle$, 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

$$\lambda(U_j) = \langle (s_j \alpha) \rangle - t_j (1 - [r_j \alpha]) + (1 - [r_j \alpha])$$

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

Proof. Let $J$ denote an interval of the form given in the statement of the Lemma. As atoms of the sequence of partitions $P_l$, the sets $U_l$ change only when the orbit of 0 hits $U_l$. By the description of Lemma 23 this does not happen over the interior of any of the intervals $J$. Suppose that $R^{-l}_\alpha U_l \cap R^{-k}_\alpha U_k \neq \emptyset$; with $l, k \in J$. Note that for such $l$ and $k$, $U_l = U_k$. Suppose $l > k$ and we obtain $U_l \cap R^{-k}_\alpha U_k \neq \emptyset$. However, the endpoints of $U_l = U_k$ are points in the orbit of zero which are reached by step $q_l$ at the latest. Therefore, for $R^{-k}_\alpha U_k$ 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 23.

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 $I_l = [q_i, q_{i+1})$).

Consider the integers $[q_{m-1}, q_m)$ for some $m$ satisfying Proposition 5: 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 (8) 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 23 $V_{j_0}$ and $V_{j_0 + q_{m-1}}$ intersect for all $c < a_m$. Points $x$ that lie in the intersection of the $V_{j_0 + q_{m-1}}$ 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 23 — convergence of the quotient (8) for almost every point fails. We sum up this failure by an arbitrary factor of $C$:

Theorem 6. For almost all $x \in [0, 1)$ and almost all $\alpha$,

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

and

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

Thus Theorem 5 is our best hope for this problem.
B.3. **Proof of Theorem 5** Towards Theorem 5 we claim the following set of inequalities:

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

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

**Lemma 27.** 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 \(G_n = \{ \alpha : a_i(\alpha) \leq n^2 \quad \forall i \leq n \}\) 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

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

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

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

**Lemma 28.** 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 26. As that Lemma shows, over each \(J_n\) the sets \(V_j\) are disjoint and hence can contribute at most one to \(\sum_{j=1}^{q_n} \chi_{V_j}(x)\).\]

For each \(i\), the interval of integers \(I_i = [q_i, q_{i+1})\) is divided (as in Lemma 26) into subintervals \(J_n\). Let us denote by \(J'_2\) the second of these intervals for each \(i\) — specifically, \(J'_2 = [q_i + q_{i-1}, 2q_i + q_{i-1})\). We remark that when \(a_{i+1} = 1\), \(J'_2\) is \([q_{i+1}, q_{i+1} + q_i)\) and is actually a subinterval of \(I_{i+1}\). Nonetheless, the collection \(\{J'_2\}\) 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 the \(J'_2\). Towards this end, let

\[h_i(x) = \sum_{j \in J'_2} \chi_{V_j}(x).\]

**Lemma 29.** For all \(i\),

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

**Proof.** As per Lemma 26 over \(J'_2\), the \(V_j\) are disjoint, so \(h_i(x) \in \{0, 1\}\). The length of the interval \(J'_2\) is \(q_i\), and for \(j \in J'_2\),

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

using the description of \(R_\alpha(U_j)\) provided by Lemma 26. By Theorem 13 in [17], \(\langle \langle q_{i-1}\alpha \rangle \rangle > \frac{1}{q_{i-1}+q_i}\). We may then bound the integral below by...
We prove with the following sequence of results that visits to the sets counted by the functions $h_i$ are (approximately) independent events.

**Lemma 30.** Let $[c, d) \subset [0, 1)$. Let $f_{[c,d]}(m) = \#\{[c,d) \cap \bigcup_{l \in J_m} R_{\alpha}^{-l}(0)\}$. 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 [??], by counting how many times the left endpoints of intervals $[i/q, i/q + 1/q_i]$ intersect $[c, d)$.

**Proposition 6.** For sufficiently large $m$ (relative to $i$)

$$\left( \frac{\lambda(V_i)|J_m| - 2}{\lambda(V_i)|J_m|} \right) \lambda(V_i) \lambda(\bigcup_{l \in J_m} V_l)$$

$$\leq \lambda(V_i) \cap \bigcup_{l \in J_m} V_l$$

$$\leq \left( \frac{\lambda(V_i)|J_m| + 2}{\lambda(V_i)|J_m|} \right) \lambda(V_i) \lambda(\bigcup_{l \in J_m} V_l).$$

**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

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

This holds for any $l \in J_m$ as all have the same measure. As $|J_m| \lambda(V_i) = \lambda(\bigcup_{l \in J_m} V_i)$ 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

$$\left( \frac{\lambda(V_i)|J_m| - 2}{\lambda(V_i)|J_m|} \right) \lambda(V_i) \lambda(\bigcup_{l \in J_m} V_l)$$

$$\leq \lambda(V_i) \cap \bigcup_{l \in J_m} V_l$$

$$\leq \left( \frac{\lambda(V_i)|J_m| + 2}{\lambda(V_i)|J_m|} \right) \lambda(V_i) \lambda(\bigcup_{l \in J_m} V_l).$$

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 $J_m$. 

$$\int_{[0,1)} h_i(x) d\lambda > \frac{q_i}{q_i + q_{i-1}} > \frac{q_i}{2q_i} = \frac{1}{2}.$$
Corollary 5. For any \( i \in J_n \), and \( J_n, J_m \) disjoint, \( n < m \)

\[
\left( \frac{\lambda(V_i)|J_m| - 2}{\lambda(V_i)|J_m|} \right) \lambda(\cup_{i \in J_n} V_i) \lambda(\cup_{l \in J_m} V_l) \\
\leq \lambda(\bigcup_{i \in J_n} V_i \cap \bigcup_{l \in J_m} V_l) \\
\leq \left( \frac{\lambda(V_i)|J_m| + 2}{\lambda(V_i)|J_m|} \right) \lambda(\cup_{i \in J_n} V_i) \lambda(\cup_{l \in J_m} V_l).
\]

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.) \( \square \)

Proposition 7. 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 \( x \in (\bigcup_{l \in J_2^i} V_l) \cap (\bigcup_{l \in J_2^j} V_l) \). Thus, we are calculating

\[
\lambda(\bigcup_{l \in J_2^i} V_l \cap \bigcup_{l \in J_2^j} V_l).
\]

By Corollary 5 we get

\[
\left( \frac{\lambda(V_i)|J_j^2| - 2}{\lambda(V_i)|J_j^2|} \right) \lambda(\cup_{l \in J_2^i} V_l) \lambda(\cup_{l \in J_2^j} V_l) \\
\leq \lambda(\bigcup_{l \in J_2^i} V_l \cap \bigcup_{l \in J_2^j} V_l) \\
\leq \left( \frac{\lambda(V_i)|J_j^2| + 2}{\lambda(V_i)|J_j^2|} \right) \lambda(\cup_{l \in J_2^i} V_l) \lambda(\cup_{l \in J_2^j} V_l)
\]

For the term \( \left( 1 \pm \frac{2}{\lambda(V_i)|J_j^2|} \right) \) we use any \( l \in J_j^2 \). By Kesten’s Theorem 22, 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 \( |J_j^2| \) we want a lower bound. From the description of these intervals given in the statement of Lemma 26 \( |J_j^2| \geq q_{j-1} \). Using these two bounds, the multiplicative error terms in the above become \( \left( 1 \pm \frac{2q_{i+1}}{q_{j-1}} \right) \).

Returning to our inequalities for \( \int h_i h_j \), as the \( V_l \) are disjoint over \( J_2^i \) or \( J_2^j \), we can translate back into integrals as so:
\[
\left(1 - \frac{2q_i - 1}{q_j - 1}\right) \int \sum_{l \in J^i} \chi_{V_l}(x) d\lambda \int \sum_{l \in J^j} \chi_{V_l}(x) d\lambda \\
\leq \int h_i h_j d\lambda \leq \\
\left(1 + \frac{2q_i - 1}{q_j - 1}\right) \int \sum_{l \in J^i} \chi_{V_l}(x) d\lambda \int \sum_{l \in J^j} \chi_{V_l}(x) d\lambda.
\]

These are the desired bounds on \(\int h_i h_j d\lambda\). \(\square\)

The independence result we want is the following.

**Proposition 8.** There exist constants \(C, b > 0\) such that

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

**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 \(q_{i+1} = a_i q_i + q_{i-1}\). As the \(a_i\) are positive integers, the \(q_i\) grow exponentially (by comparison with the Fibonacci sequence, e.g.). Thus, the terms \(\frac{2q_i - 1}{q_j - 1}\) decay exponentially in \(|i - j|\), finishing the proof. \(\square\)

We can apply this approximate independence to prove the remaining inequality in equation (9). Let \(s_i(x) = h_i(x) - \int h_i(x) d\lambda\), and note that \(s_i(x) \in (-1, 1)\). Let \(s_n(x) = \sum_{i=1}^n s_i(x)\).

**Proposition 9.** 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 29, this is bounded below by \(\frac{1}{4} n\); it is bounded above by \(n\) as \(h_i\) takes only 1 or 0 as a value. Applying Chebyshev’s inequality to \(s_n\) yields (for any \(\epsilon > 0\))

\[
\lambda(\{x : |s_n(x)| > \epsilon n\}) < \frac{\int s_n^2(x) d\lambda}{\epsilon^2 n^2} = \frac{\sum_{i=1}^n \int \hat{h}_i^2(x) d\lambda + 2 \sum_{i<j} \int \hat{h}_i(x) \hat{h}_j(x) d\lambda}{\epsilon^2 n^2} < \frac{D}{\epsilon^2 n}.
\]
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 8.

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

\[
\sum_{n \geq 1} \int \tilde{h}_i^2(x) d\lambda < n
\]

by Proposition 8. We have

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

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 \( x \in [0,1) \),

\[
\tilde{s}_{n^2}(x) \to 0 \quad \text{as } n \to \infty.
\]

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

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

as \( k \to \infty \).

We have now that for almost all \( x \),

\[
\sum_{i=1}^{n} h_i(x) - \int h_i(x) d\lambda/n \to 0.
\]

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.

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

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

The left-most inequality is Lemma 27. For the right-most:

**Lemma 31.** For all \( \alpha \),

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

**Proof.** This follows easily from Lemma 29 after noting that

\[
\sum_{j=q_n}^{q_{n+1}} \lambda(V_j) > \sum_{j=1}^{q_n} \lambda(V_j) = \int_{[0,1)} h_i(x) d\lambda.
\]

\( \square \)

It remains only to prove

**Lemma 32.** 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 $2^{a_i+1}$. 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, $\lambda(V_j) = \langle a \rangle + \langle a\rangle$. This is bounded above by $\frac{a_i}{q_i} + \frac{a_i}{q_i+1}$ by Thm 9. The total contribution of $J_1$ to the sum of $\lambda(V_j)$’s is thus bounded above by $\frac{a_i}{q_i} + \frac{a_i}{q_i+1} < 2$.

The intervals $J_2, \ldots, J_{a_i+1}$ each have length $q_i$ and over each of them $\lambda(V_j) < \langle a \rangle < \frac{1}{q_i}$. They thus each provide a contribution to the relevant sum of less than one and the result follows. \hfill $\Box$

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

Proof of Thm 5. Suppose $n \in [q_m, q_{m+1})$. Then we have the following:

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

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

Taking logs and forming the relevant quotient, we see that the log($m$) and log($m+1$) terms dominate the log(constant) and log(log($\cdot$)) terms. As $\frac{\log(m)}{\log(m+1)} \to 1$, the result follows. \hfill $\Box$

References

[1] Athreya, J: Quantitative recurrence and large deviations for Teichmüller geodesic flow. Geom. Dedicata 119 (2006), 121-140.
[2] Athreya, J. S.; Forni, G: Deviation of ergodic averages for rational polygonal billiards. Duke Math. J. 144 (2008), no. 2, 285319.
[3] Boshernitzan, M: A condition for minimal interval exchange maps to be uniquely ergodic. Duke Math. J. 52 (1985), no. 3, 723-752.
[4] Boshernitzan, M: Rank two interval exchange transformations. Ergod. Th. & Dynam. Sys. 8 (1988), no. 3, 379–394.
[5] Boshernitzan, M: A condition for unique ergodicity of minimal symbolic flows. Erg. Th. & Dynam. Sys. 12 (1992), no. 3, 425-428.
[6] Boshernitzan, M; Chaika, J: Quantitative proximality and connectedness. To appear Inventiones.
[7] Chaika, J: Shrinking targets for IETs: Extending a theorem of Kurzweil. Geom. Func. Anal. to appear.
[8] Chaika, J; Cheung, Y; Masur, H: Winning games for bounded geodesics on Teichmüller discs. arXiv:1109.5976
[9] Chernov, N; Kleinbock, D: Dynamical Borel-Cantelli lemmas for Gibbs measures. Israel J. Math. 122 (2001).
[10] Eskin, A; Masur, H: Asymptotic formulas on flat surfaces. Erg. Th. & Dynam. Sys. 21 (2001), no. 2, 443-478.
[11] Forni, G: Deviation of ergodic averages for area-preserving flows on surfaces of higher genus. Ann. of Math. (2) 155 (2002), 1–103.
[12] Galatolo, S: Hitting time and dimension in axiom A systems, generic interval exchanges and an application to Birkhoff sums. J. Stat. Phys. 123 (2006), no. 1, 111–124.
[13] Katok, A; Zemljakov, A N: Topological transitivity of billiards in polygons. Mat. Zametki 18 (1975), no. 2, 291–300.

[14] Kerckhoff, S. P: Simplicial systems for interval exchange maps and measured foliation. Ergod. Th. & Dynam. Sys. 5 (1985), 257-271.

[15] Kerckhoff, S., Masur, H., Smillie, J: Ergodicity of billiard flows and quadratic differentials. Annals of Math 1986 124 293-311.

[16] Kesten, H: On a conjecture of Erdős and Szüz related to uniform distribution mod 1. Acta Arith. 12 (1966), 193–212

[17] Khinchin: Continued fractions. Dover

[18] Kim, D.H.; Marmi, S: The recurrence time for interval exchange maps. Nonlinearity 21 (2008), no. 9, 2201–2210.

[19] Kleinbock, D, Weiss, B: Bounded geodesics in moduli space. Int. Math. Res. Not. 2004, no. 30, 1551-1560.

[20] Kurzweil, J: On the metric theory of inhomogeneous diophantine approximation. Studia. Math. 15 (1955) 84-112.

[21] Masur, H: Hausdorff dimension of the set of nonergodic foliations of a quadratic differential. Duke Math. J. 66 (1992), no. 3, 387-442.

[22] Marchese, L: The Khinchin theorem for interval exchange transformations. J. Mod. Dyn. 2011 no. 1, 123-183.

[23] Marmi, S; Moussa, P; Yoccoz, J.-C: The cohomological equation for Roth-type interval exchange maps. J. Amer. Math. Soc. 18 (2005), no. 4, 823872

[24] Masur, H: Ergodic theory of translation surfaces. Handbook of dynamical systems. Vol. 1B, 527–547, Elsevier B. V., Amsterdam, 2006.

[25] McMullen, C: Winning sets, quasiconformal maps and Diophantine approximation. Geom. Funct. Anal. 2010, 20 726-740.

[26] Schmidt, W.: A metrical theorem in diophantine approximation. Canad. J. Math. 12 1960 619-631.

[27] Veech, W: Boshernitzan’s criterion for unique ergodicity of an interval exchange transformation. Erg. Th. & Dynam. Sys. 7 (1987), no. 1, 149-153.

[28] Zorich, A: Deviation for interval exchange transformations. Erg. Th. & Dynam. Sys. 17 (1997), 14771499.

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

E-mail address: chaika@math.utah.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF UTAH, 155 S 1400 E, ROOM 233, SALT LAKER CITY, UT 84112, USA

E-mail address: dconstantine@wesleyan.edu

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, WESLEYAN UNIVERSITY, 265 CHURCH STREET, MIDDLETOWN, CT 06459, USA