NO TEMPORAL DISTRIBUTIONAL LIMIT THEOREM FOR A.E. IRRATIONAL TRANSLATION

DMITRY DOLGOPYAT AND OMRI SARIG

Abstract. Bromberg and Ulcigrai constructed piecewise smooth functions on the torus such that the set of α for which the sum $\sum_{k=0}^{n-1} f(x+k\alpha \bmod 1)$ satisfies a temporal distributional limit theorem along the orbit of a.e. x has Hausdorff dimension one. We show that the Lebesgue measure of this set is equal to zero.

1. Introduction and statement of main result

1.1. **Background.** Suppose $T : X \to X$ is a map, $f : X \to \mathbb{R}$ is a function, and $x_0 \in X$ is a *fixed* initial condition. We say that the T–*ergodic sums* $S_n = f(x_0) + f(Tx_0) + \cdots + f(T^{n-1}x_0)$ satisfy a *temporal distributional limit theorem (TDLT) on the orbit of* x_0, if there exists a non-constant real valued random variable Y, centering constants $A_N \in \mathbb{R}$ and scaling constants $B_N \to \infty$ s.t.

$$\frac{S_n - A_N}{B_N} \xrightarrow[N\to\infty]{} Y \text{ in distribution,} \tag{1.1}$$

when n is sampled uniformly from $\{1, \ldots, N\}$ and x_0 is fixed. Equivalently, for every Borel set $E \subset \mathbb{R}$ s.t. $\mathbb{P}(Y \in \partial E) = 0$,

$$\frac{1}{N}\text{Card}\{1 \leq n \leq N : \frac{S_n - A_N}{B_N} \in E\} \xrightarrow[N\to\infty]{} \mathbb{P}(Y \in E).$$

We allow and expect A_N, B_N, Y to depend on T, f, x_0.

Such limit theorems have been discovered for several zero entropy uniquely ergodic transformations, including systems where the more traditional spatial limit theorems, with x_0 is sampled from a measure on X, fail [Bec10, Bec11, ADDS15, DS17, PS, DSa]. Of particular interest are TDLT for

$$R_\alpha : [0,1] \to [0,1],\ R_\alpha(x) = x + \alpha \bmod 1,\ f_\beta(x) := 1_{[0,\beta)}(x) - \beta,$$

2010 *Mathematics Subject Classification.* 37D25 (primary), 37D35 (secondary).

This work was partially supported by the ISF grants 199/14, 1149/18 and by the BSF grant 2016105. The authors thank Adam Kanigowski for simplifying the statement and the proof of Lemma 3.2.

because the R_α–ergodic sums of f_β along the orbit of x represent the discrepancy of the sequence $x + n\alpha \bmod 1$ with respect to $[0, \beta)$ [Sch78, CK76, Bec10]. Another source of interest is the connection to the "deterministic random walk" [AK82, ADDS15].

The validity of the TDLT for R_α and f_β depends on the diophantine properties of α and β. Recall that $\alpha \in (0, 1)$ is *badly approximable* if for some $c > 0$, $|q\alpha - p| \geq c/|q|$ for all irreducible fractions p/q. Equivalently, the digits in the continued fraction expansion of α are bounded [Khi63]. Say that $\beta \in (0, 1)$ is *badly approximable with respect to* α if for some $C > 0$, $|q\alpha - \beta - p| > C/|q|$ for all $p, q \in \mathbb{Z}, q \neq 0$. If α is badly approximable then every $\beta \in \mathbb{Q} \cap (0, 1)$ is badly approximable with respect to α. The recent paper [BU] shows:

Theorem 1.1 (Bromberg & Ulcigrai)**.** *Suppose α is badly approximable and β is badly approximable with respect to α, e.g. $\beta \in \mathbb{Q} \cap (0, 1)$. Then the R_α-ergodic sums of f_β satisfy a temporal distributional limit theorem with Gaussian limit on the orbit of every initial condition.*

The set of badly approximable α has Hausdorff dimension one [Jar29], but Lebesgue measure zero [Khi24]. This leads to the following question: *Is there a β s.t. the R_α–ergodic sums of f_β satisfy a temporal distributional limit theorem for a.e. α and a.e. initial condition?*

In this paper we answer this question negatively.

1.2. **Main result.** To state our result in its most general form, we need the following terminology.

Let $\mathbb{T} := \mathbb{R}/\mathbb{Z}$. We say that $f : \mathbb{T} \to \mathbb{R}$ is *piecewise smooth* if there exists a finite set $\mathfrak{S} \subset \mathbb{T}$ s.t. f is continuously differentiable on $\mathbb{T} \setminus \mathfrak{S}$ and $\exists \psi : \mathbb{T} \to \mathbb{R}$ with bounded variation s.t. $f' = \psi$ on $\mathbb{T} \setminus \mathfrak{S}$. For example: $f_\beta(x) = 1_{[0,\beta)}(x) - \beta$ (take $\mathfrak{S} = \{0, \beta\}$, $\psi \equiv 0$). We show:

Theorem 1.2. *Let f be a piecewise smooth function of zero mean. Then there is a set of full measure $\mathcal{E} \subset \mathbb{T} \times \mathbb{T}$ s.t. if $(\alpha, x) \in \mathcal{E}$ then the R_α–ergodic sums of f do not satisfy a TDLT on the orbit of x.*

The condition $\int_\mathbb{T} f = 0$ is necessary: By Weyl's equidistribution theorem, for every $\alpha \notin \mathbb{Q}$, f Riemann integrable s.t. $\int_\mathbb{T} f = 1$, and $x_0 \in \mathbb{T}$, $S_n/N \xrightarrow[N\to\infty]{\text{dist}} \mathrm{U}[0, 1]$ as $n \sim \mathrm{U}(1, \ldots, N)$. See §1.4 for the notation.

This paper has a companion [DSb] which gives a different proof of Theorem 1.2, in the special case $f(x) = \{x\} - \frac{1}{2}$. Unlike the proof given below, [DSb] does not identify the set of α where the TDLT fails, but it does give more information on the different scaling limits for the distributions of S_n, $n \sim \mathrm{U}(1, \ldots, N_k)$ along different subsequences $N_k \to \infty$. [DSb] also shows that if we randomize both n and α by sampling (n, α)

uniformly from $\{1,\dots,N\}\times\mathbb{T}$, then $\big(S_n-\frac{1}{N}\sum_{k=1}^N S_k\big)/\sqrt{\ln N}$ converges in distribution to the Cauchy distribution.

The methods of [DSb] are specific for $f(x)=\{x\}-\frac{1}{2}$, and we do not know how to apply them to other functions such as $f_\beta(x)=1_{[0,\beta)}(x)-\beta$.

1.3. **The structure of the proof.** Suppose f is piecewise smooth and has mean zero.

We shall see below that if f is continuous, then for a.e. α, f is an R_α-coboundry, therefore S_n are bounded, hence (1.1) cannot hold with $B_N\to\infty$, Y non-constant. We remark that (1.1) does hold with $B_N\equiv 1$, $A_N=f(x_0)$, $Y=$ distribution of minus the transfer function, but this is not a TDLT since no actual scaling is involved.

The heart of the proof is to show that if f is discontinuous, then for a.e. α, the temporal distributions of the ergodic sums have different asymptotic scaling behavior on different subsequences. The proof of this has three independent parts:

(1) A reduction to the case $f(x)=\sum\limits_{m=1}^{d} b_m h(x+\beta_m)$, $h(x):=\{x\}-\frac{1}{2}$.
(2) A proof that if $\mathcal{N}\subset\mathbb{N}$ has positive lower density, then there exists $M\geq 1$ s.t. the following set has full Lebesgue measure in $(0,1)$:

$$\mathcal{A}(\mathcal{N},M):=\left\{\alpha\in(0,1):\ \begin{array}{l}\exists n_k\uparrow\infty,\ r_k\leq M \text{ s.t. for all } k:\\ r_kq_{n_k}\in\mathcal{N},\ a_{n_k+1}/(a_1+\cdots+a_{n_k})\to\infty\end{array}\right\}.$$

Here a_n and q_n are the partial quotients and principal denominators of α, see §3.1.

(3) Construction of $\mathcal{N}=\mathcal{N}(b_1,\dots,b_d;\beta_1,\dots,\beta_d)\subseteq\mathbb{N}$ with positive density, s.t. for every $\alpha\in\mathcal{A}(\mathcal{N},M)$ and a.e. x, one can analyze the temporal distributions of the Birkhoff sums of $\sum\limits_{m=1}^{d} b_m h(x+\beta_m)$.

1.4. **Notation.** $n\sim \mathrm{U}(1,\dots,N)$ means that n is a random variable taking values in $\{1,\dots,N\}$, each with probability $\frac{1}{N}$. $\mathrm{U}[a,b]$ is the uniform distribution on $[a,b]$. Lebesgue's measure is denoted by mes. $\mathbb{N}=\{1,2,3,\dots\}$ and $\mathbb{N}_0=\mathbb{N}\cup\{0\}$. If $x\in\mathbb{R}$, then $\|x\|:=\mathrm{dist}(x,\mathbb{Z})$ and $\{x\}$ is the unique number in $[0,1)$ s.t. $x\in\{x\}+\mathbb{Z}$. $\mathrm{Card}(\cdot)$ is the cardinality. If $\varepsilon>0$, then $a=b\pm\varepsilon$ means that $|a-b|\leq\varepsilon$.

2. Reduction to the case $f(x)=\sum\limits_{m=1}^{d} b_m h(x+\beta_m)$

Let $h(x)=\{x\}-\frac{1}{2}$, and let $\mathcal{G}$ denote the collection of all non-identically zero functions of the form $f(x)=\sum\limits_{m=1}^{d} b_m h(x+\beta_m)$, where $d\in\mathbb{N}, b_i,\beta_i\in\mathbb{R}$. We explain how to reduce the proof of Theorem 1.2 from the case of a general piecewise smooth $f(x)$ to the case $f\in\mathcal{G}$.

The following proposition was proved in [DSb]. Let $C(\mathbb{T})$ denote the space of continuous real-valued functions on $\mathbb{T}$ with the sup norm.

Proposition 2.1. *If $f(t)$ is differentiable on $\mathbb{T}\setminus\{\beta_1,\dots,\beta_d\}$ and f' extends to a function with bounded variation on $\mathbb{T}$, then there are $d\in\mathbb{N}_0$, $b_1,\dots,b_d\in\mathbb{R}$ s.t. for a.e. $\alpha\in\mathbb{T}$ there is $\varphi_\alpha\in C(\mathbb{T})$ s.t.*

$$f(x)=\sum_{i=1}^{d} b_i h(x+\beta_i)+\int_{\mathbb{T}} f(t)dt+\varphi_\alpha(x)-\varphi_\alpha(x+\alpha)\ \ (x\neq\beta_1,\dots,\beta_d).$$

The following proposition was proved in [DS17]. Let $(\Omega,\mathcal{B},\mu)$ be a probability space, and let $T:\Omega\to\Omega$ be a probability preserving map.

Proposition 2.2. *Suppose $f=g+\varphi-\varphi\circ T$ μ-a.e. with $f,g,\varphi:\Omega\to\mathbb{R}$ measurable. If the ergodic sums of g satisfy a TDLT along the orbit of a.e. x, then so do the ergodic sums of f.*

These results show that if Theorem 1.2 holds for every $f\in\mathcal{G}$, then Theorem 1.2 holds for any *discontinuous* piecewise smooth function with zero mean. As for *continuous* piecewise smooth functions with zero mean, these are R_α-cohomologous to $g\equiv 0$ for a.e. α because the b_i in Proposition 2.1 must all vanish. Since the zero function does not satisfy the TDLT, continuous piecewise smooth functions do not satisfy a TDLT.

3. The set $\mathcal{A}$ has full measure

3.1. **Statement and plan of proof.** Let α be an irrational number, with continued fraction expansion $[a_0;a_1,a_2,a_3,\dots]:=a_0+\dfrac{1}{a_1+\cdots}$, $a_0\in\mathbb{Z}$, $a_i\in\mathbb{N}$ $(i\geq 1)$. We call a_n the *quotients* of α. Let p_n/q_n denote the *principal convergents* of α, determined recursively by

$$q_{n+1}=a_{n+1}q_n+q_{n-1},\quad p_{n+1}=a_{n+1}p_n+p_{n-1}$$

and $p_0=a_0, q_0=1$; $p_1=1+a_1a_0$, $q_1=a_1$. We call q_n the *principal denominators* and a_i the *partial quotients* of α. Sometimes – but not always! – we will write $q_k=q_k(\alpha)$, $p_k=p_k(\alpha)$, $a_k=a_k(\alpha)$.

Given $\mathcal{N} \subset \mathbb{N}$ and $M \geq 1$, let $\mathcal{A} = \mathcal{A}(\mathcal{N}, M) \subset (0,1)$ denote the set of irrational $\alpha \in (0,1)$ s.t. for some subsequence $n_k \uparrow \infty$,

$$\exists r_k \leq M \text{ s.t. } r_k q_{n_k} \in \mathcal{N}\ ,\ \frac{a_{n_k+1}}{(a_0 + \cdots + a_{n_k})} \xrightarrow[k\to\infty]{} \infty. \tag{3.1}$$

The *lower density* of $\mathcal{N}$ is $d(\mathcal{N}) := \liminf \frac{1}{N}\mathrm{Card}(\mathcal{N} \cap [1, N])$. The purpose of this section is to prove:

Theorem 3.1. *If a set $\mathcal{N}$ has positive lower density then there exists M such that $\mathcal{A}(\mathcal{N}, M)$ has full Lebesgue measure in $(0,1)$.*

The proof consists of the following three lemmas:

Lemma 3.2. *For almost all α there is $n_0 = n_0(\alpha)$ s.t. if $k \geq n_0$ and $a_{k+1} > \frac{1}{4}k(\ln k)(\ln\ln k)$, then $a_{k+1}/(a_1 + \cdots + a_k) \geq \frac{1}{8}\ln\ln k$.*

Lemma 3.3. *Suppose $\alpha \in (0,1) \setminus \mathbb{Q}$ and $(p,q) \in \mathbb{N}_0 \times \mathbb{N}$ satisfy $\gcd(p,q) = 1$ and $|q\alpha - p| \leq \dfrac{1}{qL}$ where $L \geq 4$. Then there exists k s.t. $q = q_k(\alpha)$ and $a_{k+1}(\alpha) \geq \frac{1}{2}L$.*

Lemma 3.4. *Suppose $\psi : \mathbb{R}_+ \to \mathbb{R}$ is a non-decreasing function s.t.*

$$\sum_n \frac{1}{n\psi(n)} = \infty. \tag{3.2}$$

Suppose $\mathcal{N} \subset \mathbb{N}$ has positive lower density. For all M sufficiently large, for a.e. $\alpha \in (0,1)$ there are infinitely many pairs $(m,n) \in \mathbb{N}_0 \times \mathbb{N}$ s.t. $n \in \mathcal{N}, \gcd(m,n) \leq M$, and $|n\alpha - m| \leq \dfrac{1}{n\psi(n)}$.

Remark 1. By the monotonicity of ψ, if $e^{k-1} < n < e^k$ then $\psi\left(e^{k-1}\right) \leq \psi(n) \leq \psi\left(e^k\right)$. Hence (3.2) holds iff $\sum \frac{1}{\psi(e^k)} = \infty$.

Remark 2. If $\mathcal{N} = \mathbb{N}$, then Lemma 3.4 holds with $M = 1$ by the classical Khinchine Theorem. We do not know if Lemma 3.4 holds with $M = 1$ for any set $\mathcal{N}$ with positive lower density.

Proof of Theorem 3.1 *given Lemmas* 3.2–3.4. We apply these lemmas with $\psi(t) = c(\ln t)\,(\ln\ln t)\,(\ln\ln\ln t)$ and $c > 1/\ln(\frac{1+\sqrt{5}}{2})$.

Fix $M > 1$ as in Lemma 3.4. Then $\exists\Omega \subset (0,1)$ of full measure s.t. for every $\alpha \in \Omega$ there are infinitely many $(m,n) \in \mathbb{N}_0 \times \mathbb{N}$ as follows. Let $m^* := m/\gcd(m,n)$, $\quad n^* := n/\gcd(m,n)$, $\quad p := \gcd(m,n)$, then

(1) $pn^* \in \mathcal{N}$, $p \leq M$, $|n^*\alpha - m^*| = \frac{|n\alpha - m|}{p} \leq \frac{1}{n^*\psi(n^*)}$ $(\because n^* \leq n)$;

(2) $\exists k$ s.t. $n^* = q_k(\alpha)$ and $a_{k+1}(\alpha) \geq \frac{1}{2}\psi(q_k)$ ($\because$ Lemma 3.3). By its recursive definition, $q_k \geq k$-th Fibonacci number $\geq \frac{1}{3}(\frac{1+\sqrt{5}}{3})^k$. So for all k large enough, $a_{k+1}(\alpha) \geq \frac{1}{2}\psi(q_k) > \frac{1}{4}k(\ln k)(\ln\ln k)$;

(3) $a_{k+1}/(a_1 + \cdots + a_k) \geq \frac{1}{8}\ln\ln k \to \infty$ ($\because$ Lemma 3.2).

So every $\alpha \in \Omega$ belongs to $\mathcal{A} = \mathcal{A}(\mathcal{N}, M)$, and $\mathcal{A}$ has full measure. $\square$

Next we prove Lemmas 3.2–3.4.

3.2. **Proof of Lemma 3.2.** By [DV86], for almost every α

$$\frac{(a_1 + \cdots + a_{k+1}) - \max_{j \leq k+1} a_j}{k \ln k} \to \frac{1}{\ln 2} < 2.$$

So if k is large enough, and $a_{k+1} > \frac{1}{4}k(\ln k)(\ln\ln k)$ then

$$\max_{j \leq k+1} a_j = a_{k+1}, \quad \frac{a_1 + \cdots + a_k}{k \ln k} \leq 2, \text{ and } \frac{a_{k+1}}{a_1 + \cdots + a_k} > \frac{1}{8}\ln\ln k. \square$$

3.3. **Proof of Lemma 3.3.** For every (p, q) as in the lemma, $|q\alpha - p| < \frac{1}{2q}$. A classical result in the theory of continued fractions [Khi63, Thm 19] says that in this case $\exists k$ s.t. $q = q_k(\alpha), p = p_k(\alpha)$.

To estimate $a_{k+1} = a_{k+1}(\alpha)$ we recall the following facts, valid for the principal denominators of any irrational $\alpha \in (0, 1)$ [Khi63]:

(a) $|q_k\alpha - p_k| > \frac{1}{q_k + q_{k+1}}$;

(b) $q_{k+1} + q_k < (a_{k+1} + 2)q_k$, whence by (a) $a_{k+1} > \frac{1}{q_k|q_k\alpha - p_k|} - 2$.

In our case, $|q_k\alpha - p_k| = |q\alpha - p| \leq \frac{1}{q_k L}$, so $a_{k+1} > L - 2 \geq \frac{L}{2}$. $\square$

3.4. **Preparations for the proof of Lemma 3.4.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and $A_k \in \mathcal{F}$ be measurable events. Given $D > 1$, we say that A_k are *D-quasi-independent*, if

$$\mathbb{P}(A_{k_1} \cap A_{k_2}) \leq D\mathbb{P}(A_{k_1})\mathbb{P}(A_{k_2}) \text{ for all } k_1 \neq k_2. \tag{3.3}$$

The following proposition is a slight variation on Sullivan's Borel–Cantelli Lemma from ([Sul82]):

Proposition 3.5. *For every $D \geq 1$ there exists a constant $\delta(D) > 0$ such that the following holds in any probability space:*

(a) If A_k are D-quasi-independent measurable events s.t. $\lim\limits_{k\to\infty} \mathbb{P}(A_k) = 0$ but $\sum_k \mathbb{P}(A_k) = \infty$, then $\mathbb{P}(A_k$ occurs infinitely often$) \geq \delta(D)$.

(b) The quasi-independence assumption in (a) can be weakened to the assumption that for some $r \in \mathbb{N}$, $\mathbb{P}(A_{k_1} \cap A_{k_2}) \leq D\mathbb{P}(A_{k_1})\mathbb{P}(A_{k_2})$ for all $|k_2 - k_1| \geq r$.

(c) One can take $\delta(D) = \frac{1}{2D}$.

Proof. Since $\mathbb{P}(A_k) \to 0$ but $\sum \mathbb{P}(A_k) = \infty$, there is an increasing sequence N_j such that $\lim\limits_{j\to\infty} \sum\limits_{k=N_j+1}^{N_{j+1}} \mathbb{P}(A_k) = \frac{1}{D}$.

Let B_j be the event that at least one of events $\{A_k\}_{k=N_j+1}^{N_{j+1}}$ occurs. Since $B_j = \biguplus_{k=N_{j+1}}^{N_j+1} \big(A_k \setminus \bigcup_{j=N_j+1}^{k-1} A_j\big)$,

$$\begin{aligned}\mathbb{P}(B_j) &\geq \sum_{k=N_j+1}^{N_{j+1}} \mathbb{P}(A_k) - \sum_{N_j+1 \leq k_1 < k_2 \leq N_{j+1}} \mathbb{P}(A_{k_1} \cap A_{k_2}) \\ &\geq \sum_{k=N_j+1}^{N_{j+1}} \mathbb{P}(A_k) - D \sum_{N_j+1 \leq k_1 < k_2 \leq N_{j+1}} \mathbb{P}(A_{k_1})\mathbb{P}(A_{k_2}) \\ &\geq \sum_{k=N_j+1}^{N_{j+1}} \mathbb{P}(A_k) - \frac{D}{2}\left(\sum_{k=N_j+1}^{N_{j+1}} \mathbb{P}(A_k)\right)^2 .\end{aligned}$$

Since $\displaystyle\lim_{j\to\infty} \sum_{k=N_j+1}^{N_{j+1}} \mathbb{P}(A_k) = \frac{1}{D}$ and $D \geq 1$, $\liminf \mathbb{P}(B_j) \geq \frac{1}{2D}$.

Let E denote the event that A_j happens infinitely often. E is also the event that B_j happens infinitely often, therefore $E = \bigcap_{n=1}^{\infty} \bigcup_{j=n+1}^{\infty} B_j$. In a probability space, the measure of a decreasing intersection of sets is the limit of the measure of these sets. So $\mathbb{P}(E) \geq \liminf \mathbb{P}(B_j) \geq \frac{1}{2D}$, proving (a) and (c).

Part (b) follows from part (a) by applying it to the sets $\{A_{kr+\ell}\}$ where $0 \leq \ell \leq r-1$ is chosen to get $\displaystyle\sum_k \mathbb{P}(A_{kr+\ell}) = \infty$. □

The *multiplicity* of a collection of measurable sets $\{E_k\}$ is defined to be the largest K s.t. there are K different k_i with $\mathbb{P}(\bigcap_{i=1}^{K} E_{k_i}) > 0$.

Proposition 3.6. *Let E_k be measurable sets in a finite measure space. If the multiplicity of $\{E_k\}$ is less than K, then*

$$\operatorname{mes}\left(\bigcup_k E_k\right) \geq \frac{1}{K} \sum_k \operatorname{mes}(E_k).$$

Proof. $1_{\bigcup_i E_i} \geq \frac{1}{K} \sum_i 1_{E_i}$ almost everywhere. □

Proposition 3.7. *For every non-empty open interval $I \subset [0,1]$, $\mathrm{Card}\{(m,n) \in \{0,\dots,N\}^2 : \frac{m}{n} \in I \text{ , } \gcd(m,n) = 1\} \sim 3\mathrm{mes}(I)N^2/\pi^2$, as $N \to \infty$.*

Proof. This classical fact due to Dirichlet follows from the inclusion-exclusion principle and the identity $\zeta(2) = \pi^2/6$, see [HW08, Theorem 459]. □

Proposition 3.8. *Suppose* $\alpha = [0; a_1, a_2, \ldots]$ *and* $\overline{\alpha} = [0; a_{\ell+1}, a_{\ell+2}, \ldots]$. *Then the principal convergents* $\overline{p}_{\overline{\ell}}/\overline{q}_{\overline{\ell}}$ *of* $\overline{\alpha}$ *and the principal convergents* p_ℓ/q_ℓ *of* α *are related by* $\begin{pmatrix} p_{l+\bar{l}} & p_{l+\bar{l}+1} \\ q_{l+\bar{l}} & q_{l+\bar{l}+1} \end{pmatrix} = \begin{pmatrix} p_{l-1} & p_l \\ q_{l-1} & q_l \end{pmatrix} \begin{pmatrix} \overline{p}_{\bar{l}} & \overline{p}_{\bar{l}+1} \\ \overline{q}_{\bar{l}} & \overline{q}_{\bar{l}+1} \end{pmatrix}$

Proof. Since $a_0 = 0$, the recurrence relations for p_n/q_n imply

$$\begin{pmatrix} p_n & p_{n+1} \\ q_n & q_{n+1} \end{pmatrix} = \begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & a_{n+1} \end{pmatrix}, \quad \begin{pmatrix} p_0 & p_1 \\ q_0 & q_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & a_1 \end{pmatrix}.$$

So $\begin{pmatrix} p_n & p_{n+1} \\ q_n & q_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & a_1 \end{pmatrix} \cdot \ldots \cdot \begin{pmatrix} 0 & 1 \\ 1 & a_{n+1} \end{pmatrix}$. It follows that $\begin{pmatrix} p_{l+\bar{l}} & p_{l+\bar{l}+1} \\ q_{l+\bar{l}} & q_{l+\bar{l}+1} \end{pmatrix} = \begin{pmatrix} p_{l-1} & p_l \\ q_{l-1} & q_l \end{pmatrix} \begin{pmatrix} \overline{p}_{\bar{l}} & \overline{p}_{\bar{l}+1} \\ \overline{q}_{\bar{l}} & \overline{q}_{\bar{l}+1} \end{pmatrix}$, where $\overline{p}_i/\overline{q}_i$ are the principal convergents of $\overline{\alpha} := [0; a_{l+1}, a_{l+2}, \ldots]$. □

3.5. **Proof of Lemma 3.4.** Without loss of generality, $\lim\limits_{t\to\infty} \psi(t) = \infty$, otherwise replace $\psi(t)$ by the bigger monotone function $\psi(t) + \ln t$.

Fix $M > 1$, to be determined later. Let

$\Omega_k := \{(m,n) \in \mathbb{N}^2 : n \in \mathcal{N}, n \in [e^{k-1}, e^k], 0 < m < n, \gcd(m,n) \le M\}$,

$A_{m,n,k} := \{\alpha \in \mathbb{T} : |n\alpha - m| \le \frac{1}{e^k \psi(e^k)}\}$,

$\mathcal{A}_k := \bigcup\limits_{(m,n)\in\Omega_k} A_{m,n,k}$,

$\mathcal{A} := \{\alpha \in \mathbb{T} : \alpha \text{ belongs to infinitely many } \mathcal{A}_k\}$.

The lemma is equivalent to saying that $\mathcal{A}$ has full Lebesgue measure for a suitable choice of M.

We will prove a slightly different statement. Fix $\varepsilon > 0$ small. Given an non-empty interval $I \subset [\varepsilon, 1-\varepsilon]$, let

$$\Omega_k(I) := \{(m,n) \in \Omega_k : \frac{m}{n} \in I\}$$
$$\mathcal{A}_k(I) := \bigcup_{(m,n)\in\Omega_k(I)} A_{m,n,k}$$
$$\mathcal{A}(I) := \{\alpha \in \mathbb{T} : \alpha \text{ belongs to infinitely many } \mathcal{A}_k(I)\}.$$

We will prove that there exists a positive constant $\delta = \delta(\varepsilon, M)$ s.t. for all intervals $I \subset [\varepsilon, 1-\varepsilon]$, $\mathrm{mes}(\mathcal{A}(I) \cap I) \ge \delta \mathrm{mes}(I)$. It then follows by a standard density point argument (see below) that $\mathcal{A} \cap [\varepsilon, 1-\varepsilon]$ has full measure. Since ε is arbitrary, the lemma is proved.

Claim 1. There exist $K = K(\varepsilon)$ s.t. for every $k > K$, the multiplicity of $\{A_{m,n,k}\}_{(m,n)\in\Omega_k(I)}$ is uniformly bounded by M.

PROOF: Suppose $(m_i, n_i) \in \Omega_k(I)$ $(i = 1, 2)$ and $A_{m_1,n_1,k} \cap A_{m_2,n_2,k} \neq \varnothing$. Then there is α s.t. $|n_i\alpha - m_i| \leq \delta_k := \frac{1}{e^k\psi(e^k)}$ $(i = 1, 2)$. Choose $K = K(\varepsilon)$ so large that $k > K \Rightarrow \delta_k < \frac{\varepsilon}{4e^k}$.

If $k > K$, then $\alpha \geq \frac{m_i}{n_i} - \delta_k > \min I - \frac{\varepsilon}{2} > \frac{\varepsilon}{2}$. Let $r_i := \gcd(m_i, n_i)$ and $(n_i^*, m_i^*) := \frac{1}{r_i}(n_i, m_i)$. Then $|n_i^*\alpha - m_i^*| \leq \delta_k$ and $m_i^* \leq n_i^* \leq n_i \leq e^k$, so $|n_2^*m_1^* - n_1^*m_2^*| = \frac{1}{\alpha}|m_1^*(n_2^*\alpha - m_2^*) - m_2^*(n_1^*\alpha - m_1^*)| \leq \frac{2e^k\delta_k}{\varepsilon/2} < 1$. So $n_2^*m_1^* = n_1^*m_2^*$. Since $\gcd(n_i^*, m_i^*) = 1$, $(n_1^*, m_1^*) = (n_2^*, m_2^*)$. It follows that $(n_2, m_2) \in \{(rn_1^*, rm_1^*) : r = 1, \ldots, M\}$. So the multiplicity of $\{A_{m,n,k}\}_{(m,n)\in\Omega_k(I)}$ is uniformly bounded by M.

CLAIM 2. Let $d(\mathcal{N}) := \liminf \frac{1}{N}\mathrm{Card}(\mathcal{N} \cap [1, N]) > 0$, then there exists $M = M(\mathcal{N})$ and $\widetilde{K} = \widetilde{K}(\varepsilon, \mathcal{N}, |I|)$ s.t. for all $k > \widetilde{K}$,

$$\frac{d(\mathcal{N})\mathrm{mes}(I)}{4M\psi(e^k)} \leq \mathrm{mes}(\mathcal{A}_k(I)) \leq \frac{6\mathrm{mes}(I)}{\psi(e^k)}. \tag{3.4}$$

In particular, $\mathrm{mes}(\mathcal{A}_k(I)) \xrightarrow[k\to\infty]{} 0$ and $\sum \mathrm{mes}(\mathcal{A}_k(I)) = \infty$.

PROOF: $\mathrm{mes}(A_{m,n,k}) = \mathrm{mes}\left(\left[\frac{m}{n} - r_{m,n}, \frac{m}{n} + r_{m,n}\right]\right) = 2r_{m,n}$ where $r_{m,n} = \frac{1}{ne^k\psi(e^k)}$. Since $n \in [e^{k-1}, e^k]$,

$$\frac{\mathrm{Card}(\Omega_k(I))}{Me^{2k}\psi(e^k)} \leq \mathrm{mes}\big(\mathcal{A}_k(I)\big) \leq \frac{e\ \mathrm{Card}(\Omega_k(I))}{e^{2k}\psi(e^k)}, \tag{3.5}$$

where the lower bound uses Claim 1 and Proposition 3.6.

$\mathrm{Card}(\Omega_k(I))$ satisfies the bounds $A - B \leq \mathrm{Card}(\Omega_k(I)) \leq A$ where

$$A := \mathrm{Card}\{(m,n) : n \in \mathcal{N},\ n \in [e^{k-1}, e^k],\ \frac{m}{n} \in I\}$$
$$B := \mathrm{Card}\{(m,n) : n \in \mathcal{N},\ n \in [e^{k-1}, e^k],\ \frac{m}{n} \in I, \gcd(m,n) \geq M\}.$$

Choose $\widetilde{K} = \widetilde{K}(\varepsilon, \mathcal{N}, |I|) > K(\varepsilon)$ s.t. for all $k > \widetilde{K}$

(a) $\mathrm{Card}\{n \in \mathcal{N} : 0 \leq n \leq e^k\} \geq \frac{1}{\sqrt{2}}d(\mathcal{N})$
(b) $\mathrm{Card}\{n \in [e^{k-1}, e^k] \cap \mathbb{N} : p|n\} \leq 2(e^k - e^{k-1})/p$ for all $p \geq 1$;
(c) For all $n > e^{\widetilde{K}-1}$, $p \geq 1$,

$$\frac{n}{p\sqrt{2}}\mathrm{mes}(I) \leq \mathrm{Card}\{m \in \mathbb{N} : \frac{m}{n} \in I, p|m\} \leq \frac{2n}{p}\mathrm{mes}(I).$$

If $k > \widetilde{K}$, then $\frac{1}{2}d(\mathcal{N})e^{2k}\mathrm{mes}(I) \leq A \leq 2e^{2k}\mathrm{mes}(I)$ and

$$B \leq \sum_{p=M}^{\infty} \mathrm{Card}\{(m,n) : n \in [e^{k-1}, e^k],\ \frac{m}{n} \in I, p|m, p|n\}$$

$$\leq \sum_{p=M}^{\infty} \frac{2(e^k - e^{k-1})}{p} \cdot \frac{2e^k \mathrm{mes}(I)}{p} < 4e^{2k}\mathrm{mes}(I) \sum_{p=M}^{\infty} \frac{1}{p^2}$$

$$\leq \frac{1}{4} d(\mathcal{N}) e^{2k} \mathrm{mes}(I), \text{ provided we choose } M \text{ s.t. } \sum_{p=M}^{\infty} p^{-2} < \frac{1}{16} d(\mathcal{N}).$$

Together we get $\frac{1}{4}d(\mathcal{N})e^{2k}\mathrm{mes}(I) \leq \mathrm{Card}(\Omega_k(I)) \leq 2e^{2k}\mathrm{mes}(I)$. The claim now follows from (3.5).

Claim 3. There exists $D = D(\mathcal{N}, M)$, $r = r(M)$, and $\hat{K} = \hat{K}(\varepsilon, \mathcal{N}, I)$ s.t. for all $k_1, k_2 > \hat{K}$ s.t. $|k_1 - k_2| > r(M)$,

$$\mathrm{mes}(\mathcal{A}_{k_1}(I) \cap \mathcal{A}_{k_2}(I)|I) \leq D\mathrm{mes}(\mathcal{A}_{k_1}(I)|I)\mathrm{mes}(\mathcal{A}_{k_2}(I)|I) \tag{3.6}$$

Proof: By Claim 2, if k_1, k_2 are large enough, then

$$\mathrm{mes}(\mathcal{A}_{k_1}(I)|I)\mathrm{mes}(\mathcal{A}_{k_2}(I)|I) \geq \left(\frac{d(\mathcal{N})}{5M}\right)^2 \frac{1}{\psi(e^{k_1})\psi(e^{k_2})}, \tag{3.7}$$

where we put 5 instead of 4 in the denominator to deal with edge effects arising from $\mathrm{mes}(\mathcal{A}_k(I) \setminus I) = O\left(\frac{1}{e^k \psi(e^k)}\right)$.

To prove the claim, it remains to bound $\mathrm{mes}\left(\mathcal{A}_{k_1}(I) \cap \mathcal{A}_{k_2}(I) \middle| I\right)$ from above by $\frac{\mathrm{const}}{R_1 R_2}$, where $R_i := \psi(e^{k_i})$.

A *cylinder* is a set of the form

$$[\![a_1, \ldots, a_n]\!] = \{\alpha \in (0,1) \setminus \mathbb{Q} : a_i(\alpha) = a_i \quad (1 \leq i \leq n)\}.$$

Equivalently, $\alpha \in [\![a_1, \ldots, a_n]\!]$ iff α has an infinite continued fraction expansion of the form $\alpha = [0; a_1, \ldots, a_n, *, *, \ldots]$.

Our plan is to cover $\mathcal{A}_{k_i}(I)$ by unions of cylinders of total measure $O(1/R_i)$, and then use the following well-known fact: There is a constant $G > 1$ s.t. for any $(a_1, \ldots, a_n, b_1, \ldots, b_m) \in \mathbb{N}^{n+m}$,

$$G^{-1} \leq \frac{\mathrm{mes}[\![a_1, \ldots, a_n; b_1, \ldots, b_m]\!]}{\mathrm{mes}[\![a_1, \ldots, a_n]\!]\mathrm{mes}[\![b_1, \ldots, b_m]\!]} \leq G. \tag{3.8}$$

This is because the invariant measure $\frac{1}{\ln 2}\frac{dx}{1+x}$ of $T : (0,1) \to (0,1)$, $T(x) = \{\frac{1}{x}\}$ (the Gauss map) is a Gibbs-Markov measure, because of the bounded distortion of T, see §2 in [ADU93].

To cover $\mathcal{A}_{k_i}(I)$ by cylinders, it is enough to cover A_{m,n,k_i} by cylinders for every $(m,n) \in \Omega_{k_i}(I)$. Suppose $\alpha \in A_{m,n,k_i}$. Then $r := \gcd(m,n) \leq M$ and $(m^*, n^*) := \frac{1}{r}(m,n)$ satisfies

$$\gcd(m^*, n^*) = 1, \; n^* \in \bigcup_{|k_i^* - k_i| \leq \ln M} [e^{k_i^* - 1}, e^{k_i^*}], \; |n^*\alpha - m^*| < \frac{1}{n^* R_i}.$$

Assume k_i is so large that $R_i = \psi(e^{k_i}) \geq 4$. Then Lemma 3.3 gives $a_{l_i+1} > \frac{R_i}{2}$. Thus $\mathcal{A}_{k_i}(I) \subset \mathcal{C}_{k_i}(I, R_i)$ where

$$\mathcal{C}_k(I,R) := \bigcup_{k^* \in [k - \ln M, k]} \left\{ \alpha \in (0,1) \setminus \mathbb{Q} : \exists \ell \text{ s.t. } \begin{array}{l} q_\ell(\alpha) \in [e^{k^*-1}, e^{k^*}], \\ a_{\ell+1}(\alpha) \geq R/2 \\ p_\ell(\alpha)/q_\ell(\alpha) \in I \end{array} \right\}.$$

This is a union of cylinders, because $q_\ell(\alpha), p_\ell(\alpha), a_{\ell+1}(\alpha)$ are constant on cylinders of length $\ell + 1$.

We claim that for some $c^*(M)$ which only depends on M, for all k_i large enough,

$$\mathrm{mes}(\mathcal{C}_{k_i}(I, R_i)) \leq \frac{c^*(M)\mathrm{mes}(I)}{R_i}. \tag{3.9}$$

Every rational $\frac{m}{n} \in (0,1)$ has two finite continued fraction expansions: $[0; a_1, \ldots, a_\ell]$ and $[0; a_1, \ldots, a_\ell - 1, 1]$ with $a_\ell > 1$. We write $\ell = \ell(\frac{m}{n})$ and $a_i = a_i(\frac{m}{n})$. With this notation

$$\mathcal{C}_{k_i}(I,R_i) = \bigcup_{\substack{k_i^* \in [k_i - \ln M, k_i] \\ n \in [e^{k_i^*-1}, e^{k_i^*}]}} \bigcup_{\substack{\gcd(m,n)=1 \\ m/n \in I}} \bigcup_{b > R_i/2} [\![a_1(\tfrac{m}{n}), \ldots, a_{\ell(\frac{m}{n})}(\tfrac{m}{n}), b]\!]$$

$$\cup\, [\![a_1(\tfrac{m}{n}), \ldots, a_{\ell(\frac{m}{n})}(\tfrac{m}{n}) - 1, 1, b]\!].$$

We have $[\![a_1, \ldots, a_\ell]\!] = (\frac{p_\ell + p_{\ell-1}}{q_\ell + q_{\ell-1}}, \frac{p_\ell}{q_\ell})$ or $(\frac{p_\ell}{q_\ell}, \frac{p_\ell + p_{\ell-1}}{q_\ell + q_{\ell-1}})$, depending on the parity of ℓ [Khi63]. Since $|p_\ell q_{\ell-1} - p_{\ell-1} q_\ell| = 1$ and $q_{\ell+1} = a_{\ell+1} q_\ell + q_{\ell-1}$, we have $\mathrm{mes}([\![a_1, \ldots, a_\ell, b]\!]) = \frac{1}{q_{\ell+1}(q_{\ell+1} + q_\ell)} = \frac{1}{(bq_\ell + q_{\ell-1})((b+1)q_\ell + q_{\ell-1})} \leq \frac{1}{b(b+1)q_\ell^2}$, leading to

$$\mathrm{mes}(\mathcal{C}_{k_i}(I,R_i)) \leq \sum_{k_i^* \in [k_i - \ln M, k_i]} \sum_{n \in [e^{k_i^*-1}, e^{k_i^*}]} \sum_{\substack{\gcd(m,n)=1 \\ m/n \in I}} \sum_{b > R_i/2} \frac{2}{n^2 b(b+1)}$$

$$\leq \frac{8 \ln M}{e^{2(k_i - 1 - \ln M)} R_i} \sum_{n=1}^{e^{k_i} M} \#\{m \in \mathbb{N} : \frac{m}{n} \in I, \gcd(m,n) = 1\} \leq \frac{c^*(M)}{R_i} \mathrm{mes}(I)$$

where $c^*(M)$ only depends on M. The last step uses Prop. 3.7.

Next we cover $\mathcal{A}_{k_1}(I) \cap \mathcal{A}_{k_2}(I)$ by cylinders. Suppose without loss of generality that $k_2 > k_1$. Arguing as before one sees that if

$$k_2 > k_1 + \ln M + 1, \tag{3.10}$$

then $\mathcal{A}_{k_1}(I) \cap \mathcal{A}_{k_2}(I)$ can be covered by sets $[\![a_1, \ldots, a_\ell, b, \overline{a}_1, \ldots, \overline{a}_{\overline{l}}, \overline{b}]\!]$ as follows: The convergents p_i/q_i of (every) α in $[\![a_1, \ldots, a_\ell, b, \overline{a}_1, \ldots, \overline{a}_{\overline{l}}, \overline{b}]\!]$, $(1 \leq i \leq l + \overline{l} + 2)$, satisfy

(a) $q_l \in [e^{k_1^*-1}, e^{k_1^*}]$, $k_1^* \in [k_1 - \ln M, k_1]$, $p_l/q_l \in I$, $b \geq R_1/2$;

(b) $q_{l+\bar{l}+1} \in [e^{k_2^*-1}, e^{k_2^*}]$, $k_2^* \in [k_2 - \ln M, k_2]$, $p_{\bar{l}}/q_{\bar{l}} \in I$, $\bar{b} \geq R_2/2$
(c) $k_2^* > k_1^*$ (this is where (3.10) is used).

We claim that

$$[\![a_1, \ldots, a_\ell, b]\!] \subset \mathcal{C}_{k_1}(I), \tag{3.11}$$

$$b \leq e^{k_2^*-k_1^*+1}, \tag{3.12}$$

$$[\![\bar{a}_1, \ldots, \bar{a}_{\bar{\ell}}, \bar{b}]\!] \subset \bigcup_{|r|\leq 3} \mathcal{C}_{k_2-k_1+r-\ln b}([0,1], R_2). \tag{3.13}$$

(3.11) follows from (a). Next, $e^{k_2^*} \geq q_{l+1} \geq bq_l \geq be^{k_1^*-1}$ proving (3.12). To prove (3.13), let $\bar{p}_i/\bar{q}_i$, $1 \leq i \leq \bar{\ell}+2$, be the principal convergents of (every) $\overline{\alpha} \in [\![b, \bar{a}_1, \ldots, \bar{a}_{\bar{\ell}}, \bar{b}]\!]$. By Prop. 3.8, $q_{l+1+\bar{l}} = q_{l-1}\bar{p}_{\bar{l}+1} + q_l\bar{q}_{\bar{l}+1}$, whence $q_l\bar{q}_{\bar{l}+1} \leq q_{l+1+\bar{l}} \leq 2q_l\bar{q}_{\bar{l}+1}$. Since $q_l \in [e^{k_1^*-1}, e^{k_1^*}]$ and $q_{l+\bar{l}+1} \in [e^{k_2^*-1}, e^{k_2^*}]$,

$$e^{k_2^*-k_1^*-2} \leq \frac{q_{l+1+\bar{l}}}{2q_l} \leq \bar{q}_{\bar{l}+1} \leq \frac{q_{l+1+\bar{l}}}{q_l} \leq e^{k_2^*-k_1^*+1}. \tag{3.14}$$

Next, let $\widetilde{p}_i/\widetilde{q}_i$ $(1 \leq i \leq \bar{l})$ denote the principal convergents of (every) $\widetilde{\alpha} \in [\![\bar{a}_1, \ldots, \bar{a}_{\bar{\ell}}, \bar{b}]\!]$. Then $\frac{\bar{p}_{\bar{l}+1}}{\bar{q}_{\bar{l}+1}} = 1/(b + \frac{\widetilde{p}_{\bar{l}}}{\widetilde{q}_{\bar{l}}})$, so $\bar{q}_{\bar{l}+1} = b\widetilde{q}_{\bar{l}} + \widetilde{p}_{\bar{l}}$, whence $b\widetilde{q}_{\bar{l}} \leq \bar{q}_{\bar{l}+1} \leq (b+1)\widetilde{q}_{\bar{l}}$. Thus $\widetilde{q}_{\bar{l}} \in [(b+1)^{-1}\bar{q}_{\bar{l}+1}, b^{-1}\bar{q}_{\bar{l}+1}]$. It follows that the $\bar{l}$-th principal convergent of every $\widetilde{\alpha} \in [\![\bar{a}_1, \ldots, \bar{a}_{\bar{\ell}}, \bar{b}]\!]$ satisfies

$$\widetilde{q}_{\bar{l}} \in [e^{k_2^*-k_1^*-3-\ln b}, e^{k_2^*-k_1^*+1-\ln b}]. \tag{3.15}$$

It is now easy to see (3.13).

By (3.13), $\mathcal{A}_{k_1}(I) \cap \mathcal{A}_{k_2}(I) \cap I \subset \bigcup_{|r|\leq 3} \biguplus_{[\underline{a},b]\subset \mathcal{C}_{k_1}(I)} \biguplus_{[\underline{a}',b']\subset \mathcal{C}_{k_2-k_1+r-\ln b}([0,1])} [\![\underline{a}, b, \underline{a}', b]\!]$.

Now arguing as in the proof of (3.9) and using (3.8) we obtain

$$\mathrm{mes}(\mathcal{A}_{k_1}(I) \cap \mathcal{A}_{k_2}(I) \cap I) \leq$$

$$\leq \sum_{\substack{k_i^* \in [k_i - \ln M, k_i] \\ i=1,2; |r|\leq 3}} \sum_{n \in [e^{k_1^*-1}, e^{k_1^*}]} \sum_{\substack{\gcd(m,n)=1 \\ m/n \in I}} \sum_{b=[R_1/2]}^{[\exp(k_2^*-k_1^*+1)]} \frac{2G\mathrm{mes}(\mathcal{C}_{k_2-k_1+r-\ln b}([0,1], R_2))}{n^2 b(b+1)}$$

$$\leq \frac{\mathrm{const}\,\mathrm{mes}(I)}{R_1 R_2}. \tag{3.16}$$

(3.16) uses the estimate $\mathrm{mes}(\mathcal{C}_{k_2-k_1+r-\ln b}([0,1], R_2)) = O(1/R_2)$ which is also valid when $k_2 - k_1 + r - \ln b$ is small, provided we choose M large enough so that the asymptotic in Prop. 3.7 holds for all $N > M$ with $I = [0,1]$. See the proof of (3.9).

Combining (3.16) with (3.7), we find that under (3.10) $\mathcal{A}_{k_i}(I)$ are D-quasi-independent for sufficiently large D, proving Claim 3.

Claims 2 and 3 allow us to apply Sullivan's Borel–Cantelli Lemma (Prop. 3.5). We obtain $\delta = \delta(M)$ s.t. for every interval $I \subset [\varepsilon, 1-\varepsilon]$, $\mathrm{mes}(\mathcal{A} \cap I) \geq \delta \mathrm{mes}(I)$. This means that $[\varepsilon, 1-\varepsilon] \setminus \mathcal{A}$ has no Lebesgue density points, and therefore must have measure zero. So $\mathcal{A}$ has full measure in $[\varepsilon, 1-\varepsilon]$. Since ε is arbitrary, $\mathcal{A}$ has full measure. □

4. Proof of Theorem 1.2

As explained in Section 2, it is enough to prove Theorem 1.2 for $f(x) := \sum_{m=1}^{d} b_m h(x+\beta_m) \not\equiv 0$ with $h(x) = \{x\} - \frac{1}{2}$. Without loss of generality, β_i are different and $b_i \neq 0$. Notice that

$$f(x) = -\sum_{m=1}^{d} b_m \sum_{j=1}^{\infty} \frac{\sin(2\pi j(x+\beta_m))}{\pi j}.$$

Therefore $\|f\|_2^2 = \dfrac{1}{2\pi^2} \sum_n \dfrac{1}{n^2} D(\beta_1 n, \dots, \beta_d n)$, where $D : \mathbb{T}^d \to \mathbb{R}$ is

$$D(\gamma_1, \dots, \gamma_d) := \int_0^1 \left[\sum_{m=1}^{d} b_m \sin(2\pi(y+\gamma_m)) \right]^2 dy. \tag{4.1}$$

Since $f \not\equiv 0$, $D(\beta_1 n, \dots, \beta_d n) > 0$ for some n. Let $\mathfrak{T}$ denote the closure in $\mathbb{T}^d$ of $\mathbb{O} := \{(\beta_1 n, \dots, \beta_d n) \bmod \mathbb{Z} : n \in \mathbb{Z}\}$. This is a minimal set for the translation by $(\beta_1, \dots, \beta_d)$ on $\mathbb{T}^d$, so a standard compactness argument shows that for every $\varepsilon_0 > 0$, the set

$$\mathcal{N} := \{n \in \mathbb{N} : \; D(\beta_1 n, \dots, \beta_d n) > \varepsilon_0\} \tag{4.2}$$

is syndetic: its gaps are bounded. Thus $\mathcal{N}$ has positive lower density.

By Theorem 3.1, if M is sufficiently large then the set $\mathcal{A} := \mathcal{A}(\mathcal{N}, M)$ has full measure in $\mathbb{T}$. Let

$$S_n(\alpha, x) := \sum_{k=0}^{n-1} f(x+k\alpha).$$

The proof of Theorem 1.2 for $f(x)$ above consists of two parts:

Theorem 4.1. *Suppose $\alpha \in \mathcal{A}$, then for a.e. $x \in [0,1)$, there exist $A_k(x) \in \mathbb{R}$ and $B_k(x), N_k(x) \to \infty$ such that*

$$\frac{S_n(\alpha, x) - A_k(x)}{B_k(x)} \xrightarrow[k\to\infty]{dist} \mathrm{U}[0,1], \;\; \textit{as } n \sim \mathrm{U}(0, \dots, N_k(x)).$$

Theorem 4.2. *Suppose $\alpha \in \mathcal{A}$, then for a.e. $x \in [0,1)$, there are no $A_N(x) \in \mathbb{R}$ and $B_N(x) \to \infty$ such that*

$$\frac{S_n(\alpha, x) - A_N(x)}{B_N(x)} \xrightarrow[N\to\infty]{dist} \mathrm{U}[0,1], \;\; \textit{as } n \sim U(0, \dots, N).$$

4.1. **Preliminaries.**

Lemma 4.3. *$S_q(\alpha,\cdot):\mathbb{T}\to\mathbb{R}$ has dq discontinuities.*

Proof. The discontinuities of S_q are preimages of discontinuities of f by R_α^{-k} with $k=0,1,\dots,q-1$. $\square$

Lemma 4.4. *Let $C:=\sup|f'|\le|\sum b_m|$. If x',x'' belong to same continuity component of R_α^r then*

$$|S_r(\alpha,x')-S_r(\alpha,x'')|\le Cr|x'-x''|.$$

Proof. Since $|S_r'|=\left|\sum_{k=0}^{r-1}f'(x+k\alpha)\right|\le Cr$, the restriction of S_r to on each continuity component is Lipshitz with Lipshitz constant Cr. $\square$

Lemma 4.5. *There are constants C_1,C_2 such that the following holds. Suppose that q_n is a principal denominator of α, and $q_{n+1}>cq_n$ with $c>1$. Let $\mu_n(x):=S_{q_n}(\alpha,x)$, then*

$$\text{mes}\left\{x:\ S_{\ell q_n}(\alpha,x)=\ell\mu_n\pm C_1\frac{\ell^2}{c}\ \textit{for}\ \ell=0,\dots,k\right\}>1-C_2\frac{k}{c}. \quad (4.3)$$

Proof. If x and $x+\ell q_n\alpha$ belong to the same continuity interval of $R_\alpha^{q_n}$ for all $\ell=0,\dots,k$ then we have by Lemma 4.4 that for $\ell\le k$

$$|S_{\ell q_n}(\alpha,x)-\ell\mu_n|\le\sum_{j=0}^{\ell-1}|S_{q_n}(\alpha,x+jq_n\alpha)-S_{q_n}(\alpha,x)|\le Cq_n\sum_{j=0}^{\ell-1}\|jq_n\alpha\|$$

$$\le\frac{Cq_n}{q_{n+1}}\sum_{j=0}^{\ell-1}j\le\frac{C_1\ell^2}{c},\ \text{where } C_1:=C/2.$$

Therefore if $S_{\ell q_n}(\alpha,x)\neq\ell\mu_n\pm C_1\frac{\ell^2}{c}$ for some $\ell=0,\dots,k$, then there must exist $0\le\ell\le k$ s.t. $x,R_\alpha^{\ell q_n}(x)$ are separated by a discontinuity of $S_{q_n}(\alpha,\cdot)$. Since $\text{dist}(x,R_\alpha^{\ell q_n}(x))\le\ell/q_{n+1}$, x must belong to a ball with radius k/q_{n+1} centered at a discontinuity of $S_{q_n}(\alpha,\cdot)$. By Lemma 4.3, there are dq_n discontinuities, so the measure of such points is less than $dq_n\left(\frac{2k}{q_{n+1}}\right)\le\frac{2dk}{c}$. The lemma follows with $C_2:=2d$. $\square$

Lemma 4.6. *There is a constant $C_3=C_3(b_1,\dots,b_d)$ s.t. for every $n\ge1$ and $\alpha=[0;a_1,a_2,\dots]$, $\max\{|S_r(\alpha,x)|:0\le r\le q_n-1\}\le C_3(a_0+\dots+a_{n-1})$.*

Proof. Let $r=\sum_{j=0}^{n-1}\mathfrak{b}_jq_j$ denote the Ostrowski expansion of r. Recall that this means that $0\le\mathfrak{b}_j\le a_j$ and $\mathfrak{b}_j=a_j\Rightarrow\mathfrak{b}_{j-1}=0$. So

$$S_r=\sum_{k=0}^{\mathfrak{b}_{n-1}-1}S_{q_{n-1}}\circ R_\alpha^{q_{n-1}k}+\sum_{k=0}^{\mathfrak{b}_{n-2}-1}S_{q_{n-2}}\circ R_\alpha^{q_{n-2}k}+\dots+\sum_{k=0}^{\mathfrak{b}_0-1}S_{q_0}\circ R_\alpha^{q_0k}.$$

By the Denjoy-Koksma inequality $|S_r| \le \sum \mathfrak{b}_j \mathsf{V}(f) \le \mathsf{V}(f)\sum a_j$ where $\mathsf{V}(f) \le 2\sum b_i$ is the total variation of f on $\mathbb{T}$. $\square$

Lemma 4.7. *There exist positive constants $\varepsilon_1, \varepsilon_2$ such that for every α irrational, if q_n is a principal denominator of α and $q_n r_n \in \mathcal{N}$ with $r_n \le M$ then* $\operatorname{mes}\{x : |S_{q_n}(\alpha, x)| \ge \varepsilon_1\} \ge \varepsilon_2$.

Proof. We follow an argument from [Bec94]. Suppose q_n is a principal denominator of α and $q_n r_n \in \mathcal{N}$ for some $r_n \le M$. Let $N = q_n r_n$. Since $f(x) = -\sum_{m=1}^d b_m \sum_{j=1}^\infty \frac{\sin(2\pi j(x+\beta_m))}{\pi j}$, for each $j \in \mathbb{N}$

$$\|S_N(\alpha, \cdot)\|_{L^2}^2 \ge \frac{1}{\pi^2 j^2}\int_0^1 \left(\sum_{m=1}^d b_m \sum_{k=0}^{N-1} \sin(2\pi j(x + k\alpha + \beta_m))\right)^2 dx.$$

Using the identities $\sum_{k=1}^N \sin(y+kx) = \frac{\cos(y+x/2)-\cos(y+(2N+1)x/2)}{2\sin(x/2)}$ and $\cos A - \cos B = 2\sin(\frac{A+B}{2})\sin(\frac{B-A}{2})$ we find that

$$\|S_N(\alpha,\cdot)\|_{L^2}^2 \ge$$

$$\ge \left(\frac{\sin(\pi N j\alpha)}{\pi j \sin(\pi j \alpha)}\right)^2 \int_0^1 \left(\sum_{m=1}^d b_m \sin\left(2\pi\left(jx + j\tfrac{(N-1)\alpha}{2}\right) + 2\pi j\beta_m\right)\right)^2 dx$$

$$= \left(\frac{\sin(\pi N j\alpha)}{\pi j \sin(\pi j \alpha)}\right)^2 \int_0^1 \left(\sum_{m=1}^d b_m \sin\left(2\pi(y + j\beta_m)\right)\right)^2 dy$$

$$= \left(\frac{\sin(\pi N j\alpha)}{\pi j \sin(\pi j \alpha)}\right)^2 D(j\beta_1, \dots, j\beta_m) \text{ with } D \text{ as in (4.1).}$$

We now take $j = N = r_n q_n$. The first term is bounded below because $\|N\alpha\| \le M\|q_n\alpha\| \le \frac{M}{q_{n+1}} \le \frac{M}{a_{n+1}q_n} \le \frac{M^2}{a_{n+1}N} = o(\frac{1}{N})$, so $\frac{\sin(\pi N^2\alpha)}{\pi N \sin(\pi N\alpha)} \xrightarrow[n\to\infty]{} \pi^{-1}$. The second term is bounded below by ε_0, because $N = q_n r_n \in \mathcal{N}$. It follows that for all n large enough, $\|S_{r_n q_n}(\alpha,\cdot)\|_2 > \sqrt{\varepsilon_0}/2\pi$.

For any L^2-function φ and any $\hat\varepsilon > 0$,

$$\|\varphi\|_{L^2}^2 \le \|\varphi\|_{L^\infty}^2 \operatorname{mes}\{x : |\varphi(x)| \ge \hat\varepsilon\} + \hat\varepsilon^2.$$

Hence $\operatorname{mes}\{x : |\varphi(x)| \ge \hat\varepsilon\} \ge \dfrac{\|\varphi\|_{L^2}^2 - \hat\varepsilon^2}{\|\varphi\|_{L^\infty}^2}$. We just saw that for all n large enough, $\|S_{r_n q_n}(\alpha,\cdot)\|_2 > \sqrt{\varepsilon_0}/2\pi$, and by the Denjoy-Koksma inequality $\|S_{r_n q_n}(\alpha,\cdot)\|_{L^\infty} \le M\mathsf{V}(f)$. So for some $\hat\varepsilon > 0$ and for all n large enough, $\operatorname{mes}\{x : |S_{r_n q_n}(\alpha, x)| > \hat\varepsilon\} \ge \hat\varepsilon$.

Looking at the inequality $|S_{r_n q_n}(\alpha, x)| \le \sum_{k=0}^{r_n-1} |S_{q_n}(\alpha, x + kq_n\alpha)|$, we see that if $|S_{r_n q_n}(\alpha, x)| \ge \hat\varepsilon$, then $|S_{q_n}(\alpha, x + kq_n\alpha)| \ge \hat\varepsilon/M$ for some $0 \le k \le M-1$. So for all n large enough, $\operatorname{mes}\{x : |S_{q_n}(\alpha, x)| > \hat\varepsilon/M\} \ge \hat\varepsilon/M$. $\square$

4.2. **Proof of Theorem 4.1.** Let $\Omega^*(\alpha)$ be the set of x where the conclusion of Theorem 4.1 holds. $\Omega^*(\alpha)$ is R_α-invariant and it is measurable by Lemma A.1 in the appendix. Therefore to show that $\Omega^*(\alpha)$ has full measure, it suffices to show that it has positive measure.

Suppose $\alpha \in \mathcal{A}$ and let $n_k \uparrow \infty$ be a sequence satisfying (3.1) with $\mathcal{N}$ given by (4.2). There is no loss of generality in assuming that

$$\frac{a_{n_k+1}}{a_0 + \cdots + a_{n_k}} > k^3.$$

So $q_{n_k+1} > k^3 L_k q_{n_k}$, where $L_k := a_0 + \cdots + a_{n_k}$.

Recall that $\mu_{n_k}(x) = S_{q_{n_k}}(\alpha, x)$. For all k sufficiently large, there is a set A_k of measure at least $\varepsilon_2/2$ such that for all $x \in A_k$,

$$S_{\ell q_{n_k}}(\alpha, x) = \ell\left(\mu_{n_k}(x) \pm \frac{C_1 \ell}{k^3 L_k}\right) \text{ for all } \ell = 0, 1, \ldots, kL_k \tag{4.4}$$

$$|\mu_{n_k}(x)| \geq \varepsilon_1. \tag{4.5}$$

This is because Lemma 4.5 says that the total measure of x for which (4.4) fails is $O(1/k^2)$ while (4.5) holds on the set of measure ε_2 by Lemma 4.7.

It follows that $\mathrm{mes}(\bigcap_{n>1} \bigcup_{k>n} A_k) \geq \varepsilon_2/2$. Therefore there exists x which belongs to infinitely many A_k. After re-indexing n_k, we may assume that (4.4), (4.5) are satisfied for all $k \in \mathbb{N}$. Henceforth, we fix such an x and work with this x. Let

$$N_k(x) := kL_k q_{n_k}, \quad B_k(x) := kL_k|\mu_{n_k}(x)|, \quad A_k(x) := \frac{1}{2}(\mathrm{sgn}(\mu_{n_k}(x)) - 1)B_k.$$

Any $n \leq N_k$ can be written uniquely in the form

$$n = l(n)q_{n_k} + r(n) \quad \text{with} \quad 0 \leq l(n) \leq kL_k \text{ and } 0 \leq r(n) < q_{n_k}.$$

It is easy to see that $\frac{l(n)}{kL_k} \xrightarrow[k\to\infty]{\text{dist}} \mathrm{U}[0,1]$ as $n \sim \mathrm{U}(1, \ldots, N_k)$.

Writing $S_n(\alpha, x) = S_{l(n)q_{n_k}}(\alpha, x) + S_{r(n)}(\alpha, x + \alpha l(n) q_{n_k})$ we obtain from (4.4) and Lemma 4.6 that

$$S_n(\alpha, x) = l(n)\mu_{n_k}(x) + O(L_k).$$

So $\frac{S_n(x)}{B_k}$ is asymptotically uniform on $[0,1]$ when $\mu_{n_k} > 0$, and $[-1,0]$ when $\mu_{n_k} < 0$. So $\frac{S_n(x) - A_k}{B_k} \xrightarrow[k\to\infty]{} \mathrm{U}[0,1]$, as $n \sim \mathrm{U}(1, \ldots, N_k(x))$. $\square$

4.3. **Proof of Theorem 4.2.** Let $\Omega(\alpha)$ denote the set of $x \in \mathbb{T} := \mathbb{R}/\mathbb{Z}$ for which there are $B_N(x) \to \infty$ and $A_N(x) \in \mathbb{R}$ s.t.

$$\frac{S_n(\alpha, x) - A_N(x)}{B_N(x)} \xrightarrow[N\to\infty]{\text{dist}} \mathrm{U}[0,1], \text{ as } n \sim \mathrm{U}(1, \ldots, N). \tag{4.6}$$

$\Omega(\alpha)$ is measurable, and $A_n(\cdot), B_n(\cdot)$ can be chosen to be measurable on $\Omega(\alpha)$, see the appendix. Assume by way of contradiction that $\mathrm{mes}[\Omega(\alpha)] \neq 0$ for some $\alpha \in \mathcal{A}$.

$\Omega(\alpha)$ is invariant under $R_\alpha(x) = x + \alpha \bmod 1$ on $\mathbb{T} := \mathbb{R}/\mathbb{Z}$. Since R_α is ergodic, and $\Omega(\alpha)$ is measurable, $\mathrm{mes}[\Omega(\alpha)] = 1$.

Since $\alpha \in \mathcal{A}$, there is an increasing sequence n_k satisfying (3.1) where $\mathcal{N}$ is given by (4.2). We can choose n_k so that $q_{n_k} r_{n_k} \in \mathcal{N}$ for $r_{n_k} \leq M$, and $a_{n_k+1} > k^3 L_k$ where $L_k := a_0 + \cdots + a_{n_k}$. In particular, $q_{n_k+1} > k^3 L_k q_{n_k}$.

Recall that $\mu_{n_k}(x) := S_{q_{n_k}}(\alpha, x)$. By Lemma 4.7 we can choose x such that for infinitely many k, $|\mu_{n_k}(x)| \geq \varepsilon_1$. We will suppose that $\mu_{n_k}(x) > 0$ for infinitely many k; the case where $\mu_{n_k}(x) < 0$ for infinitely many k is similar.

Claim 1. It is possible to assume without loss of generality that $\|B_{q_{n_k}}\|_\infty := \sup_{x\in\Omega(\alpha)} |B_{q_{n_k}}(x)| \leq 3C_3 L_k$ for all k where C_3 is the constant from Lemma 4.6.

Proof. We claim that for every x with (4.6), $B_{q_{n_k}}(x) \leq 3C_3 L_k$ for all k large enough. Otherwise, by Lemma 4.6, there are infinitely many k s.t. $B_{q_{n_k}}(x) > 3\max\{|S_r(\alpha, x)| : r = 0, \ldots, q_{n_k - 1}\}$, whence $|S_n(\alpha, x)/B_{q_{n_k}}| \leq \frac{1}{3}$ for all $0 \leq n \leq q_{n_k} - 1$. In such circumstances, (4.6) does not hold (the spread is not big enough).

Since $B_{q_{n_k}}(x) \leq 3C_3 L_k$ for all k large enough, there is no harm in replacing $B_{q_{n_k}}(x)$ in (4.6) by $\min\{B_{q_{n_k}}(x), 3C_3 L_k\}$.

Claim 2. Fix $D > C = |\sum b_m|$, and let E_k denote the set of $x \in \Omega(\alpha)$ s.t. $S_r(\alpha, x) = S_r(\alpha, R_\alpha^{\ell q_{n_k}}(x)) \pm \frac{D\ell}{q_{n_k+1}}$ for all $0 \leq \ell \leq B_{q_{n_k}}(x), 0 \leq r < q_{n_k} - 1$. Then $\mathrm{mes}(E_k^c) \leq C_4 k^{-3}$.

Proof. If $x \notin E_k$, then there are $0 \leq \ell \leq B_{q_{n_k}}(x), 0 \leq r < q_{n_k} - 1$ s.t.

$$|S_r(\alpha, x) - S_r(\alpha, x + \ell q_{n_k}\alpha)| \geq \frac{D\ell}{q_{n_k+1}}.$$

By Lemma 4.4, $\{x\}$, $\{x + \ell q_{n_k}\alpha\}$ are separated by a singularity of $S_r(\alpha, \cdot)$. So x belongs to a ball of radius $2\|\ell q_{n_k}\alpha\|$ centered at one of the dq_{n_k} discontinuities of $S_{q_{n_k}}(\alpha, \cdots)$. Thus $\mathrm{mes}(E_k^c) \leq dq_{n_k} \cdot 2\|\ell q_{n_k}\alpha\|$. Now $\|\ell q_{n_k}\alpha\| \leq \ell\|q_{n_k}\alpha\| \leq \frac{\|B_{q_{n_k}}\|_\infty}{q_{n_k+1}} \leq \frac{3C_3 L_k}{q_{n_k+1}} \leq \frac{3C_3}{k^3 q_{n_k}}$ by our choice of n_k. So $\mathrm{mes}(E_k^c) \leq C_4/k^3$ with $C_4 := 6dC_3$.

Claim 3. Let F_k denote the set of $x \in \Omega(\alpha)$ s.t.

$$S_{\ell q_{n_k}}(\alpha, x) = \ell\left(\mu_{n_k}(x) \pm \frac{C_1 \ell}{k^3 L_k}\right) \text{ for all } 0 \leq \ell \leq B_{q_{n_k}}(x).$$

Then $\mathrm{mes}(F_k^c) \le C_5 k^{-2}$.

Proof. This follows from Lemma 4.5.

By Claims 2 and 3, and a Borel–Cantelli argument, for a.e. x there is $k_0(x)$ s.t. $x \in E_k \cap F_k$ for all $k \ge k_0(x)$.

Suppose $k \ge k_0(x)$, and let $N_k := q_{n_k} B_{q_{n_k}}(x)$. Every $0 \le n \le N_k - 1$ can be uniquely represented as $n = \ell q_{n_k} + r$ with $0 \le \ell \le B_{q_{n_k}}(x) - 1$ and $0 \le r \le q_{n_k} - 1$. Using the bound $\|B_{q_{n_k}}\|_\infty = O(L_k)$, we find:

$$\frac{S_n(\alpha,x) - A_{q_{n_k}}(x)}{B_{q_{n_k}}(x)} = \frac{S_{\ell q_{n_k}}(\alpha,x)}{B_{q_{n_k}}(x)} + \frac{S_r(\alpha, R_\alpha^{\ell q_{n_k}} x) - A_{q_{n_k}}(x)}{B_{q_{n_k}}(x)}$$
$$= \frac{S_{\ell q_{n_k}}(\alpha,x)}{B_{q_{n_k}}(x)} + \frac{S_r(\alpha,x) - A_{q_{n_k}}(x) + o(1)}{B_{q_{n_k}}(x)}, \text{ because } x \in E_k$$
$$= \frac{\ell(\mu_{n_k}(x) + o(1))}{B_{q_{n_k}}(x)} + \frac{S_r(\alpha,x) - A_{q_{n_k}}(x) + o(1)}{B_{q_{n_k}}(x)}, \text{ because } x \in F_k.$$

If $n \sim \mathrm{U}(0,\dots,N_k-1)$, then ℓ, r are independent random variables, $\ell \sim \mathrm{U}(0,\dots,B_{q_{n_k}}(x)-1)$ and $r \sim \mathrm{U}(0,\dots,q_{n_k}-1)$. Thus the distribution of $\frac{\ell(\mu_{n_k}(x)+o(1))}{B_{q_{n_k}}(x)}$ is close to $\mathrm{U}[0,\mu_{n_k}(x)]$, and the distribution of $\frac{S_r(\alpha,x)-A_{q_{n_k}}(x)+o(1)}{B_{q_{n_k}}(x)}$ converges to $\mathrm{U}[0,1]$ (because $x \in \Omega$).

Taking a subsequence such that $\mu_{n_k}(x) \to \bar{\varepsilon} > \varepsilon_1$ we see that the random variables $\dfrac{S_n(\alpha,x) - A_{q_{n_k}}(x)}{B_{q_{n_k}}(x)}$, where $n \sim \mathrm{U}(0,\dots,n_k-1)$, converge in distribution to the sum of two independent uniformly distributed random variables. This contradicts to (4.6), because the sum of two independent uniform random variables is not uniform. ☐

Appendix A. Measurability concerns

Let $\Omega(\alpha)$ denote the set of $x \in \mathbb{T} := \mathbb{R}/\mathbb{Z}$ such that for some $B_N(x) \to \infty$ and $A_N(x) \in \mathbb{R}$, $\frac{S_n(\alpha,x)-A_N(x)}{B_N(x)} \xrightarrow[N\to\infty]{\text{dist}} U[0,1]$, as $n \sim \mathrm{U}(1,\dots,N)$ and let $\Omega^*(\alpha)$ denote the set of $x \in \mathbb{T}$ such that along a subsequence $N_k(x)$ there exist some $B_{N_k}(x) \to \infty$ and $A_{N_k}(x) \in \mathbb{R}$, $\frac{S_n(\alpha,x)-A_{N_k}(x)}{B_{N_k}(x)} \xrightarrow[k\to\infty]{\text{dist}} \mathrm{U}[0,1]$, as $n \sim \mathrm{U}(1,\dots,N_k)$. We make no assumptions on the measurability of A_N, B_N, N_k as functions of x. The purpose of this section is to prove:

Lemma A.1. *$\Omega(\alpha)$ and $\Omega^*(\alpha)$ are measurable.*

The crux of the argument is to show that $A_N(x), B_N(x)$ can be replaced by measurable functions, defined in terms of the percentiles of the random quantities $S_n(x,\alpha)$, $n \sim \mathrm{U}(1,\dots,N)$.

Recall that given $0<t<1$, the *upper and lower t-percentiles* of a random variable X are defined by

$$\begin{aligned}\chi^+(X,t) &:= \inf\{\xi : \Pr(X \le \xi) > t\}\\ \chi^-(X,t) &:= \sup\{\xi : \Pr(X \le \xi) < t\}\end{aligned} \qquad (0<t<1).$$

Notice that $\Pr(X \le \chi^+(X,t)) \ge t$, $\Pr(X < \chi^-(X,t)) \le t$, and $\Pr(\chi^-(X,t) < X < \chi^+(X,t)) = 0$. In case X is *non-atomic* (i.e. $P(X=a)=0$ for all a), we can say more:

Lemma A.2. *Suppose X is a non-atomic real valued random variable, fix $0<t<1$ and let $\chi_t^\pm := \chi^\pm(X,t)$, then*
(a) $\Pr(X < \chi_t^+) = t$ and $\Pr(X < \chi_t^-) = t$;
(b) $\forall \varepsilon > 0$, $\Pr(\chi_t^- - \varepsilon < X < \chi_t^-), \Pr(\chi_t^+ < X < \chi_t^+ + \varepsilon)$ are positive;
(c) $\exists t_1 < t_2$ s.t. $\chi_{t_1}^- < \chi_{t_2}^+$ and $\chi_{t_1}^-, \chi_{t_2}^+$ have the same sign.

Proof. Since X is non-atomic, $\Pr(X < \chi_t^+) = \Pr(X \le \chi_t^+) \ge t$ and $\Pr(X < \chi_t^-) \le t$. If $\chi_t^+ = \chi_t^-$, part (a) holds. If $\chi_t^+ > \chi_t^-$ then for all $h>0$ small enough $\chi_t^- + h < \chi_t^+ - h$ whence

$$\begin{aligned}0 \le \Pr(\chi_t^- < X < \chi_t^+) &= \lim_{h\to 0^+} \Pr(\chi_t^- + h < X < \chi_t^+ - h)\\ &= \lim_{h\to 0^+} \Pr(X < \chi_t^+ - h) - \lim_{h\to 0^+} \Pr(X \le \chi_t^- + h) \le t - t = 0.\end{aligned}$$

Necessarily $\lim_{h\to 0^+} \Pr(X < \chi_t^+ - h) = t$ and $\lim_{h\to 0^+} \Pr(X \le \chi_t^- + h) = t$, which gives us $\Pr(X < \chi_t^+) = t$ and $\Pr(X < \chi_t^+) = \Pr(X \le \chi_t^+) = t$.

For (b) assume by contradiction that $\Pr(\chi_t^- - \varepsilon < X < \chi_t^-) = 0$, then for all $\chi_t^- - \varepsilon < \xi < \chi_t^-$, $\Pr(X \le \xi) = \Pr(X < \chi_t^-) = t$, whence $\chi_t^- \le \chi_t^- - \varepsilon$, a contradiction. Similarly, $\Pr(\chi_t^+ < X < \chi_t^+ + \varepsilon) = 0$ is impossible.

To prove (c) note that since X is non-atomic, either $\Pr(X>0)$ or $\Pr(X<0)$ is positive. Assume w.l.o.g. that $\Pr(X>0) \ne 0$. By non-atomicity, there are positive $a<b$ s.t. $\Pr(X \in (0,a)) \ne 0$ and $\Pr(X \in (a,b)) \ne 0$. Take $t_1 := \Pr(X<a)$ and $t_2 := \Pr(X<b)$. $\square$

From now on fix a non-atomic random variable Y, and choose $0 < t_1 < t_2 < 1$ as in Lemma A.2(c) s.t. $\chi^-(Y,t_1) < \chi^+(Y,t_2)$ and $\mathrm{sgn}(\chi^-(Y,t_1)) = \mathrm{sgn}(\chi^+(Y,t_2))$.

Lemma A.3. *Let S_N be (possibly atomic) random variables s.t. for some $A_N \in \mathbb{R}$ and $B_N \to \infty$, $\frac{S_N - A_N}{B_N} \xrightarrow[N\to\infty]{dist} Y$. Then $\frac{S_N - A_N^*}{B_N^*} \xrightarrow[N\to\infty]{dist} Y$,* *where A_N^*, B_N^* are the unique solution to*

$$\begin{cases} A_N^* + B_N^*\chi^-(Y,t_1) = \chi^-(S_N,t_1) \\ A_N^* + B_N^*\chi^+(Y,t_2) = \chi^+(S_N,t_2). \end{cases} \tag{A.1}$$

Proof. Without loss of generality, $\chi^-(Y,t_1), \chi^+(Y,t_2)$ are both positive.

We need the following fact (which is not automatic since S_N are allowed to be atomic):

$$\lim_{N\to\infty} \Pr(S_N < \chi^-(S_N,t)) = t \text{ for all } 0 < t < 1. \tag{A.2}$$

Indeed, given $\varepsilon > 0$, let $\xi_N := B_N\chi^-(Y,t-\varepsilon) + A_N$, then

$$\Pr(S_N < \xi_N) = \Pr\left(\tfrac{S_N - A_N}{B_N} < \chi^-(Y,t-\varepsilon)\right) \xrightarrow[N\to\infty]{} \Pr(Y < \chi^-(Y,t-\varepsilon)) = t-\varepsilon,$$

by Lemma A.2(a). So for all N large enough, $\xi_N \le \chi^-(S_N,t)$, whence $\liminf \Pr(S_N < \chi^-(S_N,t)) \ge \lim \Pr(S_N < \xi_N) = t - \varepsilon$. Since ε is arbitrary, $\liminf \Pr(S_N < \chi^-(S_N,t)) \ge t$. The other inequality $\limsup \Pr(S_N < \chi^-(S_N,t)) \le t$ is clear since $\Pr(S_N < \chi^-(S_N,t)) \le t$ for all N.

With (A.2) proved, we proceed to prove that

$$\frac{A_N^* - A_N}{B_N} \xrightarrow[N\to\infty]{} 0 \quad \text{and} \quad \frac{B_N^*}{B_N} \xrightarrow[N\to\infty]{} 1. \tag{A.3}$$

It will then be obvious that $\frac{S_N - A_N}{B_N} \xrightarrow[N\to\infty]{\text{dist}} Y$ implies $\frac{S_N - A_N^*}{B_N^*} \xrightarrow[N\to\infty]{\text{dist}} Y$.

Define two affine transformations, $\varphi_N(t) = \frac{t - A_N}{B_N}$ and $\varphi_N^*(t) = \frac{t - A_N^*}{B_N^*}$. Notice that $(\varphi_N^*)^{-1}(t) = A_N^* + B_N^* t$, so $(\varphi_N^*)^{-1}(\chi^-(Y,t_1)) = \chi^-(S_N,t_1)$, by (A.1). Since $B_N^* = \frac{\chi^+(S_N,t_2) - \chi^-(S_N,t_1)}{\chi^+(Y,t_2) - \chi^-(Y,t_1)} > 0$, φ_N^* is increasing. By (A.2), $\Pr(\varphi_N^*(S_N) < \chi^-(Y,t_1)) = \Pr(S_N < \chi^-(S_N,t_1)) \xrightarrow[N\to\infty]{} t_1$. So

$$\begin{aligned} t_1 &= \lim_{N\to\infty} \Pr(\varphi_N^*(S_N) < \chi^-(Y,t_1)) \\ &= \lim_{N\to\infty} \Pr\big(\varphi_N(S_N) < \varphi_N[(\varphi_N^*)^{-1}(\chi^-(Y,t_1))]\big) \\ &= \lim_{N\to\infty} \Pr\left(\varphi_N(S_N) < \frac{B_N^*}{B_N}\left(\chi^-(Y,t_1) + \frac{A_N^* - A_N}{B_N}\right)\right). \end{aligned}$$

We claim that this implies that

$$\liminf_{N\to\infty} \frac{B_N^*}{B_N}\left(\chi^-(Y,t_1) + \frac{A_N^* - A_N}{B_N}\right) \ge \chi^-(Y,t_1). \tag{A.4}$$

Otherwise, $\exists\varepsilon$ s.t. $\liminf\limits_{N\to\infty} \frac{B_N^*}{B_N}\left(\chi^-(Y,t_1)+\frac{A_N^*-A_N}{B_N}\right) < \chi^-(Y,t_1)-\varepsilon$, so

$$\begin{aligned}
t_1 &= \liminf_{N\to\infty} \Pr\left(\varphi_N(S_N) < \frac{B_N^*}{B_N}\left(\chi^-(Y,t_1)+\frac{A_N^*-A_N}{B_N}\right)\right)\\
&\le \liminf_{N\to\infty} \Pr\left(\varphi_N(S_N) < \chi^-(Y,t_1)-\varepsilon\right) = \Pr(Y<\chi^-(Y,t_1)-\varepsilon)\\
&= \Pr(Y<\chi^-(Y,t_1)) - \Pr(\chi^-(Y,t_1)-\varepsilon \le Y < t_1)\\
&< t_1, \text{ by Lemma A.2, parts (a),(b).}
\end{aligned}$$

Similarly, one shows that

$$\limsup_{N\to\infty} \frac{B_N^*}{B_N}\left(\chi^+(Y,t_2)+\frac{A_N^*-A_N}{B_N}\right) \le \chi^+(Y,t_2). \tag{A.5}$$

It remains to see that (A.4) and (A.5) imply (A.3). First we divide (A.4) by (A.5) to obtain

$$\limsup_{N\to\infty} \frac{\chi^-(Y,t_1)+\frac{A_N^*-A_N}{B_N}}{\chi^+(Y,t_2)+\frac{A_N^*-A_N}{B_N}} \ge \frac{\chi^-(Y,t_1)}{\chi^+(Y,t_2)}.$$

Since $x\mapsto \frac{a+x}{b+x}$ is strictly decreasing on $[0,\infty)$ when $a>b>0$, this implies that

$$\limsup_{N\to\infty} \frac{A_N^*-A_N}{B_N} \le 0. \tag{A.6}$$

Looking at (A.4), and recalling that $\chi^-(Y,t_1)>0$, we deduce that

$$\liminf_{N\to\infty} \frac{B_N^*}{B_N} \ge 1. \tag{A.7}$$

Next we look at the difference of (A.4) and (A.5) and obtain

$$\limsup_{N\to\infty} \frac{B_N^*}{B_N}\big(\chi^+(Y,t_2)-\chi^-(Y,t_1)\big) \le \chi^+(Y,t_2)-\chi^-(Y,t_1),$$

whence $\limsup(B_N^*/B_N)\le 1$. Together with (A.7), this proves that $B_N^*/B_N \xrightarrow[N\to\infty]{} 1$. Substituting this in (A.4), gives $\liminf \frac{A_N^*-A_N}{B_N}\ge 0$, which, in view of (A.6), implies that $\frac{A_N^*-A_N}{B_N} \xrightarrow[N\to\infty]{} 0$. This completes the proof of (A.3), and with it, the lemma. □

Proof of Lemma A.1. We begin with the measurability of $\Omega(\alpha)$.

Let $S_N(x)$ denote the random variable equal to $S_n(\alpha,x)$ with probability $\frac{1}{N}$ for each $1\le n\le N$.

We will apply Lemma A.3 with $Y := \mathrm{U}[0,1]$, $S_N = S_n(x)$ and (say) $t_1 := \frac{1}{3}$, $t_2 := \frac{2}{3}$. It says that

$$\Omega(\alpha) = \{x \in \mathbb{T} : \tfrac{S_N(x)-A_N^*(x)}{B_N^*(x)} \xrightarrow[N\to\infty]{\text{dist}} U[0,1]\},$$

where $A_N^*(x)$ and $B_N^*(x)$ are the unique solutions to (A.1). Since the percentiles of $S_N(x)$ are measurable as functions of x, $A_N^*(x), B_N^*(x)$ are measurable as functions of x.

We claim that $\Omega(\alpha) = \Omega_1(\alpha) \cap \Omega_2(\alpha)$ where

$$\Omega_1(\alpha) := \bigcap_{\ell=1}^{\infty} \bigcup_{M=1}^{\infty} \bigcap_{N=M+1}^{\infty} \left\{ x \in \mathbb{T} : \frac{1}{N} \sum_{n=1}^{N} 1_{(2,\infty)} \left(\left| \tfrac{S_n(\alpha,x)-A_N^*(x)}{B_N^*(x)} \right| \right) < \frac{1}{\ell} \right\}$$

$$\Omega_2(\alpha) = \bigcap_{t\in\mathbb{Q}\setminus\{0\}} \left\{ x \in \mathbb{T} : \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{it\left(\frac{S_n(\alpha,x)-A_N^*}{B_N^*(x)}\right)} = \mathbb{E}\left(e^{itY}\right) \right\}.$$

This will prove the lemma, since the measurability of $A_N^*(\cdot), B_N^*(\cdot)$ implies the measurability of $\Omega_i(\alpha)$.

If $x \in \Omega(\alpha)$ then $x \in \Omega_1(\alpha)$ because $\mathbb{P}[|\frac{S_N(x)-A_N^*}{B_N^*}| > 2] \xrightarrow[N\to\infty]{} 0$, and $x \in \Omega_2(\alpha_2)$ because $\mathbb{E}\big(e^{it\left(\frac{S_N(x)-A_N^*}{B_N^*}\right)}\big) \xrightarrow[N\to\infty]{} \mathbb{E}\left(e^{itY}\right)$ pointwise.

Conversely, if $x \in \Omega_1(\alpha) \cap \Omega_2(\alpha)$ then it is not difficult to see that $\mathbb{E}\big(e^{it\left(\frac{S_N(x)-A_N^*}{B_N^*}\right)}\big) \xrightarrow[N\to\infty]{} \mathbb{E}\left(e^{itY}\right)$ for *all* $t \in \mathbb{R}$. So $x \in \Omega(\alpha)$ by Lévy's continuity theorem. Thus $\Omega(\alpha) = \Omega_1(\alpha) \cap \Omega_2(\alpha)$, whence $\Omega(\alpha)$ is measurable.

The proof that $\Omega^*(\alpha)$ is measurable is similar. Enumerate $\mathbb{Q}\setminus\{0\} = \{t_n : n \in \mathbb{N}\}$, then $\alpha \in \Omega^*(\alpha)$ iff for every $\ell \in \mathbb{N}$ there exist $M \in \mathbb{N}$ s.t. for some $N > M$

(1) $\frac{1}{N}\sum_{n=1}^{N} 1_{(2,\infty)} \left(\left| \frac{S_n(\alpha,x)-A_N^*(x)}{B_N^*(x)} \right| \right) < \frac{1}{\ell}$

(2) $|\mathbb{E}\big(e^{it_n\left(\frac{S_N(x)-A_N^*(x)}{B_N^*(x)}\right)}\big) - \mathbb{E}\left(e^{it_nY}\right)| < \frac{1}{\ell}$ for all $n = 1, \ldots, \ell$

These are measurable conditions, because $A_N^*(\cdot), B_N^*(\cdot)$ are measurable. So $\Omega^*(\alpha)$ is measurable. □

References

[ADDS15] Artur Avila, Dmitry Dolgopyat, Eduard Duryev, and Omri Sarig. The visits to zero of a random walk driven by an irrational rotation. *Israel J. Math.*, 207(2):653–717, 2015.

[ADU93] Jon Aaronson, Manfred Denker, and Mariusz Urbański. Ergodic theory for Markov fibred systems and parabolic rational maps. *Trans. Amer. Math. Soc.*, 337(2):495–548, 1993.

[AK82] Jon Aaronson and Michael Keane. The visits to zero of some deterministic random walks. *Proc. London Math. Soc. (3)*, 44(3):535–553, 1982.

[Bec94] József Beck. Probabilistic Diophantine approximation. I. Kronecker sequences. *Ann. of Math. (2)*, 140(1):109–160, 1994.

[Bec10] József Beck. Randomness of the square root of 2 and the giant leap, Part 1. *Period. Math. Hungar.*, 60(2):137–242, 2010.

[Bec11] József Beck. Randomness of the square root of 2 and the giant leap, Part 2. *Period. Math. Hungar.*, 62(2):127–246, 2011.

[BU] Michael Bromberg and Corinna Ulcigrai. A temporal central limit theorem for real-valued cocycles over rotations. preprint ArXiv:1705.06484.

[CK76] Jean-Pierre Conze and Michael Keane. Ergodicité d'un flot cylindrique. page 7, 1976.

[DSa] Dmitry Dolgopyat and Omri Sarig. Asymptotic windings of horocycles. *to appear in Israel Math. J.*

[DSb] Dmitry Dolgopyat and Omri Sarig. Quenched and annealed temporal limit theorems for circle rotations. Preprint (2017).

[DS17] Dmitry Dolgopyat and Omri Sarig. Temporal distributional limit theorems for dynamical systems. *J. Statistical Physics*, 166(3-4):680–713, 2017.

[DV86] Harold G. Diamond and Jeffrey D. Vaaler. Estimates for partial sums of continued fraction partial quotients. *Pacific J. Math.*, 122(1):73–82, 1986.

[HW08] Godfrey Harold Hardy and Edward Maitland Wright. *An introduction to the theory of numbers*. Oxford University Press, Oxford, sixth edition, 2008. Revised by D. R. Heath-Brown and J. H. Silverman, With a foreword by Andrew Wiles.

[Jar29] Vojtech Jarník. Zur metrischen theorie der diophantischen approximationen. *Prace Matematyczno-Fizyczne*, 36(1):91–106, 1928-1929.

[Khi24] Alexander Ya. Khintchine. Einige Sätze über Kettenbrüche, mit Anwendungen auf die Theorie der Diophantischen Approximationen. *Math. Ann.*, 92(1-2):115–125, 1924.

[Khi63] Alexander Ya. Khintchine. *Continued fractions*. Translated by Peter Wynn. P. Noordhoff, Ltd., Groningen, 1963.

[PS] Elliot Paquette and Younghwan Son. Birkhoff sum fluctuations in substitution dynamical systems. *arXiv:1505.01428, preprint (2015)*.

[Sch78] Klaus Schmidt. A cylinder flow arising from irregularity of distribution. *Compositio Math.*, 36(3):225–232, 1978.

[Sul82] Dennis Sullivan. Disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics. *Acta Math.*, 149(3-4):215–237, 1982.

Department of Mathematics, University of Maryland at College Park, College Park, MD 20740, USA
E-mail address: dmitry@math.umd.edu

Faculty of Mathematics and Computer Science, The Weizmann Institute of Science, POB 26, Rehovot, Israel
E-mail address: omsarig@gmail.com