QUENCHED AND ANNEALED TEMPORAL LIMIT THEOREMS FOR CIRCLE ROTATIONS.

DMITRY DOLGOPYAT AND OMRI SARIG

In memory of Jean-Christophe Yoccoz

ABSTRACT. Let $h(x) = \{x\} - \frac{1}{2}$. We study the distribution of $\sum_{k=0}^{n-1} h(x + k\alpha)$ when x is fixed, and n is sampled randomly uniformly in $\{1, \ldots, N\}$, as $N \to \infty$. Beck proved in [Bec10, Bec11] that if $x = 0$ and α is a quadratic irrational, then these distributions converge, after proper scaling, to the Gaussian distribution. We show that the set of α where a distributional scaling limit exists has Lebesgue measure zero, but that the following *annealed* limit theorem holds: Let (α, n) be chosen randomly uniformly in $\mathbb{R}/\mathbb{Z} \times \{1, \ldots, N\}$, then the distribution of $\sum_{k=0}^{n-1} h(k\alpha)$ converges after proper scaling as $N \to \infty$ to the Cauchy distribution.

1. INTRODUCTION

We study the centered ergodic sums of functions $h : \mathbb{T} \to \mathbb{R}$ for the rotation by a an irrational angle α

$$S_n(\alpha, x) = \left(\sum_{k=1}^{n} h(x + k\alpha) \right) - n \int_{\mathbb{T}} h(z) dz. \tag{1.1}$$

Weyl's equidistribution theorem says that for every $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, and for every h Riemann integrable, $\frac{1}{n} S_n(\alpha, x) \xrightarrow[n \to \infty]{} 0$ uniformly in x. We are interested in higher-order asymptotics. We aim at results which hold for a set of full Lebesgue measure of α.

If h is sufficiently smooth, then $S_n(\alpha, x)$ is bounded for almost every α and all x (see [Her79] or appendix A). The situation for piecewise smooth h is more complicated, and not completely undertstood even for functions with a single singularity.

2010 *Mathematics Subject Classification.* 37D25 (primary), 37D35 (secondary).
D. D. acknowledges the support of the NSF grant DMS1665046. O.S. acknowledges the support of the ISF grant 199/14.

Setup. Here we study (1.1), for the simplest example of a piecewise smooth function with one discontinuity on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$:

$$h(x) = \{x\} - \frac{1}{2}. \tag{1.2}$$

The *fractional part* $\{x\}$ is the unique $t \in [0, 1)$ s.t. $x \in t + \mathbb{Z}$.

Case (1.2) is sufficient for understanding the behavior for typical α for all functions $f(t)$ on $\mathbb{T}$ which are differentiable everywhere except one point x_0, and whose derivative on $\mathbb{T} \setminus \{x_0\}$ extends to a function of bounded variation on $\mathbb{T}$. This is because of the following result proven in Appendix A:

Proposition 1.1. *If $f(t)$ is differentiable on $\mathbb{T} \setminus \{x_1, \ldots, x_\nu\}$ and f' extends to a function with bounded variation on $\mathbb{T}$, then there are $A_1, \ldots, A_\nu \in \mathbb{R}$ s.t. for a.e. α there is $\varphi_\alpha \in C(\mathbb{T})$ s.t. for all $x \neq x_i$,*

$$f(x) = \sum_{i=1}^{\nu} A_i h(x + x_i) + \int_{\mathbb{T}} f(t)dt + \varphi_\alpha(x) - \varphi_\alpha(x + \alpha).$$

Of course there are many functions h for which Proposition 1.1 holds. The choice (1.2) is convenient, because of its nice Fourier series.

Methodology. $S_n(\alpha, x)$ is very oscillatory. Therefore, instead of looking for simple asymptotic formulas for $S_n(\alpha, x)$, which is hopeless, we will look for simple *scaling limits for the distribution* of $S_n(\alpha, x)$ when x, or α, or n (or some of their combinations) are randomized. There are several natural ways to carry out the randomization:

(1) *Spatial* vs *temporal* limit theorems: In a spatial limit theorem, the initial condition x chosen randomly from the space $\mathbb{T}$. In a temporal limit theorem, the initial condition x is fixed, and the "time" n is chosen randomly uniformly in $\{1, \ldots, N\}$ as $N \to \infty$. Neither limit theorem implies the other, see [DS17].

(2) *Quenched* vs *annealed* limit theorems: In a quenched limit theorem, α is fixed. In an annealed limit theorem α is randomized. The terminology is motivated by the theory of random walks in random environment; the parameter α is the "environment parameter."

We indicate what is known and what is still open in our case.

Known results on spatial limit theorems: The quenched spatial limit theorem fails; the annealed spatial limit theorem holds.

The failure of the *quenched* spatial limit theorem is very general. Namely, It follows from the Denjoy-Koksma inequality that there are no quenched spatial distributional limit theorems for any rotation by $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, and every function of bounded variation which is not a coboundary (e.g. $h(x) = \{x\} - \frac{1}{2}$). In the coboundary case, the spatial limit theorem is trivial. Many people have looked for weaker quenched versions of spatial distributional limit theorem (e.g. along special subsequences of "times"). See [DF15, DS17] for references and further discussion.

The *annealed* spatial limit theorem is a famous result of Kesten.

Theorem 1.2. *(*[Kes60]*) If (x, α) is uniformly distributed on $\mathbb{T}\times\mathbb{T}$ then the distribution of $\frac{S_n(\alpha,x)}{\ln n}$ converges as $n \to \infty$ to a Cauchy distribution: $\exists \boldsymbol{\rho}_1 \neq 0$ s.t. for all $t \in \mathbb{R}$, $\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n(\alpha,x)}{\ln n} \leq t\right) = \frac{1}{2} + \frac{\arctan(t/\boldsymbol{\rho}_1)}{\pi}$.*

See [Kes60] for the value of constant $\boldsymbol{\rho}_1$. The same result holds for $h(x) = 1_{[0,\beta)}(\{x\}) - \beta$ with $\beta \in \mathbb{R}$, with different $\boldsymbol{\rho}_1 = \boldsymbol{\rho}_1(\beta)$ [Kes60, Kes62].

Known results on temporal limit theorems: Quenched temporal limit theorems are known for special α; There were no results on the annealed temporal limit theorem until this work.

The first temporal limit theorem for an irrational rotation (indeed for any dynamical system) is due to J. Beck [Bec10, Bec11]. Let

$$M_N(\alpha, x) := \frac{1}{N}\sum_{n=1}^{N} S_n(\alpha, x)\ ,\ \mathcal{S}_n(\alpha, x) := S_n(\alpha, x) - M_N(\alpha, x).$$

Theorem 1.3 (Beck)**.** *Let α be an irrational root of a quadratic polynomial with integer coefficients. Fix $x = 0$. If n is uniformly distributed on $\{1 \dots N\}$ then $\frac{\mathcal{S}_n(\alpha,x)}{\sqrt{\ln N}}$ converges to a normal distribution as $N \to \infty$.*

A similar result holds for the same x and α with $h(x) = \{x\} - \frac{1}{2}$ replaced by $1_{[0,\beta)}(\{x\}) - \beta$, $\beta \in \mathbb{Q}$ [Bec10, Bec11]. [ADDS15, DS17] extended this to all $x \in [0, 1)$. A remarkable recent paper by Bromberg & Ulcigrai [BU] gives a further extension to all x, *all irrational α of bounded type*, and for an uncountable collection of β (which depends on α). Recall that the set of α of bounded type is a set of full Hausdorff dimension [Jar29], but zero Lebesgue measure [Khi24].

This paper: We show that for $h(x) = \{x\} - \frac{1}{2}$, the quenched temporal limit theorem fails for a.e. α, but that the annealed temporal limit theorem holds. See §2 for precise statements.

Heuristic overview of the proof. When we expand the ergodic sums of h into Fourier series, we find that the resulting trigonometric series can be split into the contribution of "resonant" and "nonresonant" harmonics.

The non-resonant harmonics are many in number, but small in size. They tend to cancel out, and their total contribution is of order $\sqrt{\ln N}$. It is natural to expect that this contribution has Gaussian statistics. If α has bounded type, all harmonics are non-resonant, and as Bromberg and Ulcigrai show in the case $1_{[0,\beta)} - \beta$ the limiting distribution is indeed Gaussian.

The resonant harmonics are small in number, but much larger in size: Individual resonant harmonics have contribution of order $\ln N$. For typical α, the number, strength, and location of the resonant harmonics changes erratically with N in a non-universal way. This leads to the failure of temporal distributional limit theorems for typical α.

We remark that a similar obstruction to quenched limit theorems have been observed before in the theory of random walks in random environment [DG12, PS13, CGZ00].

To justify this heuristic we fix N and compute the distribution of resonances when α is uniformly distributed. Since the distribution of resonances is non-trivial, changing a scale typically leads to a different temporal distribution proving that there is no limit as $N \to \infty$. As a byproduct of our analysis we obtain some insight on the frequency with which a given limit distribution occurs.

Functions with more than one discontinuity. In a separate paper we use a different method to show that given a piecewise smooth discontinuous function with arbitrary finite number of discontinuities, the quenched temporal limit theorems fails for Lebesgue almost all α. But this method does not provide an annealed result, and it does not give us as detailed information as we get here on the scaling limits which appear along subsequences for typical α.

2. Statement of results

Let $S_n(\alpha, x) := \sum_{k=1}^{n} h(x + k\alpha)$, and

$$M_N(\alpha, x) = \frac{1}{N}\sum_{n=1}^{N} S_n(\alpha, x),$$
$$\mathcal{S}_n(\alpha, x) = S_n(\alpha, x) - M_N(\alpha, x).$$

Consider the cumulative distribution function of $\frac{\mathcal{S}_n(\alpha,x)}{\ln N}$:

$$\mathcal{F}_N(\alpha)(z) = \frac{1}{N}\text{Card}\left(1 \le n \le N : \frac{\mathcal{S}_n(\alpha, x)}{\ln N} \le z\right).$$

When α is random $\mathcal{F}_N(\alpha)$ becomes a random element in the space $\mathbb{X}$ of distribution functions endowed with Prokhorov topology.

We begin with the annealed temporal distributional limit theorem:

Theorem 2.1. *Let (α, n) be uniformly distributed on $\mathbb{T}\times\{1,\dots,N\}$. Then $\frac{S_n(\alpha,x)}{\ln N}$ converges in law to the Cauchy distribution as $N\to\infty$. That is, for all $z\in\mathbb{R}$*

$$\lim_{N\to\infty}\mathbb{P}\left(\frac{S_n(\alpha,x)}{\ln N}\le t\right)=\frac{1}{2}+\frac{\arctan(t/\boldsymbol{\rho}_2)}{\pi}$$

where

$$\boldsymbol{\rho}_2=\frac{1}{3\pi\sqrt{3}}. \tag{2.1}$$

Next we turn to the quenched result, beginning with some preparations. Recall (see e.g. [DF, Section 2.1]) that the Cauchy random variable can be represented up to scaling as

$$\mathfrak{C}=\sum_{m=1}^{\infty}\frac{\Theta_m}{\xi_m} \tag{2.2}$$

where $\Xi=\{\xi_m\}$ is a Poisson process on $\mathbb{R}$ and Θ_m are i.i.d bounded random variables with zero mean independent of Ξ. To make our exposition more self-contained we recall the derivation of (2.2) in Appendix B.

In our case the relevant distribution of Θs is the following. Let θ be uniformly distributed on $[0,1]$ and define

$$\Theta(\theta)=\sum_{k=1}^{\infty}\frac{\cos(2\pi k\theta)}{2\pi^2k^2}. \tag{2.3}$$

Θ has period one, and

$$\Theta(\theta)=\frac{\theta^2-\theta}{2}+\frac{1}{12} \tag{2.4}$$

on $[0,1]$, as can be verified by expanding $\frac{\theta^2-\theta}{2}+\frac{1}{12}$ on $[0,1]$ into a Fourier series. Notice that for every $\theta\in[0,1]$, $\Theta(\theta)=\frac{\zeta^2}{2}-\frac{1}{24}$, where $\zeta=\theta-\frac{1}{2}$. Thus $-\frac{1}{24}\le\Theta\le\frac{1}{12}$, and

$$\mathbb{P}(\Theta<t)=\begin{cases}0 & t\le-\frac{1}{24}\\ 2\sqrt{2t+\frac{1}{12}} & t\in(-\frac{1}{24},\frac{1}{12})\\ 1 & t>\frac{1}{12}.\end{cases} \tag{2.5}$$

Next, given a sequence $\Xi=\{\xi_m\}$ s.t. $\sum_m\xi_m^{-2}<\infty$ we can define

$$\mathfrak{C}_\Xi=\sum_m\frac{\Theta_m}{\xi_m}. \tag{2.6}$$

The sum in (2.6) converges almost surely due to Kolmogorov's Three Series Theorem (note that (2.3) easily implies that $\mathbb{E}(\Theta)=0$). Let $\mathfrak{F}_\Xi$ be the cumulative distribution function of $\mathfrak{C}_\Xi$. If Ξ is a Poisson process on $\mathbb{R}$, then $\mathfrak{F}_\Xi$ is a random element of Prokhorov's space $\mathbb{X}$.

Theorem 2.2. *If α has absolutely continuous distribution on $\mathbb{T}$ with bounded density then $\mathcal{F}_N(\alpha)$ converges in law as $N\to\infty$ to $\mathfrak{F}_\Xi$ where Ξ is the Poisson process on $\mathbb{R}$ with intensity*

$$\mathbf{c}=\frac{6}{\pi^2}. \tag{2.7}$$

A similar result has been proven for sub-diffusive random walks in random environment in [DG12] with the following distinctions:

(1) for random walks Θ_m+1 have exponential distribution rather than the distribution given by (2.3)
(2) For random walk the Poisson process in the denominator of (2.2) is supported on $\mathbb{R}^+$ and can have intensity $\widetilde{\mathbf{c}}x^{-s}$ with $s\neq 0$.

We now explain how to use Theorem 2.2 to show that for a.e. α, there is no non-trivial temporal distributional limit theorem for $S_n(\alpha,x)$. It is enough to show that for a.e. α, one can find several sequences N_k with different scaling limits for $S_n(\alpha,x)$ as $n\sim U\{1,\dots,N_k\}$. Let

$$\mathscr{D}(\Theta):=\overline{\{\text{finite linear combinations of i.i.d. with distribution }\Theta\}}$$

(closure in $\mathbb{X}$).

Corollary 2.3. *For a.e. α, for every $Y\in\mathscr{D}(\Theta)$, there are $N_k\to\infty$, $B_k\to\infty$, $A_k\in\mathbb{R}$ s.t.*

$$\frac{S_n(\alpha,x)-A_k}{B_k}\xrightarrow[k\to\infty]{dist}Y,\ \ \textit{as } n\sim U\{1,\dots,N_k\}. \tag{2.8}$$

In particular, for a.e. α, the distribution of Θ (2.5), and the normal distribution are distributional limit points of properly rescaled ergodic sums $S_n(\alpha,x)$ as $n\sim U\{1,\dots,N\}$.

Proof. Put on $\mathbb{X}$ the probability measure μ induced by the $\mathfrak{F}_\Xi$, when Ξ is the Poisson point process with intensity $\mathbf{c}$ as in Theorem 2.2. It is standard to see that for every $Y\in\mathscr{D}(\Theta)$, there is a decreasing sequence of open $\mathbb{U}_n\subset\mathbb{X}$ with positive measure s.t. $\mathbb{U}_n\downarrow\{Y\}$.

We claim that for any n, for almost every α, every sequence has a subsequence $\{N'_m\}$ such that $\mathcal{F}_{N'_m}(\alpha)\in\mathbb{U}_n$ for all m. A diagonal argument then produces a subsequence $\{N_k\}$ along which we have (2.8) with $A_k:=M_{N_k}(\alpha)$ and $B_k:=\ln N_k$.

Fix n and set $\mathbb{U} := \mathbb{U}_n$. To produce $\{N'_m\}$ it is enough to show that for each $\overline{N}$,

$$\text{mes}(\alpha \in \mathbb{T} : \exists N \geq \overline{N} \text{ such that } \mathcal{F}_N(\alpha) \in \mathbb{U}) = 1. \tag{2.9}$$

Let $\mu(\mathbb{U}) = 2\varepsilon$. Let α be uniformly distributed. By Theorem 2.2 there exists $n_1 \geq \overline{N}$ and a set $A_1 \subset \mathbb{T}$ such that $\text{mes}(A_1) \geq \varepsilon$ so that for every $\alpha \in A_1$, $\mathcal{F}_{n_1}(\alpha) \in \mathbb{U}$. If $\text{mes}(A_1) = 1$ we are done; otherwise we apply Theorem 2.2 with α uniformly distributed on $\mathbb{T} \setminus A_1$ and find $n_2 \geq n_1$ and a set $A_2 \subset \mathbb{T} \setminus A_1$ such that $\text{mes}(A_2) \geq \varepsilon \text{mes}(A_2)$ so that for each $\alpha \in A_2$, $\mathcal{F}_{n_2}(\alpha) \in \mathbb{U}$. Continuing in this way, we obtain $n_m \uparrow \infty$ such that for $\alpha \in A_j$, $\mathcal{F}_{n_j}(\alpha) \in \mathbb{U}$ and $\text{mes}(\mathbb{T} \setminus \cup_{j=1}^k A_j) \leq (1-\varepsilon)^k$. Letting k to infinity we obtain (2.9). □

Corollary 2.3 shows for a.e. α, there is no non-trivial temporal distributional limit theorem for $S_n(\alpha, x)$.

3. The main steps in the proofs of Theorems 2.1, 2.2

We state the main steps in the proofs of Theorems 2.1, 2.2. The technical work needed to carry out these steps is in the next section.

Step 1: Identifying the resonant harmonics. It is a classical fact that the Fourier series of $h(x) = \{x\} - \frac{1}{2}$ converges to $h(x)$ everywhere:

$$h(x) = -\sum_{j=1}^{\infty} \frac{\sin(2\pi j x)}{\pi j} \quad \text{for all } x \in \mathbb{R}. \tag{3.1}$$

We will use this identity to represent

$$\mathcal{S}_n(\alpha, x) := \sum_{k=1}^{n} h(x + k\alpha) - \frac{1}{N} \sum_{k=1}^{N} h(x + k\alpha)$$

as a trigonometric sum, and then work to separate the "resonant frequencies", which contribute to the asymptotic distributional behavior of $\mathcal{S}_n(\alpha)$, from those which do not. We need the following definitions:

$$\begin{aligned} T &:= N \ln^2 N, \\ g_{j,n} &:= \frac{\cos((2n+1)\pi j \alpha + 2\pi j x)}{2\pi j \sin(\pi j \alpha)}, \\ \mathcal{S}_{n,T}(\alpha) &:= \sum_{j=1}^{T} g_{j,n}. \end{aligned}$$

Proposition 3.1. *Let (α, n) be uniformly distributed on $\mathbb{T} \times \{1, \ldots, N\}$. Then $\mathcal{S}_n(\alpha, x) = \mathcal{S}_{n,T}(\alpha) + \hat{\varepsilon}_n$ where $\frac{\hat{\varepsilon}_n}{\ln N} \xrightarrow[N\to\infty]{dist} 0$.*

The proof is given in §4.2. We follow the analysis of [Bec10, Bec11], but we obtain weaker estimates since we consider a larger set of rotation numbers than in [Bec10, Bec11].

In what follows, indices j in $g_{j,n}$ are called "*harmonics*." We will separate the harmonics into different classes, according to their contribution to $\mathcal{S}_{n,T}(\alpha)$. We begin with some standard definitions:

Given $x \in \mathbb{R}$ there is a unique pair $y \in (-\frac{1}{2}, \frac{1}{2}]$, $m \in \mathbb{Z}$ such that $x = m + y$. We will call *y the signed distance from x to the nearest integer* and denote it by $((x))$. We let $\|x\| = |((x))|$, $\langle\langle x \rangle\rangle = (-1)^m((x))$ where m is as above.

Fix N. An integer $1 \le j \le T = N \ln^2 N$ is called a *prime harmonic*, if $j\alpha = ((j\alpha)) + m$ where $\gcd(j, m) = 1$. If j and m are not coprime, that is $r := \gcd(j, m) \neq 1$, then we call j/r the *prime harmonic associated to j*.

Definition 3.2. *Fix $\delta > 0$, and $N \gg 1$.*

(1) $p \in \mathbb{N}$ *is called a* prime resonant harmonic*, if $p \le N$, p is a prime harmonic, and $\|p\alpha\| \le (\delta p \ln N)^{-1}$.*
(2) $j \in \mathbb{N}$ *is called a* resonant harmonic*, if $j \le N \ln^2 N$, and the prime harmonic associated to j is a prime resonant harmonic.*

Let $\mathcal{R} = \mathcal{R}(\delta, N)$ denote the set of resonant harmonics, $\mathcal{P} = \mathcal{P}(\delta, N)$ the set of prime resonant harmonics, and $\mathcal{O} = \mathcal{O}(\delta, N)$ be the set of non resonant harmonics which are less than $T(N) = N \ln^2 N$. Split

$$\mathcal{S}_{n,T}(\alpha) = \mathcal{S}^{\mathcal{R}}_{n,T}(\alpha) + \mathcal{S}^{\mathcal{O}}_{n,T}(\alpha) \quad \text{where} \quad \mathcal{S}^{\mathcal{J}}_{n,T} = \sum_{j \in \mathcal{J}} g_{j,n}.$$

Let $\mathcal{V}^{\mathcal{O}}_N(\alpha) := \mathbb{E}_n[\mathcal{S}^{\mathcal{O}}_{n,T}(\alpha)^2] \equiv \frac{1}{N}\sum_{n=1}^N \mathcal{S}^{\mathcal{O}}_{n,T}(\alpha)^2$.

Proposition 3.3. *Suppose $\alpha \in \mathbb{T}$ is distributed according to an absolutely continuous measure with bounded density. For every $\varepsilon > 0$ there are $\delta_0 > 0$ and $E_N(\varepsilon) \subset \mathbb{T}$ Borel with the following properties:*

(1) $\mathrm{mes}(E_N(\varepsilon)) > 1 - \varepsilon$ *for all N large enough;*
(2) for all $0 < \delta < \delta_0$, $\displaystyle\lim_{N\to\infty}\left(\sup_{\alpha \in E_N(\varepsilon)} \frac{\mathcal{V}^{\mathcal{O}(N,\delta)}_N(\alpha)}{\ln^2 N}\right) \le \varepsilon$.

The proof is given in §4.2. Here is a corollary.

Corollary 3.4. *For every $\varepsilon > 0$ there is a $\delta_0 > 0$ s.t. for all N large enough and $0 < \delta < \delta_0$, $\mathbb{P}_{\alpha,n}\left(\frac{|\mathcal{S}^{\mathcal{O}}_{n,T}|}{\ln N} > \varepsilon\right) < \varepsilon$, where $1 \le n \le N$ is distributed uniformly, $\alpha \in \mathbb{T}$ is sampled from an absolutely continuous measure with bounded density, and α, n are independent.*

Proof. Without loss of generality $3\varepsilon^2 < \varepsilon$. Fubini's theorem gives

$$\mathbb{P}_{\alpha,n}\left(\frac{|\mathcal{S}^{\mathcal{O}}_{n,T}|}{\ln N} > \varepsilon\right) \le \varepsilon^4 + \int_{E_N(\varepsilon^4)} \mathbb{P}_n\left(\frac{|\mathcal{S}^{\mathcal{O}}_{n,T}|}{\ln N} > \varepsilon\right) d\alpha.$$

By Chebyshev's Inequality, the integrand is less than $\frac{1}{\varepsilon^2}\cdot\frac{1}{\ln^2 N}\mathcal{V}^{\mathcal{O}}_N$. On $E_N(\varepsilon^4)$, this is less than $2\varepsilon^2$ for all N large enough (uniformly in α). So $\mathbb{P}_{\alpha,n}(\frac{1}{\ln}|\mathcal{S}^{\mathcal{O}}_{n,T}| > \varepsilon) \le \varepsilon^4 + 2\varepsilon^2 < \varepsilon$. $\square$

Proposition 3.1 and Corollary 3.4 say that the asymptotic distributional behavior of $\mathcal{S}_n(\alpha)$ is determined by the behavior of the sum of the resonant terms $\mathcal{S}^{\mathcal{R}}_{n,T}(\alpha)$.

Step 2: An identity for the sum of resonant terms. Let

$$\xi_j := j\langle\langle j\alpha\rangle\rangle \ln N \text{ and } \Theta_j(n) := \Theta\left(\tfrac{[(2n+1)j\alpha+2jx] \bmod 2}{2}\right),$$

where $\Theta(t) = \displaystyle\sum_{k=1}^{\infty}\frac{\cos(2\pi kt)}{2\pi^2k^2}$, see (2.3).

Proposition 3.5. *For all δ small enough,*

$$\frac{\mathcal{S}^{\mathcal{R}(\delta,N)}_{n,T}}{\ln N} = \sum_{j\in\mathcal{P}(\delta,N)} \frac{\Theta_j(n) + \mathcal{O}\left(\frac{1}{\ln N}\right)}{\xi_j}. \tag{3.2}$$

The big Oh is uniform in j but not in δ.

For the proof, see §4.3.

Step 3: Limit theorems for resonant harmonics. We will use (3.2) to study of the distributional behavior of $\mathcal{S}^{\mathcal{R}}_{n,T}$.

First we will describe the distribution of the set of denominators $\{\xi_j = j\langle\langle j\alpha\rangle\rangle \ln N\}_{j\in\mathcal{P}(\delta,N)}$, and then we will describe the conditional joint distribution of set of numerators, given $\{\xi_j\}$. Notice that we need information on the *point process* ("random set") $\{\xi_j : j \in \mathcal{P}(\delta,N)\}$, not just on individual terms.

Proposition 3.6. *Suppose that α is distributed according to a bounded density on $\mathbb{T}$. For each δ the point process*

$$\left\{\left(\frac{\ln j}{\ln N}, \langle\langle j\alpha\rangle\rangle j \ln N\right)\right\}_{j\in\mathcal{P}(\delta,N)}$$

converges in distribution as $N\to\infty$ to the Poisson Point Process on $[0,1]\times[-\frac{1}{\delta},\frac{1}{\delta}]$, with constant intensity $\mathbf{c} = \frac{6}{\pi^2}$.

The second coordinate contains the information we need on $\{\xi_j\}$. The information contained in the first coordinate is needed in the proof of Proposition 3.7 below.

Proposition 3.6 is proven in [DF, Theorem 4.1] with $\langle\langle x\rangle\rangle$ replaced by $((x))$. The proof given in [DF] relies on the Poisson limit theorem for the sum along the orbit of the diagonal flow on $SL_2(\mathbb{R})/SL_2(\mathbb{Z})$ of the Siegel transform of functions of the form $F(x,y)=\mathbf{1}_{I_N}(x)\mathbf{1}_{J_N}(y)$ where $\{I_N\}$ and $\{J_N\}$ are sequences of shrinking intervals. To obtain Proposition 3.6 one needs to change slightly the definition of I_N but all the estimates used in [DF] remain valid in the present context.

Proposition 3.7. *Suppose that α is distributed according to a bounded density on $\mathbb{T}$. For every $r>1$ and $\varepsilon>0$ there are δ, N_0 and $A(N,\delta,r)\subset\mathbb{T}$ such that $\mathrm{mes}(A(N,\delta,r))>1-\varepsilon$ and*

(1) If $\alpha\in A(N,\delta,r)$ then $|\mathcal{P}(\delta,N)|>r$ for all $N>N_0$.
(2) For each neighborhood $\mathbb{V}$ of the uniform distribution on $[0,2]^r$ there exists $N_{\mathbb{V}}$ such that for $N\geq N_{\mathbb{V}}$ the following holds. Let $\alpha\in A(N,\delta,r)$ and j_k be an enumeration of the prime resonant harmonics in $\mathcal{P}(\delta,N)$ which orders $\|j_k\alpha\|j_k$ in decreasing order, then the distribution of the random vector

$$(j_1(\alpha(2n+1)+2x),\dots,j_r(\alpha(2n+1)+2x)) \mod 2$$

where $n\sim$ Uniform$\{1,\dots,N\}$ belongs to $\mathbb{V}$.

Proposition 3.7 is proved in §4.4.

Proof of Theorem 2.1. This theorem describes the distributional behavior of $\frac{1}{\ln N}\mathcal{S}_n(\alpha)$ as $(\alpha,n)\sim U(\mathbb{T}\times\{1,\dots,N\})$, when $N\to\infty$. Step 1 says that for every ε there are δ,N_0 such that for all $N>N_0$,

$$\frac{1}{\ln N}\mathcal{S}_n(\alpha,x)=\frac{1}{\ln N}\mathcal{S}_{n,T}^{\mathcal{R}(\delta)}+\Delta_n(\alpha)$$

where $\mathbb{P}(|\Delta_n(\alpha)|\geq\varepsilon)\leq\varepsilon$, as $(\alpha,n)\sim U(\mathbb{T}\times\{1,\dots,N\})$. To see this take $\Delta_n:=\frac{1}{\ln N}(\widehat{\varepsilon}_n+\mathcal{S}_{n,T}^{\mathcal{O}})$, and use Proposition 3.1 and Corollary 3.4.

We will prove Theorem 2.1 by showing that $\frac{1}{\ln N}\mathcal{S}_n^{\mathcal{R}(\delta,N)}(\alpha)\xrightarrow[N\to\infty]{\text{dist}}\mathfrak{C}_\delta$ as $(\alpha,n)\sim U(\mathbb{T}\times\{1,\dots,N\})$, where $\mathfrak{C}_\delta$ are random variables such that $\mathfrak{C}_\delta\xrightarrow[\delta\to0]{\text{dist}}$Cauchy.

Let j_k be an enumeration of $\mathcal{P}(\delta,N)$ which orders $\|j_k\alpha\|j_k$ in decreasing order. By step 2,

$$\frac{\mathcal{S}_{n,T}^{\mathcal{R}(\delta,N)}}{\ln N}=\sum\frac{\Theta_{j_k}(n)+O(\frac{1}{\ln N})}{\xi_{j_k}}.$$

Proposition 3.6 says that the point process $\{\xi_{j_k}\}$ converges in law to the Poisson Point Process on $[-\frac{1}{\delta}, \frac{1}{\delta}]$. Proposition 3.7 says that given $\{\xi_{j_k}\}$, $(\Theta_{j_1}(n)) + O(\frac{1}{\ln N}), \dots, \Theta_{j_{|\mathcal{P}|}}(n) + O(\frac{1}{\ln N})) \xrightarrow[N\to\infty]{\text{dist}} (\Theta_1, \dots, \Theta_{|\mathcal{P}|})$ where Θ_i are are independent identically distributed random variables with distribution (2.5).

It follows that $\frac{\mathcal{S}^{\mathcal{R}(\delta,N)}_{n,T}}{\ln N} \xrightarrow[N\to\infty]{\text{dist}} \sum \frac{\Theta_m}{\xi_m}$ where Θ_i are independent, distributed like (2.5), and $\{\xi_m\}$ is a Poisson Point Process $\mathfrak{C}_\delta$ on $[-\frac{1}{\delta}, \frac{1}{\delta}]$ with density $\mathbf{c} = 6/\pi^2$. In the limit $\mathfrak{C}_\delta \xrightarrow[\delta\to 0]{}$ Cauchy random variable, see Appendix B. This completes the proof of Theorem 2.1 except for the formula (2.1) which is proven in the appendix. □

Proof of Theorem 2.2. Theorem 2.2 describes the convergence in distribution of the ($\mathbb{X}$–valued) random variable

$$\mathcal{F}_N(\alpha)(\cdot) := \frac{1}{N}\text{Card}\left(1 \le n \le N : \tfrac{\mathcal{S}_n(\alpha,x)}{\ln N} \le \cdot\right)$$

to $\mathfrak{F}_\Xi$ as $N \to \infty$, when α is sampled from an absolutely continuous distribution on $\mathbb{T}$ with bounded density. We will assume for simplicity that the density is constant, the changes needed to treat the general case are routine and are left to the reader.

Again we claim that it is enough to prove the result with $\mathcal{S}^{\mathcal{R}}_{n,T}$ replacing $\mathcal{S}_n$. Let $\Delta_n(\alpha)$ be as above. By step 1, for every ε there are δ and N_0 s.t. for all $N > N_0$, $\mathbb{P}(|\Delta_n(\alpha)| \ge \varepsilon) \le \varepsilon$ as $(\alpha, n) \sim U(\mathbb{T} \times \{1, \dots, N\})$. By Fubini's theorem for such N

$$\text{mes}\left(\alpha : \frac{\text{Card}(1 \le n \le N : |\Delta_n(\alpha)| \ge \varepsilon)}{N} \ge \sqrt{\varepsilon}\right) \le \sqrt{\varepsilon}.$$

It follows that the set of α where the asymptotic distributional behavior of $\frac{1}{\ln N}\mathcal{S}^{\mathcal{R}}_{n,T}(\alpha)$ is different from that of $\frac{1}{\ln N}\mathcal{S}_n(\alpha)$ in the limit $N \to \infty, \delta \to 0$ has measure zero.

Thus to prove Theorem 2.2, it is enough to show that

$$\mathcal{F}^\delta_N(\alpha)(\cdot) := \frac{1}{N}\text{Card}\left(1 \le n \le N : \tfrac{\mathcal{S}^{\mathcal{R}}_{n,T}(\alpha)}{\ln N} \le \cdot\right)$$

converges in law, as $N \to \infty$, to an $\mathbb{X}$-valued random variable $\mathfrak{F}^\delta_\Xi$ such that $\mathfrak{F}^\delta_\Xi \xrightarrow[\delta\to 0]{\text{dist}} \mathfrak{F}_\Xi$, where $\mathfrak{F}_\Xi$ is the cumulative distribution function of the random variable $\mathfrak{C}_\Xi$ defined in (2.6) . This is done as before, using Propositions 3.6 and 3.7. □

4. Proofs of the Key Steps.

4.1. **Preliminaries.** The following facts are elementary:

Lemma 4.1.

(a) If the harmonic j is not prime and it is associated to the prime harmonic p, $j = rp$, then $((j\alpha)) = r((p\alpha))$.

(b) For every $x \in \mathbb{R}$, $\frac{2}{\pi} \le \frac{|\sin(\pi x)|}{\pi\|x\|} \le \frac{\pi}{2}$, and

$$|\sin(\pi x)| = \pi\|x\| + O(\|x\|^3) \text{ as } x \to 0.$$

(c) For every $x \in \mathbb{R}$ and $m, N \in \mathbb{N}$,

$$\sum_{j=1}^{m} \sin(y + jx) = \frac{\cos(y + x/2) - \cos(y + (2m+1)x/2)}{2\sin(x/2)} \tag{4.1}$$

$$\sum_{j=1}^{m} \cos(y + 2jx) = \frac{\sin(mx)\cos((m+1)x + y)}{\sin x} \tag{4.2}$$

(d) If α is uniformly distributed on $\mathbb{T}$ then for each $j \neq 0$ $j\alpha$ is also uniformly distributed on $\mathbb{T}$.

(e) For every $0 < a < \frac{1}{2}$, $\mathrm{mes}(\alpha \in \mathbb{T} : \|j\alpha\| < a) = 2a$. (mes =Lebesgue)

(f) $\int_{\mathbb{T}\cap[\|j\alpha\|>a]} \frac{d\alpha}{\|j\alpha\|} = 2\ln\left(\frac{1}{2a}\right)$, $\quad \int_{\mathbb{T}\cap[\|j\alpha\|>a]} \frac{d\alpha}{\|j\alpha\|^2} = \frac{2}{a} - 1$.

Part (e) of Lemma 4.1 implies the following estimates:

$$\lim_{N\to\infty} \mathrm{mes}\left\{\alpha \in \mathbb{T} : \ j\|j\alpha\| > \ln^{-1.1} j \text{ for all } j \ge N/\ln^{10} N\right\} = 1, \tag{4.3}$$

$$\lim_{N\to\infty} \mathrm{mes}(\alpha \in \mathbb{T} : \ j\|j\alpha\| > \ln^{-2} N \text{ for all } j \le 2T) = 1, \tag{4.4}$$

$$\lim_{N\to\infty} \mathrm{mes}\big\{\alpha \in \mathbb{T} : \#\big(j : \|j\alpha\| < \tfrac{\ln^6 N}{N}\ , j < 2T\big) \le \ln^9 N\big\} = 1. \tag{4.5}$$

We prove (4.5), and leave the proofs of (4.3),(4.4) (which are easier) to the reader. Let $\widetilde{F}_N$ denote the set of α in (4.5), then

$$\widetilde{F}_N^c = \{\alpha \in \mathbb{T} : \sum_{j=1}^{2T} \mathbf{1}_{[\|j\alpha\| < \frac{\ln^6 N}{N}]}(\alpha) > \ln^9 N\}.$$

By Lemma 4.1(e), the sum has expectation $\displaystyle\sum_{j=1}^{2N\ln^2 N} \frac{2\ln^6 N}{N}$. By Markov's inequality, $\mathrm{mes}(\widetilde{F}_N^c) \le \frac{1}{\ln^9 N} \displaystyle\sum_{j=1}^{2N\ln^2 N} \frac{2\ln^6 N}{N} \xrightarrow[N\to\infty]{} 0.$

We also observe the following consequnce of (4.2)

$$\left|\sum_{j=1}^{m} \cos(y + 2\pi m x)\right| \leq \min\left(\frac{\pi}{2||x||}, m\right). \tag{4.6}$$

To see that (4.6) is less than $\frac{\pi}{2||x||}$ we estimate the numerator of (4.2) by 1 and the denominator by Lemma 4.1(b). To see that (4.6) is less than m, note that each term in the LHS is less than 1 in absolute value.

We note that (4.3) is a very special case of Khinchine's Theorem on Diophantine approximations (see e.g. [BRV17, Thm 2.3]). This theorem says that if $\varphi : \mathbb{N} \to \mathbb{R}^+$ is a function such that $\sum_q \varphi(q) < \infty$, then for almost every α the inequality

$$\|q\alpha\| < \varphi(q) \tag{4.7}$$

has only finitely many solutions, while if $\sum_q \varphi(q) = \infty$ and φ is non increasing then (4.7) has infinitely many solutions.

Next we list some tightness estimates. Recall that a family of real-valued functions $\{f_n\}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called *tight*, if for every $\varepsilon > 0$ there is an $a > 0$ s.t. $\mathbb{P}(|f_n| > a) < \varepsilon$ for all n.

Lemma 4.2. *Let* $\alpha \sim U(\mathbb{T})$ *then* $\left\{\frac{1}{N \ln N} \sum_{j=1}^{N} \frac{1}{\|j\alpha\|}\right\}_{N \in \mathbb{N}}$ *is tight.*

Proof. For every $\varepsilon > 0$,

$$\text{mes}\left(\alpha : \sum_{j=1}^{N} \frac{1}{\|j\alpha\|} \neq \sum_{j=1}^{N} \left[\frac{1}{\|j\alpha\|} \mathbf{1}_{[\varepsilon/(4N),1/2]}\left(\|j\alpha\|\right)\right]\right) \leq \frac{\varepsilon}{2}, \tag{4.8}$$

because the event in the brackets equals $\bigcup_{j=1}^{N} \{\alpha : \|j\alpha\| < \frac{\varepsilon}{4N}\}$ up to measure zero, and $\text{mes}\{\alpha : \|j\alpha\| < \frac{\varepsilon}{4N}\} = \frac{\varepsilon}{2N}$ by Lemma 4.2(e).

On the other hand, one can check using Lemma 4.2(f) that

$$\mathbb{E}\left(\sum_{j=1}^{N} \left[\frac{1}{\|j\alpha\|} \mathbf{1}_{[\varepsilon/(4N),1/2]}\left(\|j\alpha\|\right)\right]\right) = 2N\left(\ln N + \ln \frac{2}{\varepsilon}\right) \leq 3N \ln N$$

if N is sufficiently large. Hence by Markov's inequality

$$\mathbb{P}\left(\sum_{j=1}^{N} \left[\frac{1}{\|j\alpha\|} \mathbf{1}_{[\varepsilon/(4N),1/2]}\left(\|j\alpha\|\right)\right] \geq \frac{6N \ln N}{\varepsilon}\right) \leq \frac{\varepsilon}{2}. \tag{4.9}$$

Combining (4.8) and (4.9) we see that $\mathbb{P}\left(\sum_{j=1}^{N} \frac{1}{\|j\alpha\|} \geq \frac{6N \ln N}{\varepsilon}\right) \leq \varepsilon$. $\square$

Lemma 4.3. *Let* $\alpha \sim U(\mathbb{T})$, *then the following families of functions are tight as* $N \to \infty$ *(recall that* $T := N \ln^2 N$*):*

(a) $\displaystyle \frac{1}{\ln^2 N}\sum_{j=1}^{T}\frac{1}{j\|j\alpha\|}$.

(b) $\displaystyle \frac{1}{(\ln\ln T)^2}\sum_{j=T\ln^{-\beta_1}T}^{T\ln^{\beta_2}T}\frac{1}{j\|j\alpha\|}\mathbf{1}_{[\ln^{-\beta_3}T,\ln^{\beta_4}T]}(j\|j\alpha\|)$, *for every* $\beta_1,\beta_2,\beta_3,\beta_4$.

(c) $\displaystyle \frac{1}{\ln T\quad\ln\ln T}\sum_{j=1}^{T}\frac{1}{j\|j\alpha\|}\mathbf{1}_{[\ln^{-\beta_3}T,\ln^{\beta_4}T]}(j\|j\alpha\|)$, *for every* β_3,β_4.

We omit the proof, because it is similar to the proof of Lemma 4.2.

4.2. **Step 1 (Propositions 3.1 and 3.3).** The proofs of these Propositions follow [Bec11] closely, but we decided to give all the details, since our assumptions are different.

Proof of Proposition 3.1. The starting point is the Fourier series expansion of $h(x)=\{x\}-\frac{1}{2}$ given in (3.1). Let $h_T(x)=-\displaystyle\sum_{j=1}^{T}\frac{\sin(2\pi jx)}{\pi j}$. Summation by parts (see [Bec11, formula (8.6)]) gives us that

$$\left|\sum_{k=1}^{n}h(x+k\alpha)-\sum_{k=1}^{n}h_T(x+k\alpha)\right|\le\frac{1}{T}\sum_{k=1}^{n}\frac{1}{\|k\alpha\|}\le\frac{1}{T}\sum_{k=1}^{N}\frac{1}{\|k\alpha\|}.$$

The last expression converges to 0 in distribution as $\alpha\sim U(\mathbb{T})$ and $N\to\infty$, by Lemma 4.2 (recall that $T=N\ln^2 N$). Thus it suffices to study the distribution of

$$\sum_{k=1}^{n}h_T(x+k\alpha)-\frac{1}{N}\sum_{n=1}^{N}\left(\sum_{k=1}^{n}h_T(x+k\alpha)\right)\tag{4.10}$$

as $\alpha\sim U(\mathbb{T})$ and $N\to\infty$.

A direct calculation using Lemma 4.1(c) shows that

$$(4.10)=\mathcal{S}_{n,T}(\alpha)+\sum_{j=1}^{T}f_j$$

where

$$f_j=\frac{1}{N}\sum_{n=1}^{N}g_{j,n}=\frac{\sin(2\pi jx+2\pi j\alpha)-\sin(2\pi jx+2\pi(N+1)j\alpha)}{4\pi Nj\sin^2(\pi j\alpha)},$$

(cf. [Bec11, Lemma 8.2]).

Observe that by (4.6) there is a universal constant C such that

$$|f_j| \le C \min\left(\frac{1}{jN\|j\alpha\|^2}, \frac{1}{j\|j\alpha\|}\right). \tag{4.11}$$

Let $\overline{T} := N/\ln^{10} N$. By (4.4), with probability close to 1 we have $1/(j\|j\alpha\|) < \ln^2 N$ for all $j \le T$, so with probability close to 1, all $j \le T$ satisfy $\|j\alpha\| \in [\frac{1}{T\ln^2 N}, \frac{1}{2}]$. We split $\left[\frac{1}{T\ln^2 N}, \frac{1}{2}\right] = \left[\frac{1}{T\ln^2 N}, \frac{1}{N}\right] \cup \left[\frac{1}{N}, \frac{1}{2}\right]$ and apply the bounds in (4.11) to each piece. Thus on

$$F_N := \{\alpha : \forall j \le T,\ j\|j\alpha\| > \ln^{-2} N\} \tag{4.12}$$

we have

$$\sum_{j=\overline{T}+1}^{T} |f_j| \le C \sum_{j=\overline{T}+1}^{T} |\widetilde{f}_j|$$

where

$$\widetilde{f}_j = \frac{\mathbf{1}_{[\frac{1}{N},\frac{1}{2}]}(\|j\alpha\|)}{Nj\|j\alpha\|^2} + \left(\frac{\mathbf{1}_{[\frac{1}{T\ln^2 N},\frac{1}{2}]}(\|j\alpha\|)}{j\|j\alpha\|} - \frac{\mathbf{1}_{[\frac{1}{N},\frac{1}{2}]}(\|j\alpha\|)}{j\|j\alpha\|}\right).$$

Since by Lemma 4.1(f) $\mathbb{E}_\alpha\left(\sum_{j=\overline{T}+1}^{T} \widetilde{f}_j\right) \le C(\ln\ln N)^2$, and by (4.4) $\mathrm{mes}(F_N) \xrightarrow[N\to\infty]{} 1$ we conclude that $\displaystyle\frac{1}{\ln N}\sum_{j=\overline{T}+1}^{T} |f_j| \xrightarrow[N\to\infty]{\text{dist}} 0$ as $(\alpha, n) \sim U(\mathbb{T}\times\{1,\dots,N\})$.

Next, we show that $\displaystyle\frac{1}{\ln T}\sum_{j=1}^{\overline{T}} f_j \xrightarrow[N\to\infty]{\text{dist}} 0$. By (4.11), $\displaystyle\sum_{j=1}^{\overline{T}} |f_j| \le CB_N$ with $\displaystyle B_N = \sum_{j=1}^{\overline{T}} \frac{1}{Nj\|j\alpha\|^2}$, so it is enough to show that $\frac{B_N}{\ln N} \xrightarrow[N\to\infty]{\text{dist}} 0$.

Here is the proof. Split $B_N = B_N^- + B_N^+$ where the first term contains the j s.t. $j\|j\alpha\| \ge \frac{\delta^{-1}}{\ln N}$, and B_N^+ contains the j s.t. $j\|j\alpha\| < \frac{\delta^{-1}}{\ln N}$.

By Lemma 4.1(f)

$$\mathbb{E}_\alpha\left(\frac{B_N^-}{\ln N}\right) \le C\left(\frac{\delta\overline{T}\ln N}{N\ln N}\right) = O\left(\frac{1}{\ln^{10} N}\right).$$

Hence by Markov's inequality $\frac{1}{\ln N}B_N^- \xrightarrow[N\to\infty]{\text{dist}} 0$.

Fix $\varepsilon > 0$ and let F_N be as in (4.12) above, then (4.4) says that $\mathrm{mes}(F_N) > 1-\varepsilon$ for all N large enough.

Let $\overline{\mathcal{R}}(N) = \left\{j \le \overline{T} : \delta j\|j\alpha\|\ln N \le 1\right\}$. By Lemma 4.1(a)

$$\overline{\mathcal{R}}(N) \subset \bigcup_{p\in\mathcal{P}(\delta,N)} \{kp : k\in\mathbb{Z},\ p\in\mathcal{P}(\delta,N),\ kp \le \overline{T}\}. \tag{4.13}$$

Given $p \in \mathcal{P}$, let

$$\mathcal{R}_p(\delta) := \{\text{resonant harmonics in } \mathcal{R}, \text{ associated to } p\}.$$

(4.13) implies that for some constants $C, \overline{C}$

$$B_N^+ \leq \sum_{p\in\mathcal{P}(\delta,N)} \sum_{j\in\mathcal{R}_p(\delta)} \frac{C}{Nj\|j\alpha\|^2} \overset{(\text{Lemma } 4.1(a))}{=} \sum_{p\in\mathcal{P}(\delta,N)} \sum_{kp\in R_p(\delta)} \frac{C}{Nk^3 p\|p\alpha\|^2}$$

$$\leq \sum_{p\in\mathcal{P}(\delta,N)} \frac{C}{Np\|p\alpha\|^2}\left(\sum_{k=1}^{\infty}\frac{1}{k^3}\right) \leq \sum_{p\in\mathcal{P}(\delta,N)} \frac{\overline{C}}{Np\|p\alpha\|^2}$$

$$\leq \sum_{p\in\mathcal{P}(\delta,N)} \frac{\overline{C}\,\overline{T}}{N(p\|p\alpha\|)^2} \overset{(\alpha\in F_N)}{\leq} \sum_{p\in\mathcal{P}(\delta,N)} \frac{\overline{C}\,\overline{T}(\ln^2 N)^2}{N} \leq \frac{\overline{C}\mathrm{Card}(\mathcal{P}(\delta,N))}{\ln^6 N}.$$

By Proposition 3.6, if α is distributed according to a bounded density on $\mathbb{T}$, then $\mathrm{Card}(\mathcal{P}(\delta,N))$ converges in law as $N\to\infty$ to a Poissonian random variable. Therefore there is $K(\varepsilon)$ so that

$$\mathrm{mes}\{\alpha\in\mathbb{T} : \mathrm{Card}(\mathcal{P}(\delta,N)) \leq K(\varepsilon)\} > 1-\varepsilon.$$

It follows that $\mathrm{mes}\{\alpha \in \mathbb{T} : B_N^+ \leq K(\varepsilon)\overline{C}\ln^{-6} N\} > 1-2\varepsilon$, whence $\frac{1}{\ln N}B_N^+ \xrightarrow[N\to\infty]{\text{dist}} 0$.

To summarize,

$$\widehat{\varepsilon}_N := \mathcal{S}_n(\alpha,x) - \mathcal{S}_{n,T}(\alpha) = \sum_{j=1}^{T} f_j + \text{error which tends to 0 in distribution,}$$

and $\displaystyle\frac{1}{\ln N}\sum_{j=1}^{T}|f_j| = \frac{1}{\ln N}\left(\sum_{j=\overline{T}+1}^{T}|f_j| + B_N^- + B_N^+\right) \xrightarrow[N\to\infty]{\text{dist}} 0$ by the arguments above. ∎

Proof of Proposition 3.3. We give the proof assuming that $\alpha \sim U(\mathbb{T})$. The modifications needed to deal with absolutely continuous measures with bounded densities are routine, and are left ot the reader.

Simple algebra gives $\mathcal{V}_N^{\mathcal{O}} = \frac{1}{N}\sum\limits_{j\in\mathcal{O}}\sum\limits_{n=1}^{N} g_{j,n}^2 + \frac{1}{N}\sum\limits_{\substack{j_1,j_2\in\mathcal{O}\\ j_1\neq j_2}} g_{j_1,n}g_{j_2,n}$, where $g_{j,n} = \frac{\cos((2n+1)\pi j\alpha + 2\pi jx)}{2\pi j\sin(\pi j\alpha)}$. To bound the sum of diagonal terms, we use that by (4.6) $|g_{j,n}| < \frac{1}{j\|j\alpha\|}$, whence

$$\frac{1}{N}\sum_{j\in\mathcal{O}}\sum_{n=1}^{N} g_{j,n}^2 \leq \sum_{j\in\mathcal{O}}\frac{1}{N}\sum_{n=1}^{N}\frac{1}{j^2\|j\alpha\|^2} =: \mathrm{Diag}(N). \tag{4.14}$$

To bound the sum of off-diagonal terms, we first collect terms to get

$$\frac{1}{N}\sum_{\substack{j_1,j_2\in\mathcal{O}\\ j_1\neq j_2}} g_{j_1,n}g_{j_2,n} = \sum_{\substack{j_1,j_2\in\mathcal{O}\\ j_1\neq j_2}} \frac{1}{4\pi^2 j_1 j_2 \sin(\pi j_1\alpha)\sin(\pi j_2\alpha)}\widetilde{\Gamma}_{j_1,j_2,N},$$

where

$$\widetilde{\Gamma}_{j_1,j_2,N} := \frac{1}{N}\sum_{n=1}^{N}\cos((2n+1)\pi j_1\alpha + 2\pi j x)\cos((2n+1)\pi j_2\alpha + 2\pi j x).$$

Now the identity $\cos\frac{A+B}{2}\cos\frac{A-B}{2} = \frac{1}{2}[\cos A + \cos B]$ and (4.6) give

$$\widetilde{\Gamma}_{j_1,j_2,N} \leq C\left[\min\left(\frac{1}{N\|(j_1-j_2)\alpha\|},1\right) + \min\left(\frac{1}{N\|(j_1+j_2)\alpha\|},1\right)\right].$$

This and the estimate $j|\sin(\pi j\alpha)| \geq 2j\|j\alpha\|$ (Lemma 4.1(b)) implies that for some universal constant C

$$\frac{1}{N}\sum_{\substack{j_1,j_2\in\mathcal{O}\\ j_1\neq j_2}} g_{j_1,n}g_{j_2,n} \leq C\left[\mathrm{OffDiag}^-(N) + \mathrm{OffDiag}^+(N)\right]$$

where

$$\mathrm{OffDiag}^\pm := \sum_{\substack{j_1,j_2\in\mathcal{O}\\ j_1\neq j_2}} \frac{\min\left(\frac{1}{N\|(j_1\pm j_2)\alpha\|}, N\|(j_1\pm j_2)\alpha\|\right)}{j_1\|j_1\alpha\|\cdot j_2\|j_2\alpha\|}. \tag{4.15}$$

Since $\min(x,x^{-1}) \leq 1$ for all x, the numerator is bounded by one.

Thus $\mathcal{V}_N^{\mathcal{O}} \leq \mathrm{Diag}(N) + C[\mathrm{OffDiag}^- + \mathrm{OffDiag}^+]$. We need the following additional decomposition:

$$\mathrm{OffDiag}^\pm := \mathrm{OffDiag}_i^\pm + \mathrm{OffDiag}_{ii}^\pm + \mathrm{OffDiag}_{iii}^\pm + \mathrm{OffDiag}_{iv}^\pm \tag{4.16}$$

where

(i) $\mathrm{OffDiag}_i^-$: sum over the terms s.t. $\|(j_1-j_2)\alpha\| \geq \frac{\ln^6 N}{N}$;

(ii) $\mathrm{OffDiag}_{ii}^-$: $\|(j_1-j_2)\alpha\| \leq \frac{\ln^6 N}{N}$, and $j_1\|j_1\alpha\| \geq \ln^{10} N$;

(iii) $\mathrm{OffDiag}_{iii}^-$: $\|(j_1-j_2)\alpha\| \leq \frac{\ln^6 N}{N}$, and

$$j_1\|j_1\alpha\| \leq \ln^{10} N, \quad j_2\|j_2\alpha\| \geq \ln^{10} N;$$

(iv) $\mathrm{OffDiag}_{iv}^-$: $\|(j_1-j_2)\alpha\| \leq \frac{\ln^6 N}{N}$,

$$j_1\|j_1\alpha\| \leq \ln^{10} N, \quad j_2\|j_2\alpha\| \leq \ln^{10} N.$$

Similarly for $\mathrm{OffDiag}^+$ with $\|(j_1+j_2)\alpha\|$ instead of $\|(j_1-j_2)\alpha\|$.

We now have the following upper bound

$$\mathcal{V}_N^{\mathcal{O}} \leq \mathrm{Diag} + C\sum_{k=i}^{iv}\left(\mathrm{OffDiag}_k^- + \mathrm{OffDiag}_k^+\right). \tag{4.17}$$

For every $\varepsilon > 0$, and for each of the nine summands $D_1, \dots, D_9$ above, we will construct $\delta_0 > 0$ and Borel sets $A_1(\varepsilon, N), \dots, A_9(\varepsilon, N) \subset \mathbb{T}$ s.t. $\mathrm{mes}[A_i] > 1 - \frac{\varepsilon}{9}$, and with the following property:

$$\forall 0 < \delta < \delta_0, \ \lim_{N\to\infty} \left(\sup_{\alpha \in A_i} \frac{D_i}{\ln^2 N} \right) \leq \frac{\varepsilon}{9}. \tag{4.18}$$

This will prove the proposition, with $E_N(\varepsilon) := \bigcap_{i=1}^{9} A_i$.

We begin by recalling some facts on the typical behavior of $\|j\alpha\|$ for $\alpha \sim U(\mathbb{T})$. Recall that $T := N \ln^2 N$, and let

$$E_N^* := \left\{ \alpha \in \mathbb{T} : \begin{array}{ll} (A) & \forall j > \frac{N}{\ln^{10} N}, \quad j\|j\alpha\| > \ln^{-1.1} j \\ (B) & \forall j \leq 2T, \ j\|j\alpha\| > \ln^{-2} N \\ (C) & \#\{1 \leq j \leq 2T : \|j\alpha\| < \frac{\ln^6 N}{N}\} \leq \ln^9 N \end{array} \right\}.$$

Then $\mathrm{mes}(E_N^*) \xrightarrow[N\to\infty]{} 1$, by (4.3),(4.4), and (4.5). Most of our sets A_i will be subsets of E_N^*.

The summand $\mathrm{Diag} = \sum_{j\in\mathcal{O}} \frac{1}{N} \sum_{n=1}^{N} \frac{1}{j^2\|j\alpha\|^2}$.

Suppose $\alpha \in E_N^*$. By the definition of resonant harmonics, for every $j \in \mathcal{O}$ either $\delta j\|j\alpha\| \ln N \geq 1$, or $N \leq j \leq T$.

In the first case, $j\|j\alpha\| \geq (\delta \ln N)^{-1}$. In the second case, by property (A) of E_N^*, $j\|j\alpha\| > \ln^{-1.1} j \geq \ln^{-1.1} T$. Accordingly,

$$\begin{aligned} \mathbb{E}_\alpha(\mathrm{Diag} \cdot \mathbf{1}_{E_N^*}) &\leq \Bigg[\sum_{j=1}^{N} \int_{\mathbb{T}} \frac{1}{j^2\|j\alpha\|^2} \mathbf{1}_{[j\|j\alpha\| > (\delta \ln N)^{-1}]} d\alpha \\ &\qquad + \sum_{j=N+1}^{T} \int_{\mathbb{T}} \frac{1}{j^2\|j\alpha\|^2} \mathbf{1}_{[j\|j\alpha\| > \ln^{-1.1} T]} d\alpha \Bigg] \\ &\leq \left[\sum_{j=1}^{N} \frac{2\delta j \ln N}{j^2} + \sum_{j=N+1}^{T} \frac{2j \ln^{1.1} T}{j^2} \right] \quad \text{(by Lemma 4.1(f))} \\ &\leq 2\delta \ln^2 N + 4 \ln^{1.1} N \ln \ln N. \end{aligned}$$

Let $\delta_0 := \varepsilon^2/1000$, and choose N_0 so large that for all $N > N_0$, $\mathrm{mes}(E_N^*) > 1 - \varepsilon/9$ and $\frac{4 \ln^{1.1} N \ln \ln N}{\ln^2 N} < \varepsilon^2/1000$. By Markov's inequality, for all $N > N_0$ and $\delta < \delta_0$,

$$\mathrm{mes}\{\alpha \in E_N^* : \mathrm{Diag} > \tfrac{\varepsilon}{9} \ln^2 N\} \leq \frac{2\delta \ln^2 N + 4 \ln \ln N}{(\varepsilon/9) \ln^2 N} < \frac{\varepsilon}{9}.$$

We obtain (4.18) with $D_i = \mathrm{Diag}$, $A_i := \{\alpha \in E_N^* : \mathrm{Diag} \leq \frac{\varepsilon}{9} \ln^2 N\}$, and $\delta_0 := \varepsilon^2/1000$. Notice that this A_i depends on ε.

THE SUMMAND $\mathrm{OffDiag}_i^-$: This is the part of (4.15) with j_1, j_2 s.t. $\|(j_1 - j_2)\alpha\| \geq \frac{\ln^6 N}{N}$.

$$\mathrm{OffDiag}_i^- \leq \sum_{j_1, j_2 \in \mathcal{O}} \frac{1/\ln^6 N}{j_1\|j_1\alpha\| \cdot j_2\|j_2\alpha\|}$$

$$\leq \frac{1}{\ln^6 N}\left(\sum_{j=1}^{T} \frac{1}{j\|j\alpha\|}\right)^2 = \frac{1}{\ln^2 N}\left(\frac{1}{\ln^2 N}\sum_{j=1}^{T} \frac{1}{j\|j\alpha\|}\right)^2.$$

By Lemma 4.3(a), the term in the brackets is tight: there is a constant $K = K(\varepsilon)$ s.t. for all N, $A := \{\alpha \in \mathbb{T} : \frac{1}{\ln^2 N}\sum_{j=1}^{T} \frac{1}{j\|j\alpha\|} \leq K\}$ has measure more than $1 - \frac{\varepsilon}{9}$.

For all $N > \exp\sqrt[4]{\frac{9K^2}{\varepsilon}}$, for every $\alpha \in A$, $\frac{1}{\ln^2 N}\mathrm{OffDiag}_i^- \leq \frac{K^2}{\ln^4 N} < \frac{\varepsilon}{9}$, and we get (4.18) with $A_i = A$, $N_0 = \exp\sqrt[4]{\frac{9K^2}{\varepsilon}}$.

THE SUMMAND $\mathrm{OffDiag}_{ii}^-$: Suppose $\alpha \in E_N^*$, and let $j_2 \in \mathcal{O}$ be an index which appears in $\mathrm{OffDiag}_{ii}^-$. Let $I(j_2)$ be the set of j_1 s.t. (j_1, j_2) appear in $\mathrm{OffDiag}_{ii}^-$, namely

$$I(j_2) := \{j_1 \in \mathcal{O} : \|(j_1 - j_2)\alpha\| \leq \tfrac{\ln^6 N}{N},\ j_1\|j_1\alpha\| \geq \ln^{10} N\}.$$

By property (C) in the definition of E_N^* (applied to $|j_1 - j_2|$), the cardinality of $I(j_2)$ is bounded by $2\ln^9 N$.

Since the numerator in (4.15) is bounded by one, and since in case (ii) $j_1\|j_1\alpha\| \geq \ln^{10} N$, we have

$$\frac{\mathrm{OffDiag}_{ii}^-}{\ln^2 N} \leq \frac{1}{\ln^2 N}\sum_{j_2=1}^{T} \frac{1}{j_2\|j_2\alpha\|}\left(\sum_{j_1 \in I(j_2)} \tfrac{1}{j_1\|j_1\alpha\|}\right)$$

$$\leq \frac{1}{\ln^2 N}\sum_{j_2=1}^{T} \frac{1}{j_2\|j_2\alpha\|}\left(\frac{|I(j_2)|}{\ln^{10} N}\right) \leq \frac{2}{\ln^3 N}\sum_{j=1}^{T} \frac{1}{j\|j\alpha\|} = \frac{2}{\ln N}\left(\tfrac{1}{\ln^2 N}\sum_{j=1}^{T} \tfrac{1}{j\|j\alpha\|}\right).$$

The term in the brackets is tight by Lemma 4.3(a), and we can continue to obtain (4.18) as we did in case (i).

THE SUMMAND $\mathrm{OffDiag}_{iii}^-$: This is similar to $\mathrm{OffDiag}_{ii}^-$.

THE SUMMAND $\mathrm{OffDiag}_{iv}^-$: Suppose the pair of indices (j_1, j_2) appears in $\mathrm{OffDiag}_{iv}^-$: $\|(j_1 - j_2)\alpha\| \leq \frac{\ln^6 N}{N}$, $j_1\|j_1\alpha\| \leq \ln^{10} N$, $j_2\|j_2\alpha\| \leq \ln^{10} N$. Let $j_{\max} := \max(j_1, j_2)$ and $j_{\min} := \min(j_1, j_2)$.

We claim that if $\alpha \in E_N^*$, then

$$\ln^{-2} N \leq j_{\min}\|j_{\min}\alpha\| \leq \ln^{10} N; \tag{4.19}$$

$$\ln^{-1.1} T \leq j_{\max}\|j_{\max}\alpha\| \leq \ln^{10} N; \tag{4.20}$$

$$j_{\max} \geq \frac{N}{\ln^8 N}. \tag{4.21}$$

Here is the proof. The left side of (4.19) is because $j_{\min} \leq T := N \ln^2 N$ by definition of $\mathcal{O}$ and by property (B) in the definition of E_N^*. The right side is because we are in case (iv). Let $\Delta j := j_{\max} - j_{\min}$, then $\Delta j \leq j_{\max} \leq T$, whence by property (B), $\Delta j\|(\Delta j)\alpha\| > \ln^{-2} N$. So $\Delta j > (\|(\Delta j)\alpha\| \ln^2 N)^{-1}$. In case (iv), $\|(\Delta j)\alpha\| \leq \frac{\ln^6 N}{N}$, so $\Delta j \geq \frac{N}{\ln^8 N}$. Since $j_{\max} \geq \Delta j$, (4.21) follows. The right side of (4.20) is by the definition of case (iv). The left side is because of (4.21), property (A) in the definition of E_N^*, and because $j_{\max} \leq T$.

We can now see that for every $\alpha \in E_N^*$,

$$\frac{\mathrm{OffDiag}_{iv}^-(N)}{\ln^2 N} \leq$$

$$\frac{2}{\ln^2 N}\left(\sum_{j=\frac{N}{\ln^8 N}}^{T} \frac{1}{j\|j\alpha\|}\mathbf{1}_{[\ln^{-1.1} T \leq j\|j\alpha\| \leq \ln^{10} N]}\right)\left(\sum_{j=1}^{T} \frac{1}{j\|j\alpha\|}\mathbf{1}_{[\ln^{-2} T \leq j\|j\alpha\| \leq \ln^{10} N]}\right)$$

$$\leq \frac{2(\ln\ln T)^3 \ln T}{\ln^2 N} \times \left(\frac{1}{(\ln\ln T)^2}\sum_{j=\frac{T}{\ln^{10} T}}^{T} \frac{1}{j\|j\alpha\|}\mathbf{1}_{[\ln^{-1.1} T \leq j\|j\alpha\| \leq \ln^{10} N]}\right) \times$$
$$\times \left(\frac{1}{\ln T \ln\ln T}\sum_{j=1}^{T} \frac{1}{j\|j\alpha\|}\mathbf{1}_{[\ln^{-2} T \leq j\|j\alpha\| \leq \ln^{10} N]}\right).$$

The first term tends to zero (because $T = N \ln^2 N$), and the second and third terms are tight by Lemma 4.3(b),(c). We can now proceed as before to obtain (4.18) for $D_i = \mathrm{OffDiag}_{iv}^-$.

The summands $\mathrm{OffDiag}_k^+, k = i, \ldots, iv$: These can be handled in the same way as $\mathrm{OffDiag}_k^-$, except that in cases (ii),(iii) and (iv) we need to apply properties (B) and (C) to $j_1 + j_2$ instead of $|j_1 - j_2|$. This is legitimate since $j_1 + j_2 \leq 2T$. □

4.3. **Step 2 (Proposition 3.5).**

Proof of Proposition 3.5. Fix $\delta > 0$ small, $\alpha \in \mathbb{T} \setminus \mathbb{Q}$, and $N \gg 1$, and set $\mathcal{P} = \mathcal{P}(\delta, N), \mathcal{R} = \mathcal{R}(\delta, N)$. Recall that

$$\mathcal{R}_p(\delta) := \{\text{resonant harmonics in } \mathcal{R}, \text{ associated to } p\}.$$

Then we have the following decomposition for $\mathcal{S}_{n,T}^{\mathcal{R}} / \ln N$

$$\frac{\mathcal{S}_{n,T}^{\mathcal{R}}}{\ln N} \equiv \frac{1}{\ln N} \sum_{j \in \mathcal{R}} g_{j,n} = \frac{1}{\ln N} \sum_{p \in \mathcal{P}} \sum_{kp \in \mathcal{R}_p(\delta)} \frac{\cos((kp)t_n)}{2\pi kp \sin(\pi kp\alpha)}. \tag{4.22}$$

where

$$t_n = (2n+1)\pi\alpha + 2\pi x \tag{4.23}$$

To continue, we need the following observations on $\mathcal{R}_p(\delta)$:

Claim 1: Suppose $p \in \mathcal{P}$, and $L_p := \min(\lfloor \frac{1}{2}\delta p \ln N \rfloor, \lfloor N \ln^2 N/p \rfloor)$. For every $1 \leq k \leq L_p$

(a) $kp \in \mathcal{R}_p(\delta)$ and $((kp\alpha)) = k((p\alpha))$;

(b) $\frac{1}{2\pi kp \sin(\pi kp\alpha)} = \frac{1}{p\langle\langle p\alpha \rangle\rangle} \left(\frac{1}{2\pi^2 k^2} + O\left(\|p\alpha\|^2\right)\right)$.

Proof: Write $p\alpha = m + ((p\alpha))$ with $m \in \mathbb{Z}$. If $j = kp$ with $1 \leq k \leq L_p$, then $j\alpha = km + k((p\alpha))$ and, since p is prime resonant,

$$|k((p\alpha))| = k\|p\alpha\| \leq L_p(\delta p \ln N)^{-1} \leq \frac{1}{2}.$$

Since α is irrational, this is a strict inequality, whence $((j\alpha)) = k((p\alpha))$.

This also shows that $j\alpha = km + ((j\alpha))$ whence $r := \gcd(j, km) = k$, and j is associated to $p \equiv j/r$. To complete the proof that $j \in \mathcal{R}_p(\delta)$ we just need to check that $j \leq N \ln^2 N$. This is immediate from the definition of L_p. This proves (a).

For (b), we write again $p\alpha = m + ((p\alpha))$ with $m \in \mathbb{Z}$ and note that $\sin(\pi kp\alpha) = \sin(\pi km + \pi((kp\alpha))) = (-1)^m \sin \pi((kp\alpha))$. We saw above that if $k \leq L_p$, then $((kp\alpha)) = k((p\alpha))$. So

$$\sin(\pi kp\alpha) = (-1)^m \sin \pi k((p\alpha)) = (-1)^m \pi k((p\alpha)) + O(k^3\|p\alpha\|^3).$$

Note that $(-1)^m((p\alpha)) = \langle\langle p\alpha \rangle\rangle$ because of the decomposition $p\alpha = m + ((p\alpha))$ above. So $\sin(\pi kp\alpha) = \pi k \langle\langle p\alpha \rangle\rangle + O(k^3\|p\alpha\|^3)$. Thus

$$\frac{1}{2\pi kp \sin(\pi kp\alpha)} = \frac{1}{2\pi^2 k^2 p \langle\langle p\alpha \rangle\rangle \left(1 + O\left(k^2||p\alpha||^3\right)\right)} = \frac{1 + O\left(k^2||p\alpha||^3\right)}{2\pi^2 k^2 p \langle\langle p\alpha \rangle\rangle}$$

proving (b).

Claim 2: If $p \in \mathcal{P}$ and $kp \in \mathcal{R}_p(\delta)$, then $\frac{1}{2\pi kp \sin(\pi kp\alpha)} = O\left(\frac{1}{k^2 p \|p\alpha\|}\right)$.

Proof: Write $kp\alpha = \ell + ((kp\alpha))$ with $\ell \in \mathbb{Z}$. Since p is associated to kp, $\gcd(kp,\ell) = k$. We have

$$\|kp\alpha\| = |kp\alpha - \ell| = \gcd(kp,\ell)\left|\frac{kp}{\gcd(kp,\ell)}\alpha - \frac{\ell}{\gcd(kp,\ell)}\right|$$

$$= k\left|p\alpha - \frac{\ell}{\gcd(kp,\ell)}\right| \geq k\|p\alpha\|.$$

In particular, $k\|p\alpha\| \leq \|kp\alpha\| \leq \frac{1}{2}$. Therefore $\|kp\alpha\| = k\|p\alpha\|$, and $|\sin(\pi kp\alpha)| \asymp \pi\|kp\alpha\| = \pi k\|p\alpha\|$, whence $\frac{1}{2\pi kp\sin(\pi kp\alpha)} = O\left(\frac{1}{k^2 p\|p\alpha\|}\right)$.

Claim 3: Suppose $p \in \mathcal{P}$, then

$$\{kp : k = 1, \ldots, L_p\} \subset \mathcal{R}_p(\delta) \subset \left\{kp : k = 1, \ldots, \frac{N\ln^2 N}{p}\right\}.$$

Proof: The first inclusion is by Claim 1(a), the second is because by definition, every $j \in \mathcal{R}$ is less than $T = N\ln^2 N$.

We now return to (4.22). Claims 2 and 3 say that the inner sum is $\displaystyle\sum_{kp\in\mathcal{R}_p(\delta)} g_{kp,n} = \sum_{k=1}^{L_p} g_{kp,n} + O\left(\frac{1}{p\|p\alpha\|}\sum_{k=L_p+1}^{N\ln^2 N/p}\frac{1}{k^2}\right)$. If $L_p = \lfloor N\ln^2 N/p\rfloor$, then the error term vanishes. If not, then $L_p = \lfloor\frac{1}{2}\delta p\ln N\rfloor$ and the error term can be found by summation to be equal to $O\left(\frac{1}{\delta p\|p\alpha\|}\cdot\frac{1}{p\ln N}\right)$. Thus $\sum_{kp\in\mathcal{R}_p(\delta)} g_{kp,n} = \sum_{k=1}^{L_p} g_{kp,n} + O\left(\frac{1}{p\|p\alpha\|}\cdot\frac{1}{p\ln N}\right)$ (where the implied constant is not uniform in δ).

Applying Claim 1 to $\sum_{k=1}^{L_p} g_{kp,n}$ we obtain (recall (4.23))

$$\frac{\mathcal{S}^{\mathcal{R}}_{n,T}}{\ln N} = \frac{1}{\ln N}\sum_{p\in\mathcal{P}}\frac{1}{p\langle\langle p\alpha\rangle\rangle}\left(\sum_{k=1}^{L_p}\left(\frac{\cos((kp)t_n)}{2\pi^2k^2} + O(\|p\alpha\|^2)\right) + O\left(\frac{1}{p\ln N}\right)\right)$$

$$= \frac{1}{\ln N}\sum_{p\in\mathcal{P}}\frac{1}{p\langle\langle p\alpha\rangle\rangle}\left(\sum_{k=1}^{L_p}\frac{\cos((kp)t_n)}{2\pi^2k^2} + O(\|p\alpha\|^2 L_p) + O\left(\frac{1}{p\ln N}\right)\right).$$

For $p \in \mathcal{P}$, $\|p\alpha\| < \frac{1}{\delta p\ln N}$, so $\|p\alpha\|^2 L_p \leq \frac{\frac{1}{2}\delta p\ln N}{\delta^2 p^2\ln^2 N} = O\left(\frac{1}{p\ln N}\right)$. Thus

$$\frac{\mathcal{S}^{\mathcal{R}}_{n,T}}{\ln N} = \sum_{p\in\mathcal{P}}\frac{1}{p\langle\langle p\alpha\rangle\rangle\ln N}\left(\sum_{k=1}^{L_p}\frac{\cos((kp)t_n)}{2\pi^2k^2} + O\left(\frac{1}{p\ln N}\right)\right).$$

Every $p \in \mathcal{P}$ satisfies $p \leq N$. So $L_p \geq \min(\lfloor \frac{1}{2}\delta p \ln N\rfloor, \lfloor \ln^2 N\rfloor)$, whence $\sum\limits_{k=L_p+1}^{\infty} \dfrac{\cos((kp)t_n)}{2\pi^2 k^2} = O\left(\dfrac{1}{p\ln N}\right)$. So

$$\frac{\mathcal{S}_{n,T}^{\mathcal{R}}}{\ln N} = \sum_{p\in\mathcal{P}} \frac{1}{p\langle\langle p\alpha\rangle\rangle \ln N}\left(\sum_{k=1}^{\infty} \frac{\cos(k(pt_n))}{2\pi^2 k^2} + O\left(\frac{1}{\ln N}\right)\right).$$

This is equation (3.2). □

4.4. **Step 3 (Proposition 3.7).**

Proof of Proposition 3.7. We carry out the proof assuming $\alpha \sim U(\mathbb{T})$, and the leave the routine modifications needed to treat the general case to the reader. Fix an integer $r > 1$ and a small real number $\varepsilon > 0$.

Claim 1: There exist $\delta > 0, N_0 \geq 1$ s.t. for all $N > N_0$, $\Omega(N, \delta, r) := \{\alpha \in \mathbb{T} : |\mathcal{P}(\delta, N)| \geq r\}$ has measure bigger than $1 - \varepsilon$.

Proof: Proposition 3.6 says that $\{j\langle\langle j\alpha\rangle\rangle \ln N : j \in \mathcal{P}(\delta, N)\}$ converges as a point process to a Poisson point process with density $\frac{6}{\pi^2}$ on $[-\frac{1}{\delta}, \frac{1}{\delta}]$. So $\#\mathcal{P}(\delta, N) \xrightarrow[N\to\infty]{\text{dist}}$Poisson distribution with expectation $\frac{12}{\pi^2\delta}$. If we choose δ small enough, the probability that this random variable is less than $2r$ is smaller than $\frac{\varepsilon}{2}$. So the claim holds with some N_0.

Let $\mu_{\alpha,r,N} := \frac{1}{N}\sum_{n=1}^{N} \delta_{(j_1 t_n,\ldots,j_r t_n)}$ where $t_n = (2n+1)\pi\alpha + 2\pi x$ (see (4.23)) and fix some weak-star open neighborhood $\mathbb{V}$ of the normalized uniform Lebesgue on $[0, 2]^r$. We will construct $N_{\mathbb{V}}$ s.t. for all $N \geq N_{\mathbb{V}}$,

$$\text{mes}\,(\alpha \in \Omega(N, \delta, r) : \mu_{\alpha,r,N} \in \mathbb{V}) > 1 - 2\varepsilon$$

where $j_k := j_k(\alpha, \delta, N)$ enumerate $\mathcal{P}(\delta, N)$ as in Proposition 3.7.

Without loss of generality, $\mathbb{V} = \bigcap_{\nu=1}^{\nu_0}\{\mu : |\int_{[0,2]^r} f_\nu d\mu| < \overline{\varepsilon}\}$ where $\overline{\varepsilon} > 0$, $\nu_0 \in \mathbb{N}$, and $f_\nu(x_1,\ldots,x_r) = e^{i\pi\sum_{k=1}^{r} l_k^{(\nu)} x_k}$, where

$$\underline{l}^{(\nu)} := (l_1^{(\nu)},\ldots,l_r^{(\nu)}) \in \mathbb{Z}^r \setminus \{\underline{0}\} \quad (1 \leq \nu \leq \nu_0).$$

This is because any weak star open neighborhood of Lebesgue's measure on $[0, 2]^r$ contains a neighborhood of this form. Let

$$L_\nu := L_\nu(\alpha, \delta, N) := \sum_{k=1}^{r} l_k^{(\nu)} j_k.$$

Claim 2: $\text{mes}\,(\alpha \in \Omega(N, \delta, r) : L_\nu = 0 \text{ for some } 1 \leq \nu \leq \nu_0) \xrightarrow[N\to\infty]{} 0.$

Proof. Proposition 3.6 says that the sequence j_k is superlacunary in the sense that for each R

$$\text{mes}\left(\alpha :\ \forall 1 \le k' < k'' \le r :\ \frac{\max(j_{k'}, j_{k''})}{\min(j_{k'}, j_{k''})} \ge R\right) \xrightarrow[N\to\infty]{} 1. \tag{4.24}$$

To see this recall that the gaps between neighboring points of a Poisson process have an exponential distribution, therefore for any ε there exists δ such that the gaps between $\ln j_{k'}$ and $\ln j_{k''}$ are, with probability $1-\varepsilon$, bounded below by $\delta \ln N$.

Let $k^*_\nu = \arg\max(j_k : l_k^{(\nu)} \neq 0)$. Applying (4.24) with

$$R = R_\nu := 2(r-1)\sum_k |l_k^{(\nu)}|$$

we see that for all N large enough, with almost full probability, $j_{k^*_\nu} \ge 2\sum_{k\neq k^*} |l_k^{(\nu)} j_k|$, whence $L_\nu \neq 0$. This proves the claim.

Notice that this argument also shows that with almost full probability, $|L_\nu| \le 2 l_{k^*}^{(\nu)} j_{k^*_\nu}$.

Claim 3: $\text{mes}\left(\alpha \in \Omega(N,\delta,r) : \exists \nu \le \nu_0 \text{ s.t. } \|L_\nu \alpha/2\| \le \frac{\ln N}{N}\right) \xrightarrow[N\to\infty]{} 0.$

Proof. By Proposition 3.6, $\{\frac{\ln j}{\ln N} : j \in \mathcal{P}(\delta, N)\}$ converges as a point process to a Poisson point process with intensity $\frac{12}{\pi^2\delta}$ on $[0,1]$, therefore

$$\begin{aligned}
&\text{mes}\left(\alpha \in \Omega(N,\delta,r) :\ j_{k^*_\nu} > \tfrac{N}{\ln^4 N}\right)\\
&\le \text{mes}\left(\alpha \in \mathbb{T} : \exists j \in \mathcal{P}(\delta,N) \text{ s.t. } \tfrac{\ln j}{\ln N} > 1 - \tfrac{4\ln\ln N}{\ln N}\right) \xrightarrow[N\to\infty]{} 0.
\end{aligned}$$

Therefore for all N large enough, with almost full probability in $\Omega(N,\delta,r)$, $j_{k^*_\nu} \le N/\ln^4 N$, whence (for N large enough) also

$$1 \le |L_\nu| \le 2 l_{k^*_\nu} j_{k^*_\nu} < \frac{N}{\ln^4 N} \max_\nu \|\underline{l}^{(\nu)}\|_\infty < \frac{N}{\ln^3 N}.$$

A simple modification of (4.4) gives

$$\text{mes}(\alpha \in \mathbb{T} :\ j\|j\alpha/2\| > \ln^{-2} N \text{ for all } 1 \le j \le N) \xrightarrow[N\to\infty]{} 1.$$

It follows that for all N large enough, with almost full probability in $\Omega(N,\delta,r)$, $L_\nu \|L_\nu \alpha/2\| > 1/\ln^2 N$, whence

$$\|L_\nu \alpha/2\| > \frac{1}{L_\nu \ln^2 N} > \frac{1}{\frac{N}{\ln^3 N}\cdot \ln^2 N} = \frac{\ln N}{N},$$

proving the claim.

Fix $N_{\mathbb{V}}$ s.t. for every $N > N_{\mathbb{V}}$, for each $1 \le \nu \le \nu_0$, $\|L_\nu \alpha/2\| > \ln N/N$ with almost full probability in $\Omega(N, \delta, r)$. Make $N_{\mathbb{V}}$ so large that $\frac{1}{\ln N_{\mathbb{V}}} < \overline{\varepsilon}$. Then for every $N > N_{\mathbb{V}}$

$$\left|\int f_\nu d\mu_{\alpha,r,N}\right| = \frac{1}{N}\left|\sum_{n=1}^{N} e^{\pi i L_\nu(\alpha(2n+1)+2x)}\right| \le \frac{1}{N}\frac{2}{|1-e^{2\pi i L_\nu \alpha}|}$$
$$= \frac{1}{N\|L\alpha/2\|} \le \frac{1}{N \cdot \frac{\ln N}{N}} = \frac{1}{\ln N} < \overline{\varepsilon},$$

whence $\mu_{\alpha,r,N} \in \mathbb{V}$ as required. $\square$

Appendix A. Proof of proposition 1.1

Suppose $f : \mathbb{T} \to \mathbb{R}$ is differentiable on $\mathbb{T} \setminus \{x_1, \ldots, x_N\}$, and f' extends to a function of bounded variation on $\mathbb{T}$. Since f' has bounded variation, f' is bounded. So f is Lipschitz between its singularities. It follows that $L_i^- := \lim\limits_{t \to x_i^-} f(t)$, $L_i^+ := \lim\limits_{t \to x_i^+} f(t)$ exist for each i.

Let $\varphi(t) := f(t) - \sum_{i=1}^{N}(L_i^+ - L_i^-)h(t + x_1) - \int_{\mathbb{T}} f(s)ds$. It is easy to see that $\varphi|_{\mathbb{T}\setminus\{x_1,\ldots,x_N\}}$ extends to a continuous function ψ on $\mathbb{T}$ s.t.

$$\psi(t) = f(t) - \sum_{i=1}^{N}(L_i^+ - L_i^-)h(t + x_1) - \int_{\mathbb{T}} f(s)ds$$

for every $t \in \mathbb{T} \setminus \{x_1, \ldots, x_N\}$. (But maybe the identity breaks at x_i.)

By construction, $\psi \in C(\mathbb{T})$ and $\psi'|_{\mathbb{T}\setminus\{x_1,\ldots,x_N\}} = f'|_{\mathbb{T}\setminus\{x_1,\ldots,x_N\}}$ extends to a function with bounded variation on $\mathbb{T}$.

Expand ψ to a Fourier series: $\psi(x) = \sum\limits_{k \in \mathbb{Z}-0} a_k e^{2\pi i k x}$. Our assumptions imply $a_k = O(k^{-2})$. A formal solution of $\psi = \kappa_\alpha - \kappa_\alpha \circ R_\alpha$ gives

$$\kappa_\alpha(x) \sim \sum_{k \in \mathbb{Z}\setminus\{0\}} b_{k,\alpha} e^{2\pi i k x}, \quad \text{where} \quad b_{k,\alpha} = \frac{a_k}{1 - e^{2\pi i k \alpha}}.$$

We claim that for a.e. α, $\sum\limits_k |b_{k,\alpha}| < \infty$ so that κ_α is a well-defined continuous function. Indeed $b_{k,\alpha} = O\left(\frac{|a_k|}{||k\alpha||}\right) = O\left(\frac{1}{k^2||k\alpha||}\right)$ so it suffices to check that $\sum\limits_{k=1}^{\infty} \frac{1}{k^2||k\alpha||}$ converges for a.e. α.

By Khinchine theorem for a.e. α we have $||k\alpha|| > k^{-1.1}$ for all large k, so $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2||k\alpha||}$ converges for a.e. α iff

$$\sum_{k=1}^{\infty} \frac{1}{k^2||k\alpha||} \mathbf{1}_{[||k\alpha||>k^{-1.1}]}$$

converges for a.e. α. This is indeed the case, because

$$\int_{\mathbb{T}} \sum_{k=1}^{\infty} \frac{1}{k^2||k\alpha||} \mathbf{1}_{[||k\alpha||>k^{-1.1}]} d\alpha = \sum \frac{1}{k^2} \cdot 2 \ln \left(\frac{k^{1.1}}{2} \right) < \infty. \qquad \square$$

Appendix B. Cauchy and Poisson.

Let $\Xi = \{\xi_n\}$ be a Poisson process on $\mathbb{R}$ with intensity $\mathbf{c}$ and let

$$Y_L = \sum_{|\xi_n|<L} \frac{\Theta_n}{\xi_n}$$

where Θ_n are i.i.d bounded random variables with zero mean independent of Ξ. Our goal is to compute the distributional limit of Y_L as $L \to \infty$. Rewrite $Y_L = \displaystyle\sum_{\xi_n^*<L} \frac{\Theta_n^*}{\xi_n^*}$ where $\xi_n^* = |\xi_n|$, $\Theta_n^* = \frac{|\xi_n|}{\xi_n}\Theta_n$. Note that $\{\xi_n^*\}$ is a Poisson process on $\mathbb{R}^+$ with intensity $2\mathbf{c}$, $|\Theta_n^*| = |\Theta_n|$ and the distribution of Θ^* is symmetric. Let F be the distribution function of Θ^* and $\varphi(t)$ be its characteristic function. Note that φ is real, namely, $\varphi(t) = \mathbb{E}(\cos(t\Theta^*))$ because Θ^* is symmetric. Let $\Phi_R(t)$ be the characteristic function of Y_L. By [RY99, Proposition XII.1.12]

$$\Phi_L(t) = \exp\left[2\mathbf{c} \int_0^L (\varphi(t/y) - 1)\, dy\right].$$

Introducing $x = 1/y$, $\delta = 1/L$ we can rewrite this expression as

$$\Phi_L(t) = \exp\left[2\mathbf{c} \int_\delta^\infty \frac{\varphi(xt) - 1}{x^2} dx\right].$$

The integral in the above expression equals to

$$R_\delta(t) = \int_\delta^\infty \int_{-K}^{K} \frac{\cos(xt\theta) - 1}{x^2} dx dF(\theta)$$

where $K = ||\Theta||_\infty$. Using that the cosine function is even we obtain

$$\int_\delta^\infty \frac{\cos(xt\theta) - 1}{x^2} dx = |t\theta| \int_{\delta|t\theta|}^\infty \frac{\cos(x) - 1}{x^2} dx$$

$$= |t\theta| \int_0^\infty \frac{\cos(x) - 1}{x^2} dx + O\left(\delta t\theta\right) = -\frac{\pi|t||\theta|}{2} + O\left(\delta t\theta\right).$$

Integrating with respect to θ and using that $\mathbb{E}(|\Theta^*|) = \mathbb{E}(|\Theta|)$ we see that

$$\lim_{\delta\to 0} R_\delta(t) = -\frac{\pi|t|}{2}\mathbb{E}(|\Theta|).$$

Hence

$$\lim_{L\to\infty} \Phi_L(t) = \exp\left(-\mathbf{c}\pi|t|\mathbb{E}(|\Theta|)\right).$$

It follows that Y_L converges as $L\to\infty$ to $\boldsymbol{\rho}_2\mathfrak{C}$ where

$$\boldsymbol{\rho}_2 = \mathbf{c}\pi\mathbb{E}(|\Theta|) \tag{B.1}$$

and $\mathfrak{C}$ is the standard Cauchy distribution with density $\frac{1}{\pi(1+x^2)}$.

In particular, for Θ defined by (2.3) it follows from (2.4) that

$$\mathbb{E}(|\Theta|) = \frac{1}{2}\int_0^1 \left|\theta^2 - \theta + \frac{1}{6}\right| d\theta = \frac{1}{18\sqrt{3}}. \tag{B.2}$$

Combining (B.1), (2.7) and (B.2) we get (2.1).

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.

[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.

[BRV17] Victor Beresnevich, Felipe Ramirez, and Sanju Velani. Metric diophantine approximation: aspects of recent work. *London Math. Soc. Lect. Notes*, 437:1–95, 2017.

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

[CGZ00] Francis Comets, Nina Gantert, and Ofer Zeitouni. Quenched, annealed and functional large deviations for one-dimensional random walk in random environment. *Probab. Theory Related Fields*, 118(1):65–114, 2000.

[DF] Dmitry Dolgopyat and Bassam Fayad. Deviations of ergodic sums for toral translations: Boxes. preprint.

[DF15] Dmitry Dolgopyat and Bassam Fayad. Limit theorems for toral translations. In *Hyperbolic dynamics, fluctuations and large deviations*, volume 89 of *Proc. Sympos. Pure Math.*, pages 227–277. Amer. Math. Soc., Providence, RI, 2015.

[DG12] Dmitry Dolgopyat and Ilya Goldsheid. Quenched limit theorems for nearest neighbour random walks in 1D random environment. *Comm. Math. Phys.*, 315(1):241–277, 2012.

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

[Her79] Michael Robert Herman. Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations. *Inst. Hautes Études Sci. Publ. Math.*, (49):5–233, 1979.

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

[Kes60] Harry Kesten. Uniform distribution mod 1. *Ann. of Math. (2)*, 71:445–471, 1960.

[Kes62] Harry Kesten. Uniform distribution mod 1. II. *Acta Arith*, 7:355–380, 1961/1962.

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

[PS13] Jonathon Peterson and Gennady Samorodnitsky. Weak quenched limiting distributions for transient one-dimensional random walk in a random environment. *Ann. Inst. Henri Poincaré Probab. Stat.*, 49(3):722–752, 2013.

[RY99] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

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`