



INSTITUT
POLYTECHNIQUE
DE PARIS

*J*ournal de l'École polytechnique *Mathématiques*

Ayesha HUSSAIN & Youness LAMZOURI

The limiting distribution of Legendre paths

Tome 11 (2024), p. 589-611.

<https://doi.org/10.5802/jep.260>

© Les auteurs, 2024.



Cet article est mis à disposition selon les termes de la licence
LICENCE INTERNATIONALE D'ATTRIBUTION CREATIVE COMMONS BY 4.0.
<https://creativecommons.org/licenses/by/4.0/>

Publié avec le soutien
du Centre National de la Recherche Scientifique



Publication membre du
Centre Mersenne pour l'édition scientifique ouverte
www.centre-mersenne.org
e-ISSN : 2270-518X

THE LIMITING DISTRIBUTION OF LEGENDRE PATHS

BY AYESHA HUSSAIN & YOUNESS LAMZOURI

ABSTRACT. — Let p be a prime number and $(\frac{\cdot}{p})$ be the Legendre symbol modulo p . The *Legendre path* attached to p is the polygonal path whose vertices are the normalized character sums $\frac{1}{\sqrt{p}} \sum_{n \leq j} (\frac{n}{p})$ for $0 \leq j \leq p-1$. In this paper, we investigate the distribution of Legendre paths as we vary over the primes p such that $Q \leq p \leq 2Q$, when Q is large. Our main result shows that as $Q \rightarrow \infty$, these paths converge in law, in the space of real-valued continuous functions on $[0, 1]$, to a certain random Fourier series constructed using Rademacher random completely multiplicative functions. This was previously proved by the first author under the assumption of the Generalized Riemann Hypothesis.

RÉSUMÉ (La répartition limite des chemins de Legendre). — Soient p un nombre premier et $(\frac{\cdot}{p})$ le symbole de Legendre modulo p . Le *chemin de Legendre* attaché à p est le chemin polygonal dont les sommets sont les sommes de caractères normalisées $\frac{1}{\sqrt{p}} \sum_{n \leq j} (\frac{n}{p})$ pour $0 \leq j \leq p-1$. Dans cet article, nous étudions la répartition des chemins de Legendre lorsqu'on varie le premier p dans un intervalle $[Q, 2Q]$, où Q est grand. Notre résultat principal montre que lorsque $Q \rightarrow \infty$, ces chemins convergent en loi, dans l'espace des fonctions continues à valeurs réelles sur $[0, 1]$, vers une certaine série de Fourier aléatoire construite en utilisant des fonctions aléatoires complètement multiplicatives de Rademacher. Ceci a été démontré précédemment par le premier auteur sous l'hypothèse de Riemann généralisée.

CONTENTS

1. Introduction.....	589
2. Properties of the random Fourier series $F_{\mathbb{X}}(t)$	595
3. Proof of convergence in the sense of finite distributions.....	599
4. Tightness of the sequence of processes \mathcal{F}_Q	605
References.....	610

1. INTRODUCTION

A central question in number theory is to understand the behavior and the size of the character sum

$$(1.1) \quad S_p(x) := \sum_{n \leq x} \left(\frac{n}{p} \right),$$

where $(\frac{\cdot}{p})$ is the Legendre symbol modulo an odd prime p and $1 \leq x \leq p$. Such sums encode important information on the distribution of quadratic residues and

MATHEMATICAL SUBJECT CLASSIFICATION (2020). — 11L40, 11N64, 11K65.

KEYWORDS. — Legendre symbol, character sums, Rademacher random multiplicative functions, random Fourier series.

non-residues modulo p . In particular, bounds for $S_p(x)$ lead to results on the size of the least quadratic non-residue modulo p (see for example the works of Ankeny [1]; Burgess [5]; Graham and Ringrose [8]; and Montgomery [24]). Moreover, a recent work of Granville and Mangerel [9] shows that cancellations in the sum $S_p(x)$ in short initial intervals are more or less equivalent to improved bounds for the value of the Dirichlet L -function $L(s, (\frac{\cdot}{p}))$ at $s = 1$, which is connected to class numbers of the quadratic extensions $\mathbb{Q}(\sqrt{\pm p})$.

The Legendre symbol appears to exhibit random behavior over short intervals. This is for example illustrated by the work of Davenport and Erdős [7], who studied the distribution of $S_p(x)$ over short moving intervals. More precisely, they showed that if p is large and x varies among the integers $\{0, 1, \dots, p-1\}$, the short sum $S_p(x+H) - S_p(x)$ tends to a normal distribution with mean zero and variance H , provided $\log H = o(\log p)$ and $H \rightarrow \infty$ as $p \rightarrow \infty$. Recently, Harper [12] showed that this is no longer the case if H is much larger, namely when $H \geq p/(\log p)^A$, and $A > 0$ is a fixed constant.

Our goal in this paper is to gain some understanding on the distribution of $S_p(x)$ over long intervals. In particular we will attempt to answer the following question: do these character sums, viewed as functions of x , possess a limiting distribution as p varies? and if so, can we describe this distribution?

Asked this way, this question does not make sense since one has $\max_{1 \leq x \leq p} |S_p(x)| \gg \sqrt{p}$ (this is an easy consequence of Parseval's theorem applied to Pólya's Fourier expansion (1.4) below), and $\max_{1 \leq x \leq p} |S_p(x)| \ll_\varepsilon \sqrt{p}$ for all but at most $O(\varepsilon \pi(Q))$ primes $p \leq Q$ (this was proved by Montgomery and Vaughan [25]), where $\pi(Q)$ is the number of primes $p \leq Q$. Therefore, we need to normalize the sum $S_p(x)$ by \sqrt{p} . Moreover, since $S_p(x)$ is periodic with period p , we shall normalize the length of the sum by p , and consider the function $t \rightarrow \frac{1}{\sqrt{p}} S_p(tp)$. However, a final complication remains: this function is discontinuous, with jumps at every $t \in \frac{1}{p}\mathbb{N}$. Instead, we shall consider the continuous function f_p on $[0, 1]$, where we concatenate the points where $\frac{1}{\sqrt{p}} S_p(tp)$ changes. We call such a function the *Legendre path* associated to p . Note that for $t \in [0, 1]$ we have

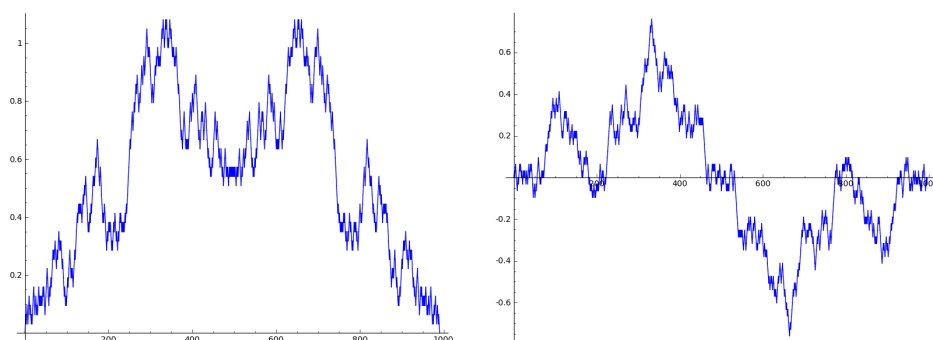
$$(1.2) \quad f_p(t) = \frac{1}{\sqrt{p}} S_p(tp) + \frac{\{pt\}}{\sqrt{p}} \left(\frac{\lceil pt \rceil}{p} \right) = \frac{1}{\sqrt{p}} S_p(tp) + O\left(\frac{1}{\sqrt{p}}\right),$$

where $\{x\}$ is the fractional part of x .

In this paper, we shall prove that these paths, viewed as random processes on $[0, 1]$, have a nice limiting distribution. Before stating our result, we shall construct a probabilistic random model for the Legendre path f_p . A good random model for the Legendre symbol, which was extensively studied in recent years, is a *Rademacher random completely multiplicative function*. This is defined as

$$(1.3) \quad \mathbb{X}_n := \prod_{q^k | n} \mathbb{X}_q^k$$

for $n \in \mathbb{N}$, where $\mathbb{X}_1 = 1$, and $\{\mathbb{X}_q\}_{q \text{ prime}}$ is a sequence of I. I. D. random variables taking the values ± 1 with probability $1/2$ each. The distribution of sums of



The Legendre path for $p = 991 \equiv 3 \pmod{4}$ The Legendre path for $p = 997 \equiv 1 \pmod{4}$.

FIGURE 1. Legendre paths for $p = 991$ and $p = 997$, where the x -axis is pt .

Rademacher random multiplicative functions was extensively studied in recent years, in particular by Harper [10] and [11]. In studying the distribution of the character sum $S_p(x)$, our first guess is to model the normalized sum $\frac{1}{\sqrt{p}}S_p(tp)$ by the following normalized sum of random variables

$$\frac{1}{\sqrt{p}} \sum_{n \leq tp} \mathbb{X}_n.$$

However, it turns out that this is not a good model for $\frac{1}{\sqrt{p}}S_p(tp)$, since it does not account for the periodicity of the Legendre symbol modulo p . To exploit this periodicity, we shall use the Fourier series expansion of f_p , which was first established by Pólya in the following quantitative form [26, eq. (9.19), p. 311]

$$(1.4) \quad f_p(t) = \frac{\tau\left(\left(\frac{\cdot}{p}\right)\right)}{2\pi i \sqrt{p}} \sum_{1 \leq |n| \leq Z} \left(\frac{n}{p}\right) \frac{1 - e(-nt)}{n} + O\left(\frac{1}{\sqrt{p}} + \frac{\sqrt{p} \log p}{Z}\right),$$

for any real number $Z \geq 1$, where $e(t) := e^{2\pi i t}$ and $\tau\left(\left(\frac{\cdot}{p}\right)\right)$ is the Gauss sum associated to the Legendre symbol modulo p , more precisely

$$(1.5) \quad \tau\left(\left(\frac{\cdot}{p}\right)\right) := \sum_{a=1}^p \left(\frac{a}{p}\right) e(a/p) = \begin{cases} \sqrt{p}, & \text{if } p \equiv 1 \pmod{4}, \\ i\sqrt{p}, & \text{if } p \equiv 3 \pmod{4}. \end{cases}$$

REMARK 1.1. — The value of the prime modulo 4 influences the shape of the Legendre path, as shown in Figure 1. In fact, a simple calculation implies that for all $j \in \{0, 1, \dots, p-1\}$ we have $S_p(p-1-j) = S_p(j)$ if $p \equiv 3 \pmod{4}$, and $S_p(p-1-j) = -S_p(j)$ if $p \equiv 1 \pmod{4}$. This shows that if we extend the Legendre path f_p to \mathbb{R} by periodicity, the periodic extension is an even function if $p \equiv 3 \pmod{4}$, and is an odd function if $p \equiv 1 \pmod{4}$.

Inserting the values of the Gauss sum (1.5) in the Fourier expansion (1.4) gives

$$(1.6) \quad f_p(t) = \frac{\varepsilon_p}{2\pi i} \sum_{1 \leq |n| \leq Z} \left(\frac{n}{p}\right) \frac{1 - e(-nt)}{n} + O\left(\frac{1}{\sqrt{p}} + \frac{\sqrt{p} \log p}{Z}\right),$$

where $\varepsilon_p = 1$ if $p \equiv 1 \pmod{4}$, and equals i if $p \equiv 3 \pmod{4}$. Let $\{\mathbb{X}_q\}_{q \text{ prime}}$ be a sequence of I. I. D. random variables taking the values ± 1 with equal probability $1/2$. Since the primes split evenly into the residue classes 1 and 3 $\pmod{4}$, we shall model the value $\left(\frac{-1}{p}\right)$ by a random variable \mathbb{X}_{-1} , which is independent from the \mathbb{X}_q 's for q prime, and takes the values ± 1 with equal probability $1/2$. We also extend the definition of Rademacher random completely multiplicative functions to the negative integers using multiplicativity, namely by setting $\mathbb{X}_{-n} = \mathbb{X}_{-1}\mathbb{X}_n$ for $n \in \mathbb{N}$. Combining all these elements together leads us to model the Legendre path f_p by the following random Fourier series

$$F_{\mathbb{X}}(t) := \frac{\mathbb{Y}}{2\pi i} \sum_{n \in \mathbb{Z} \setminus \{0\}} \mathbb{X}_n \frac{1 - e(-nt)}{n}$$

for $t \in [0, 1]$, where the random variable \mathbb{Y} is defined by $\mathbb{Y} = 1$ if $\mathbb{X}_{-1} = 1$, and $\mathbb{Y} = i$ if $\mathbb{X}_{-1} = -1$. For a fixed $t \in [0, 1]$, it follows from [15, Lem. 1] that the series defining $F_{\mathbb{X}}(t)$ converges almost surely as the limit of the partial sums

$$F_{\mathbb{X},N}(t) := \frac{\mathbb{Y}}{2\pi i} \sum_{0 < |n| \leq N} \mathbb{X}_n \frac{1 - e(-nt)}{n}.$$

Note that $F_{\mathbb{X},N}$ are $C([0, 1])$ -valued random variables. We will show, in Proposition 2.1 below, that the sequence of processes $(F_{\mathbb{X},N})_{N \geq 1}$ converges in $C([0, 1])$ to the process $F_{\mathbb{X}}$ as $N \rightarrow \infty$. As a consequence, we deduce that $F_{\mathbb{X}}$ is almost surely the Fourier series of a continuous function. Figure 2 shows samples of the random process $F_{\mathbb{X}}$ depending on the value of \mathbb{X}_{-1} .

Let Q be a large positive integer. We shall view each $p \mapsto f_p$ for $Q \leq p \leq 2Q$ as a random variable on the finite probability space $\{Q \leq p \leq 2Q\}$ endowed with the uniform probability measure. We denote this random variable by \mathcal{F}_Q . Our main result shows that the stochastic process $(\mathcal{F}_Q)_Q$ converges in law, in the space of continuous functions $C([0, 1])$, to the random process $F_{\mathbb{X}}$.

THEOREM 1.2. — *As $Q \rightarrow \infty$, the sequence of processes $(\mathcal{F}_Q)_Q$ converges to the process $F_{\mathbb{X}}$ in the sense of convergence in law in the Banach space $C([0, 1])$ endowed with the topology of uniform convergence. More precisely, for any continuous and bounded map*

$$\phi : C([0, 1]) \rightarrow \mathbb{R},$$

we have

$$\lim_{Q \rightarrow \infty} \mathbb{E}(\phi(\mathcal{F}_Q)) = \mathbb{E}(\phi(F_{\mathbb{X}})),$$

where

$$\mathbb{E}(\phi(\mathcal{F}_Q)) := \frac{1}{\pi^*(Q)} \sum_{Q \leq p \leq 2Q} \phi(f_p),$$

and $\pi^(Q)$ is the number of primes p in the interval $[Q, 2Q]$.*

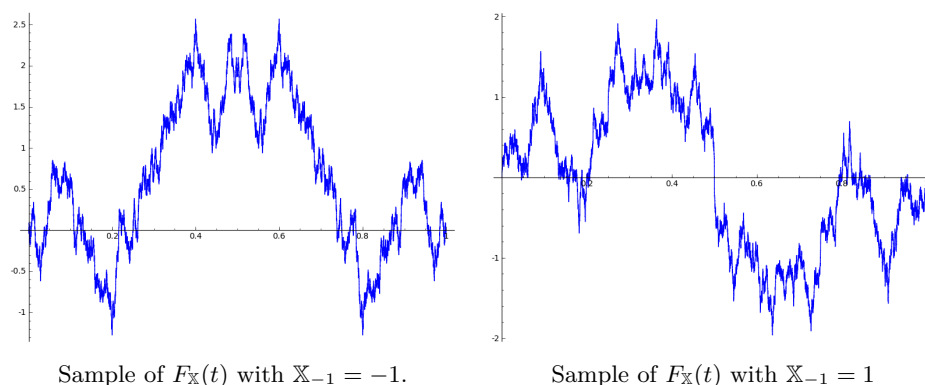


FIGURE 2. Samples, with 10,000 points, of $F_{\mathbb{X}}(t)$, where the x axis is t .

The proof of this result is split into two main parts. We first show in Section 3 that the sequence of processes $(\mathcal{F}_Q)_Q$ converges to the random process $F_{\mathbb{X}}$ in the sense of finite distributions, as $Q \rightarrow \infty$. By Prokhorov’s theorem (see for example [3, Th. 5.1]) convergence in law in $C([0, 1])$ is then equivalent to showing that the sequence of processes $(\mathcal{F}_Q)_Q$ is *tight*, which we establish in Section 4.

REMARK 1.3. — The “graphs” of various exponential sums were investigated by several authors since the seventies. See for example the works of Lehmer [20] and Loxton [21] and [22]. Our work is motivated by the recent papers of Kowalski and Sawin [17] on Kloosterman paths of prime moduli (paths constructed using partial sums of Kloosterman sums), and Ricotta-Royer [27], Ricotta-Royer-Shparlinski [28], and Milicevich-Zhang [23] on generalizations to Kloosterman paths of prime power moduli. However, unlike these works, where the corresponding equidistribution problem can be seen as “vertical”, ours can be viewed as a “horizontal” problem, which is usually believed to be more challenging.

In [14] the first author investigated the distribution of character paths (paths constructed using partial sums of non-principal Dirichlet characters modulo a prime q), and proved the analogue of Theorem 1.2 for these paths. In this case, the sequence of corresponding processes converges in law to an analogous random Fourier series as $q \rightarrow \infty$, where the Fourier coefficients are Steinhaus random multiplicative functions, and where the random variable \mathbb{Y} is uniformly distributed on the unit circle. We should note that since the Legendre symbol is a real character, the Legendre path is a time graph, while the Kloosterman paths and character paths (for non-real Dirichlet characters) are instead polygonal curves in the complex plane.

In her thesis, the first author proved Theorem 1.2 assuming the Generalized Riemann Hypothesis GRH. Our unconditional proof is completely different and uses several new ideas to remove GRH from the different parts of the proof. More precisely, to prove that the sequence of processes $(\mathcal{F}_Q)_Q$ converges in law to the random

process $F_{\mathbb{X}}$ in the sense of finite distributions, we compute in Section 3 the multi-variate moments of $(f_p(t_1), \dots, f_p(t_k))$ (for $0 \leq t_1 < \dots < t_k \leq 1$) unconditionally, by using a recent result of the second author [19] to considerably shorten the Pólya Fourier expansion (1.4) of f_p for almost all p . We also need to bound the contribution of the possible exceptional discriminant. At the end of Section 3, we describe an alternative approach to establishing convergence in the sense of finite distributions, which is discussed in Section 6.2 of Kowalski's book [16], and relies instead on showing that the sequence of "Fourier coefficients" of the Legendre path converges in the sense of finite distributions to those of $F_{\mathbb{X}}$. Although this approach is somewhat simpler, we decided to include both since the original argument gives a quantitative asymptotic formula for the joint moments of the Legendre path f_p at different points, which is interesting on its own, and might have further applications. Furthermore, in Section 4 we prove tightness of the sequence of processes $(\mathcal{F}_Q)_Q$ by appealing to a variant of Kolmogorov's Tightness Criterion from [16], and then using Burgess's bound for short character sums along with the quadratic large sieve inequalities of Heath-Brown [13] and Montgomery and Vaughan [26], in order to bound several moments of $|f_p(t) - f_p(s)|$ depending on the range of $|t - s| \in [0, 1]$. We should note that our proof is very different from the work of the first author [14] on character paths associated to the family of non-principal characters modulo a large prime q , since this family behave very differently from that of Legendre symbols. Indeed, the crucial ingredients used in [14] are the orthogonality relations for characters modulo q , as well as Deligne's bound for hyper-Kloosterman sums, which is used by the first author to control certain twisted moments of Gauss sums.

REMARK 1.4. — Recall that the behavior of the Legendre path attached to p depends on the value of the Legendre symbol $\left(\frac{-1}{p}\right)$. Indeed, it follows from (1.6) that for any real number $Z \geq 1$ we have

$$(1.7) \quad f_p(t) = \frac{1}{\pi} \sum_{1 \leq |n| \leq Z} \left(\frac{n}{p}\right) \frac{\sin(2\pi nt)}{n} + O\left(\frac{1}{\sqrt{p}} + \frac{\sqrt{p} \log p}{Z}\right),$$

if $p \equiv 1 \pmod{4}$, and

$$(1.8) \quad f_p(t) = \frac{1}{\pi} \sum_{1 \leq |n| \leq Z} \left(\frac{n}{p}\right) \frac{1 - \cos(2\pi nt)}{n} + O\left(\frac{1}{\sqrt{p}} + \frac{\sqrt{p} \log p}{Z}\right),$$

if $p \equiv 3 \pmod{4}$. As before, we can view each $p \mapsto f_p$ for $Q \leq p \leq 2Q$ and $p \equiv 1 \pmod{4}$ (respectively $p \equiv 3 \pmod{4}$) as a random variable on the finite probability space $\{Q \leq p \leq 2Q \text{ and } p \equiv 1 \pmod{4}\}$ (respectively $\{Q \leq p \leq 2Q \text{ and } p \equiv 3 \pmod{4}\}$) endowed with the uniform probability measure, and we denote this random variable by $\mathcal{F}_{Q,+}$ (respectively $\mathcal{F}_{Q,-}$). We also do the same with the random Fourier series, where we fix $\mathbb{X}_{-1} = 1$ or -1 . To this end we define for $t \in [0, 1]$

$$F_{\mathbb{X},+}(t) = \frac{1}{\pi} \sum_{n \geq 1} \mathbb{X}_n \frac{\sin(2\pi nt)}{n}, \quad F_{\mathbb{X},-}(t) = \frac{1}{\pi} \sum_{n \geq 1} \mathbb{X}_n \frac{1 - \cos(2\pi nt)}{n}.$$

Then, it follows from the proof of Theorem 1.2 that the sequence of processes $(\mathcal{F}_{Q,\pm})_Q$ converges to the process $F_{\mathbb{X},\pm}$.

The second author investigated in [19] the distribution of $\frac{1}{\sqrt{p}} \max_{1 \leq x \leq p} |S_p(x)|$, as p varies over the primes satisfying $Q \leq p \leq 2Q$. In particular, he proved that the tail of the distribution is double exponentially decreasing in a large uniform range, which is believed to be best possible. Since the maximum of the Legendre path f_p occurs at its vertices and $\|\cdot\|_\infty$ is a continuous function on $C([0, 1])$, a direct consequence of Theorem 1.2 is the existence of a limiting distribution of $\frac{1}{\sqrt{p}} \max_{1 \leq x \leq p} |S_p(x)|$.

COROLLARY 1.5. — *Let μ be the probability measure on $[0, \infty)$ associated to the random variable $\|F_{\mathbb{X}}\|_\infty$. Then for any bounded continuous function h on $[0, \infty)$, we have*

$$\lim_{Q \rightarrow \infty} \frac{1}{\pi^*(Q)} \sum_{Q \leq p \leq 2Q} h\left(\frac{1}{\sqrt{p}} \max_{1 \leq x \leq p} |S_p(x)|\right) = \int_0^\infty h(x) d\mu(x).$$

This result is the analogue of [4, Th. 1.4], which was proved (using a different and more direct approach) by Bober, Goldmakher, Granville and Koukoulopoulos for the family of non-principal characters modulo a large prime q .

Acknowledgements. — The authors would like to thank the anonymous referees for carefully reading the paper, and for several helpful suggestions, which led in particular to the simplification of the proof of Proposition 4.5, as well as to the new Section 3.2 on the convergence of the sequence of Fourier coefficients of the path in the sense of finite distributions. The second author is supported by a junior chair of the Institut Universitaire de France. Part of this work was completed while the second author was on a Délégation CNRS at the IRL3457 CRM-CNRS in Montréal. The second author would like to thank the CNRS for its support and the Centre de Recherches Mathématiques for its excellent working conditions.

2. PROPERTIES OF THE RANDOM FOURIER SERIES $F_{\mathbb{X}}(t)$

2.1. ALMOST SURE CONTINUITY OF $F_{\mathbb{X}}$. — Recall that $F_{\mathbb{X}}(t)$ is defined as the limit of the finite sum $F_{\mathbb{X},N}(t)$ as $N \rightarrow \infty$, and that for fixed $t \in [0, 1]$, the series defining $F_{\mathbb{X}}(t)$ is almost surely convergent by [15, Lem. 1]. In order to prove that $F_{\mathbb{X}}$ is almost surely continuous, we use the fact that $F_{\mathbb{X},N}$ are $C([0, 1])$ -valued random variables, and then prove that the sequence of processes $(F_{\mathbb{X},N})_N$ converges in law to the process $F_{\mathbb{X}}$. To this end we will show that $(F_{\mathbb{X},N})_N$ converges to $F_{\mathbb{X}}$ in the sense of finite distributions, and then show that the sequence $(F_{\mathbb{X},N})_N$ is tight.

PROPOSITION 2.1. — *As $N \rightarrow \infty$, the sequence of processes $(F_{\mathbb{X},N})_N$ converges in the Banach space $C([0, 1])$. In particular, its limit $F_{\mathbb{X}}$ is almost surely the Fourier series of a continuous function.*

Proof. — Let $t \in [0, 1]$. Then, we have

$$\begin{aligned} \mathbb{E}(|F_{\mathbb{X}}(t) - F_{\mathbb{X},N}(t)|^2) &= \frac{1}{(2\pi)^2} \mathbb{E}\left(\left|\sum_{|n|>N} \mathbb{X}_n \frac{1 - e(-nt)}{n}\right|^2\right) \\ &= \frac{1}{(2\pi)^2} \sum_{\substack{|n_1|, |n_2| > N \\ n_1 n_2 = \square}} \frac{(1 - e(-n_1 t))(1 - e(n_2 t))}{n_1 n_2} \mathbb{E}(\mathbb{X}_{n_1 n_2}). \end{aligned}$$

Since $\mathbb{E}(\mathbb{X}_n) = 1$ if n is a square, and equals 0 otherwise, we deduce that

$$\mathbb{E}(|F_{\mathbb{X}}(t) - F_{\mathbb{X},N}(t)|^2) \ll \sum_{\substack{|n_1|, |n_2| > N \\ n_1 n_2 = \square}} \frac{1}{|n_1 n_2|} \ll \sum_{n > N} \frac{d(n^2)}{n^2} \ll \sum_{n > N} \frac{1}{n^{3/2}} \ll \frac{1}{N^{1/2}},$$

upon using the bound $d(m) \ll m^{1/4}$, where $d(m)$ is the divisor function. Therefore, $F_{\mathbb{X},N}$ converges to $F_{\mathbb{X}}$ in $L^2([0, 1])$ and hence in $L^1([0, 1])$ as $N \rightarrow \infty$. By [16, Lem. 3] we deduce that the sequence of processes $(F_{\mathbb{X},N})_N$ converges to the process $F_{\mathbb{X}}$ in the sense of finite distributions.

To complete the proof, we need to show that the sequence of processes $(F_{\mathbb{X},N})_{N \geq 1}$ is tight. By Kolmogorov's tightness criterion (see for example [16, Prop. B.11.10]), we need to prove the existence of constants $C \geq 0$, $\alpha > 0$ and $\delta > 0$ such that for any $N \geq 1$, and real numbers $0 \leq s < t \leq 1$ we have

$$\mathbb{E}(|F_{\mathbb{X},N}(t) - F_{\mathbb{X},N}(s)|^\alpha) \leq C|t - s|^{1+\delta}.$$

We will prove this with $\alpha = 4$ and $\delta = 1/2$. Let $0 \leq s < t \leq 1$ be real numbers. For an integer $n \neq 0$ we define

$$(2.1) \quad g(n) := \sum_{\substack{n_1, n_2 \in \mathbb{Z} \\ n_1 n_2 = n}} (e(-n_1 s) - e(-n_1 t))(e(-n_2 s) - e(-n_2 t)).$$

Then we have

$$\begin{aligned} \mathbb{E}(|F_{\mathbb{X},N}(t) - F_{\mathbb{X},N}(s)|^4) &\ll \mathbb{E}\left(\left|\sum_{1 \leq |n| \leq N} \mathbb{X}_n \frac{e(-ns) - e(-nt)}{n}\right|^4\right) \\ (2.2) \quad &= \mathbb{E}\left(\left|\sum_{1 \leq |n| \leq N^2} \frac{g(n) \mathbb{X}_n}{n}\right|^2\right) = \sum_{\substack{1 \leq |n_1|, |n_2| \leq N^2 \\ n_1 n_2 = \square}} \frac{g(n_1) \overline{g(n_2)}}{n_1 n_2}. \end{aligned}$$

Let $\varepsilon > 0$ be a small real number. Note that $|e(\alpha_1 n) - e(\alpha_2 n)| \ll \min(1, |n(\alpha_1 - \alpha_2)|)$, and hence

$$\begin{aligned} (2.3) \quad g(n) &\ll \sum_{\substack{n_1, n_2 \in \mathbb{Z} \\ n_1 n_2 = n}} \min(1, |n_1(t - s)|) \min(1, |n_2(t - s)|) \ll d(|n|) \min(1, |n|(t - s)^2) \\ &\ll_\varepsilon |n|^{\varepsilon/4} \min(1, |n|(t - s)^2), \end{aligned}$$

since $d(|n|) \ll_\varepsilon |n|^{\varepsilon/4}$. This implies

$$\begin{aligned} \sum_{\substack{1 \leq |n_1|, |n_2| \leq N^2 \\ n_1 n_2 = \square}} \left| \frac{g(n_1)g(n_2)}{n_1 n_2} \right| &\ll_\varepsilon \sum_{\substack{1 \leq |n_1|, |n_2| \leq N^2 \\ n_1 n_2 = \square}} \frac{\min(1, |n_1 n_2|(t-s)^4)}{|n_1 n_2|^{1-\varepsilon/4}} \\ &\ll_\varepsilon \sum_{n=1}^{\infty} d(n^2) \frac{\min(1, n^2(t-s)^4)}{n^{2-\varepsilon/2}}. \end{aligned}$$

We now use the bound $d(n^2) \leq d(n)^2 \ll_\varepsilon n^{\varepsilon/2}$ to get

$$(2.4) \quad \sum_{\substack{1 \leq |n_1|, |n_2| \leq N^2 \\ n_1 n_2 = \square}} \left| \frac{g(n_1)g(n_2)}{n_1 n_2} \right| \ll_\varepsilon \sum_{n \geq 1/|t-s|^2} \frac{1}{n^{2-\varepsilon}} + |t-s|^4 \sum_{n \leq 1/|t-s|^2} n^\varepsilon \ll_\varepsilon |t-s|^{2-2\varepsilon}.$$

Inserting this estimate in (2.2) and choosing $\varepsilon = 1/4$ implies that

$$\mathbb{E} \left(|F_{\mathbb{X},N}(t) - F_{\mathbb{X},N}(s)|^4 \right) \leq C|t-s|^{3/2},$$

for all positive integers N and real numbers $0 \leq s < t \leq 1$, where $C > 0$ is an absolute constant. This completes the proof. \square

2.2. THE SEQUENCE OF JOINT MOMENTS OF THE RANDOM PROCESS IS DETERMINATE

In Section 3 we shall use the method of moments to prove that the sequence of processes $(\mathcal{F}_Q)_Q$ converges to the process $F_{\mathbb{X}}$ in the sense of convergence of finite distributions, as $Q \rightarrow \infty$. More precisely, given an integer $k \geq 1$ and a k -tuple $0 \leq t_1 < \dots < t_k \leq 1$, we will prove that the joint moments of $(\mathcal{F}_Q(t_1), \dots, \mathcal{F}_Q(t_k))$ converge to those of $(F_{\mathbb{X}}(t_1), \dots, F_{\mathbb{X}}(t_k))$ as $Q \rightarrow \infty$. However, we first need to verify that these moments are determinate, i.e. that they have only one representing measure. For $\mathbf{n} = (n_1, \dots, n_k) \in (\mathbb{Z}_{\geq 0})^k$ we let $n = \sum_{j=1}^k n_j$, and define

$$(2.5) \quad M_{\mathbb{X}}(\mathbf{n}) := \mathbb{E} \left(\prod_{i=1}^k F_{\mathbb{X}}(t_i)^{n_i} \right), \quad M_{\mathbb{X},\pm}(\mathbf{n}) := \mathbb{E} \left(\prod_{i=1}^k F_{\mathbb{X},\pm}(t_i)^{n_i} \right).$$

Then we have

$$\begin{aligned} M_{\mathbb{X}}(\mathbf{n}) &= \mathbb{E} \left(\prod_{i=1}^k F_{\mathbb{X}}(t_i)^{n_i} \mid \mathbb{X}_{-1} = 1 \right) \mathbb{P}(\mathbb{X}_{-1} = 1) \\ (2.6) \quad &+ \mathbb{E} \left(\prod_{i=1}^k F_{\mathbb{X}}(t_i)^{n_i} \mid \mathbb{X}_{-1} = -1 \right) \mathbb{P}(\mathbb{X}_{-1} = -1) \\ &= \frac{1}{2} (M_{\mathbb{X},+}(\mathbf{n}) + M_{\mathbb{X},-}(\mathbf{n})), \end{aligned}$$

since $\mathbb{P}(\mathbb{X}_{-1} = 1) = \mathbb{P}(\mathbb{X}_{-1} = -1) = 1/2$. Therefore, using the definition of $F_{\mathbb{X},+}(t)$ and expanding the moment $M_{\mathbb{X},+}(\mathbf{n})$ we obtain

$$(2.7) \quad \begin{aligned} M_{\mathbb{X},+}(\mathbf{n}) &= \frac{1}{\pi^n} \mathbb{E} \left(\sum_{a_{1,1}, \dots, a_{k,n_k} \geq 1} \prod_{i=1}^k \prod_{j=1}^{n_i} \mathbb{X}_{a_{i,j}} \frac{\sin(2\pi a_{i,j} t_i)}{a_{i,j}} \right) \\ &= \frac{1}{\pi^n} \mathbb{E} \left(\sum_{a=1}^{\infty} \frac{\mathbb{X}_a}{a} \mathcal{B}_{\mathbf{n},\mathbf{t},+}(a) \right) = \frac{1}{\pi^n} \sum_{a=1}^{\infty} \frac{\mathcal{B}_{\mathbf{n},\mathbf{t},+}(a^2)}{a^2}, \end{aligned}$$

where

$$\mathcal{B}_{\mathbf{n},\mathbf{t},+}(a) := \sum_{a_1 \cdots a_k = a} \prod_{i=1}^k \sum_{b_{i,1} \cdots b_{i,n_i} = a_i} \prod_{j=1}^{n_i} \sin(2\pi b_{i,j} t_i).$$

We observe that uniformly for \mathbf{t} we have

$$(2.8) \quad |\mathcal{B}_{\mathbf{n},\mathbf{t},+}(a)| \leq \sum_{a_1 \cdots a_k = a} \prod_{i=1}^k d_{n_i}(a_i) = d_n(a),$$

where d_n is the n -th divisor function. Similarly, one has

$$(2.9) \quad M_{\mathbb{X},-}(\mathbf{n}) = \frac{1}{\pi^n} \sum_{a=1}^{\infty} \frac{\mathcal{B}_{\mathbf{n},\mathbf{t},-}(a^2)}{a^2},$$

where

$$\mathcal{B}_{\mathbf{n},\mathbf{t},-}(a) := \sum_{a_1 \cdots a_k = a} \prod_{i=1}^k \sum_{b_{i,1} \cdots b_{i,n_i} = a_i} \prod_{j=1}^{n_i} (1 - \cos(2\pi b_{i,j} t_i)).$$

We also have the following analogous bound to (2.8)

$$(2.10) \quad |\mathcal{B}_{\mathbf{n},\mathbf{t},-}(a)| \leq 2^n d_n(a).$$

We now prove the following lemma.

LEMMA 2.2. — *The sequence of moments $(M_{\mathbb{X}}(\mathbf{n}))_{\mathbf{n} \in (\mathbb{Z}_{\geq 0})^k}$ have only one representing measure.*

Proof. — It is sufficient to show that the moments satisfy the Carleman condition [29, Th. 15.11]:

$$(2.11) \quad \sum_{n=1}^{\infty} |\mathbb{E}(F_{\mathbb{X}}(t_i)^{2n})|^{-1/2n} = \infty,$$

for all $1 \leq i \leq k$.

Let $t \in [0, 1]$ be fixed. Combining the identities (2.6), (2.7), and (2.9) with the upper bounds (2.8) and (2.10) (with $k = 1$) we deduce that

$$|\mathbb{E}(F_{\mathbb{X}}(t)^{2n})| \leq \frac{2^{2n} + 1}{2\pi^{2n}} \sum_{a=1}^{\infty} \frac{d_{2n}(a^2)}{a^2}.$$

Now, it follows from [18, Lem. 3.3] that

$$\sum_{a=1}^{\infty} \frac{d_{2n}(a^2)}{a^2} \leq \sum_{a=1}^{\infty} \frac{d_{2n}(a)^2}{a^2} = \exp(O(n \log \log n)).$$

Combining these estimates implies the existence of positive constants C_1, C_2 such that for all integers $n \geq 2$

$$|\mathbb{E}(F_{\mathbb{X}}(t)^{2n})|^{-\frac{1}{2n}} \geq C_2 \frac{1}{(\log n)^{C_1}},$$

and hence

$$\sum_{n=2}^{\infty} |\mathbb{E}(F_{\mathbb{X}}(t)^{2n})|^{-\frac{1}{2n}} \geq C_2 \sum_{n=2}^{\infty} \frac{1}{(\log n)^{C_1}} = \infty.$$

This concludes the proof. \square

3. PROOF OF CONVERGENCE IN THE SENSE OF FINITE DISTRIBUTIONS

3.1. CONVERGENCE IN THE SENSE OF FINITE DISTRIBUTIONS OF THE PROCESS. — In this section we will prove the following theorem.

THEOREM 3.1. — *As $Q \rightarrow \infty$ the sequence of processes $(\mathcal{F}_Q)_Q$ converges to the process $F_{\mathbb{X}}$ in the sense of convergence of finite distributions. In other words, for every $k \geq 1$ and for every k -tuple $0 \leq t_1 < \dots < t_k \leq 1$, the vectors $(\mathcal{F}_Q(t_1), \dots, \mathcal{F}_Q(t_k))$ converge in law, as $Q \rightarrow \infty$, to $(F_{\mathbb{X}}(t_1), \dots, F_{\mathbb{X}}(t_k))$.*

In order to prove this result it suffices to show that the joint moments of $(\mathcal{F}_Q(t_1), \dots, \mathcal{F}_Q(t_k))$ converge to those of $(F_{\mathbb{X}}(t_1), \dots, F_{\mathbb{X}}(t_k))$ as $Q \rightarrow \infty$, since it follows from Section 2.2 that the distribution of the latter random vector is completely determined by its joint moments. Therefore, Theorem 3.1 is a direct consequence of the following proposition.

PROPOSITION 3.2. — *Let $A > 1$ be a fixed constant. Let k be a positive integer, $\mathbf{n} = (n_1, \dots, n_k) \in (\mathbb{Z}_{\geq 0})^k$ and $0 \leq t_1 < \dots < t_k \leq 1$ be real numbers. Then we have*

$$(3.1) \quad \frac{1}{\pi^*(Q)} \sum_{Q \leq p \leq 2Q} \prod_{i=1}^k f_p(t_i)^{n_i} = M_{\mathbb{X}}(\mathbf{n}) + O_{\mathbf{n}, A}\left(\frac{1}{(\log Q)^A}\right),$$

where $M_{\mathbb{X}}(\mathbf{n})$ is defined in (2.5), and the implicit constant in the error term is not effective.

Conditionally on GRH, the first author established this result in her thesis [14], with a much stronger error term $Q^{-1/2+o(1)}$. In order to compute these moments unconditionally, the main ingredient is a recent work of the second author [19], which allows us to substantially reduce the length of the Pólya Fourier expansion of f_p , for almost all primes p such that $Q \leq p \leq 2Q$ (see Lemma 3.4 below). Another difficulty in proving Proposition 3.2 (which explains the weak error term of (3.1)) arises from the possible existence of Landau-Siegel zeros. Let x be a large real number. By the

results of [6, Chap. 20], it follows that for all square-free integers $|d| \leq \exp(\sqrt{\log x})$ with at most one exception q_1 , we have

$$(3.2) \quad \sum_{p \leq x} \chi_d(p) \ll x \exp(-c\sqrt{\log x}),$$

for some positive constant c , where χ_d denotes the Kronecker symbol attached to d . Moreover, by Siegel's theorem this exceptional discriminant if it exists must satisfy

$$(3.3) \quad |q_1| \gg_A (\log x)^{4A},$$

for any fixed constant $A > 0$, where the implicit constant is not effective. To lighten our notation we shall denote by $\chi_p := \left(\frac{\cdot}{p}\right)$ the Legendre symbol modulo a prime p , throughout the remaining part of the paper. The following lemma is standard (see [6]) but we include its proof for the sake of completeness.

LEMMA 3.3. — *Let Q be large and $n \leq \exp(\sqrt{\log Q})$ be a positive integer. Then, we have*

$$\begin{aligned} & \sum_{\substack{Q \leq p \leq 2Q \\ p \equiv 1 \pmod{4}}} \chi_p(n) \\ &= \begin{cases} \frac{\pi^*(Q)}{2} + O\left(Q \exp(-c\sqrt{\log Q})\right), & \text{if } n = m^2, \\ O\left(Q \exp(-c\sqrt{\log Q})\right), & \text{if } n \text{ is not of the form } m^2 \text{ or } |q_1| \cdot m^2, \end{cases} \end{aligned}$$

for some positive constant c . Moreover, the same estimate holds for $\sum_{\substack{Q \leq p \leq 2Q \\ p \equiv 3 \pmod{4}}} \chi_p(n)$.

Proof. — We only prove the estimate for the character sum over primes $p \equiv 1 \pmod{4}$ since the proof in the case $p \equiv 3 \pmod{4}$ is similar. The first estimate when n is a square follows simply from the prime number theorem in arithmetic progressions together with the trivial bound $\omega(n) \leq \log n$, where $\omega(n)$ is the number of prime factors of n . Now, suppose that n is not a square, and write $n = dm^2$ where d is square-free and $d \neq |q_1|$. By the law of quadratic reciprocity we deduce that

$$\sum_{\substack{Q \leq p \leq 2Q \\ p \equiv 1 \pmod{4}}} \chi_p(n) = \sum_{\substack{Q \leq p \leq 2Q \\ p \equiv 1 \pmod{4}}} \chi_p(d) + O(\omega(n)) = \sum_{\substack{Q \leq p \leq 2Q \\ p \equiv 1 \pmod{4}}} \chi_d(p) + O(\sqrt{\log Q}).$$

The second estimate follows from (3.2), together with the fact that

$$\sum_{\substack{Q \leq p \leq 2Q \\ p \equiv 1 \pmod{4}}} \chi_d(p) = \frac{1}{2} \sum_{Q \leq p \leq 2Q} \chi_d(p) + \frac{1}{2} \sum_{p \leq Q} \chi_{-d}(p). \quad \square$$

In order to compute the joint moments of $f_p(t_1), \dots, f_p(t_k)$ we need to considerably shorten the sum in the main term of (1.4), for all $t \in \{t_1, t_2, \dots, t_k\}$. To this end we shall use the following result from a recent work of the second author [19] on the distribution of large values of quadratic character sums.

LEMMA 3.4. — Let $B > 0$ be a fixed constant. For all real numbers N_1, N_2 verifying $\exp((\log \log Q)^2) \leq N_1 < N_2 \leq 2N_1 \leq Q^{21/40}$, the number of primes p such that $Q \leq p \leq 2Q$ and

$$\max_{\alpha \in [0,1)} \left| \sum_{N_1 \leq n \leq N_2} \frac{\chi_p(n) e(\alpha n)}{n} \right| \geq \frac{1}{(\log N_1)^B},$$

is

$$\ll_B Q \exp\left(-\frac{\log Q}{10 \log \log Q}\right).$$

Proof. — This is a direct consequence of [19, Prop. 4.2 & 4.3] by embedding the set of primes p such that $Q \leq p \leq 2Q$ into the set of all fundamental discriminants $|d| \leq 2Q$. \square

We now have all the ingredients to prove Proposition 3.2.

Proof of Proposition 3.2. — Let k be a positive integer, $\mathbf{n} = (n_1, \dots, n_k) \in (\mathbb{Z}_{\geq 0})^k$ and $0 \leq t_1 < \dots < t_k \leq 1$ be real numbers. Throughout the proof we shall assume that Q is sufficiently large in terms of $n = n_1 + \dots + n_k$. To shorten our notation we define

$$M_Q(\mathbf{n}) := \frac{1}{\pi^*(Q)} \sum_{Q \leq p \leq 2Q} \prod_{i=1}^k f_p(t_i)^{n_i}.$$

Then, we note that

$$M_Q(\mathbf{n}) = \frac{1}{2} (M_{Q,+}(\mathbf{n}) + M_{Q,-}(\mathbf{n})),$$

where

$$(3.4) \quad \begin{aligned} M_{Q,+}(\mathbf{n}) &:= \frac{2}{\pi^*(Q)} \sum_{\substack{Q \leq p \leq 2Q \\ p \equiv 1 \pmod{4}}} \prod_{i=1}^k f_p(t_i)^{n_i}, \\ M_{Q,-}(\mathbf{n}) &:= \frac{2}{\pi^*(Q)} \sum_{\substack{Q \leq p \leq 2Q \\ p \equiv 3 \pmod{4}}} \prod_{i=1}^k f_p(t_i)^{n_i}. \end{aligned}$$

Therefore, in view of (2.6) it suffices to establish the following asymptotic formulas

$$(3.5) \quad \begin{aligned} M_{Q,+}(\mathbf{n}) &= M_{\mathbb{X},+}(\mathbf{n}) + O_{\mathbf{n},A}\left(\frac{1}{(\log Q)^A}\right), \\ M_{Q,-}(\mathbf{n}) &= M_{\mathbb{X},-}(\mathbf{n}) + O_{\mathbf{n},A}\left(\frac{1}{(\log Q)^A}\right). \end{aligned}$$

We shall only prove this estimate for $M_{Q,+}(\mathbf{n})$, since the proof for $M_{Q,-}(\mathbf{n})$ is similar.

Using (1.4) with $Z := Q^{21/40}$ we get

$$M_{Q,+}(\mathbf{n}) = \frac{1}{\pi^n} \frac{2}{\pi^*(Q)} \sum_{\substack{Q \leq p \leq 2Q \\ p \equiv 1 \pmod{4}}} \prod_{i=1}^k \left(\sum_{a \leq Z} \chi_p(a) \frac{\sin(2\pi a t_i)}{a} + O(Q^{-1/50}) \right)^{n_i}.$$

We now define $Y := \exp((\log Q)^{1/3})$. Let $L_1 := \lfloor \log Y / \log 2 \rfloor$, $L_2 := \lfloor \log Z / \log 2 \rfloor$, and put $s_{L_1} := Y$, $s_{L_2+1} := Z$, and $s_\ell := 2^\ell$ for $L_1 + 1 \leq \ell \leq L_2$. Then we have

$$(3.6) \quad \max_{t \in [0,1)} \left| \sum_{Y \leq a \leq Z} \chi_p(a) \frac{\sin(2\pi at)}{a} \right| \leq \sum_{L_1 \leq \ell \leq L_2} \max_{t \in [0,1)} \left| \sum_{s_\ell \leq a \leq s_{\ell+1}} \frac{\chi_p(a) e(at)}{a} \right|.$$

Furthermore, it follows from Lemma 3.4 with the choice $B = 6An + 1$ that

$$(3.7) \quad \sum_{L_1 \leq \ell \leq L_2} \max_{t \in [0,1)} \left| \sum_{s_\ell \leq a \leq s_{\ell+1}} \frac{\chi_p(a) e(at)}{a} \right| \ll \sum_{L_1 \leq \ell \leq L_2} \frac{1}{\ell^B} \ll \frac{1}{(\log Q)^{2An}},$$

for all primes p such that $Q \leq p \leq 2Q$ except for a set $\mathcal{E}(Q)$ of size

$$|\mathcal{E}(Q)| \ll_n L_2 Q \exp\left(-\frac{\log Q}{10 \log \log Q}\right) \ll_n Q \exp\left(-\frac{\log Q}{20 \log \log Q}\right).$$

Combining the estimates (3.6) and (3.7) we deduce that $M_{Q,+}(\mathbf{n})$ equals

$$\begin{aligned} & \frac{2}{\pi^n \pi^*(Q)} \sum_{\substack{p \in [Q, 2Q] \setminus \mathcal{E}(Q) \\ p \equiv 1 \pmod{4}}} \prod_{i=1}^k \left(\sum_{a \leq Z} \chi_p(a) \frac{\sin(2\pi at_i)}{a} + O(Q^{-1/50}) \right)^{n_i} \\ & \quad + O_n \left(\exp\left(-\frac{\log Q}{30 \log \log Q}\right) \right) \\ & = \frac{2}{\pi^n \pi^*(Q)} \sum_{\substack{p \in [Q, 2Q] \setminus \mathcal{E}(Q) \\ p \equiv 1 \pmod{4}}} \prod_{i=1}^k \left(\sum_{a \leq Y} \chi_p(a) \frac{\sin(2\pi at_i)}{a} + O\left(\frac{1}{(\log Q)^{2An}}\right) \right)^{n_i} \\ & \quad + O_n \left(\exp\left(-\frac{\log Q}{30 \log \log Q}\right) \right) \\ & = \frac{2}{\pi^n \pi^*(Q)} \sum_{\substack{p \leq Q \\ p \equiv 1 \pmod{4}}} \prod_{i=1}^k \left(\sum_{a \leq Y} \chi_p(a) \frac{\sin(2\pi at_i)}{a} + O\left(\frac{1}{(\log Q)^{2An}}\right) \right)^{n_i} \\ & \quad + O_n \left(\exp\left(-\frac{\log Q}{30 \log \log Q}\right) \right). \end{aligned}$$

Expanding the main term we obtain

$$\begin{aligned} M_{Q,+}(\mathbf{n}) &= \frac{2}{\pi^n \pi^*(Q)} \sum_{\substack{Q \leq p \leq 2Q \\ p \equiv 1 \pmod{4}}} \sum_{a \leq Y^n} \chi_p(a) \frac{\mathcal{B}_{Y,\mathbf{n},\mathbf{t},+}(a)}{a} + O_n \left(\frac{(\log Y)^n}{(\log Q)^{2An}} \right) \\ (3.8) \quad &= \frac{2}{\pi^n} \sum_{a \leq Y^n} \frac{\mathcal{B}_{Y,\mathbf{n},\mathbf{t},+}(a)}{a} \frac{1}{\pi^*(Q)} \sum_{\substack{Q \leq p \leq 2Q \\ p \equiv 1 \pmod{4}}} \chi_p(a) + O_n \left(\frac{1}{(\log Q)^{An}} \right), \end{aligned}$$

where $\mathcal{B}_{Y,\mathbf{n},\mathbf{t},+}(a)$ is defined by

$$\mathcal{B}_{Y,\mathbf{n},\mathbf{t},+}(a) = \sum_{a_1 \cdots a_k = a} \prod_{i=1}^k \sum_{\substack{b_{i,1}, \dots, b_{i,n_i} = a_i \\ b_{i,j} \leq Y}} \prod_{j=1}^{n_i} \sin(2\pi b_{i,j} t_i).$$

We shall first estimate the contribution of the squares $a = m^2$ which gives rise to the main term of $M_{Q,+}(\mathbf{n})$. Indeed, the contribution of these terms to the right hand side

of (3.8) equals

$$\begin{aligned} \frac{1}{\pi^n} \sum_{m \leq Y^{n/2}} \frac{\mathcal{B}_{Y,\mathbf{n},t,+}(m^2)}{m^2} \left(1 + O\left(\exp(-c_0 \sqrt{\log Q})\right) \right) \\ + O\left(\frac{1}{\pi^*(Q)} \sum_{m \leq Y^{n/2}} \frac{|\mathcal{B}_{Y,\mathbf{n},t,+}(m^2)|\omega(m)}{m^2}\right), \end{aligned}$$

for some positive constant c_0 , by the prime number theorem for arithmetic progressions. Let $\varepsilon > 0$ be small. By (2.8) we have $\mathcal{B}_{Y,\mathbf{n},t,+}(m^2), \mathcal{B}_{\mathbf{n},t,+}(m^2) \ll_{\varepsilon,n} m^\varepsilon$. Since $\omega(m) \leq \log m$ and $\mathcal{B}_{Y,\mathbf{n},t,+}(\ell) = \mathcal{B}_{\mathbf{n},t,+}(\ell)$ if $\ell \leq Y$, we deduce that this contribution equals

$$\begin{aligned} (3.9) \quad & \frac{1}{\pi^n} \sum_{m \leq \sqrt{Y}} \frac{\mathcal{B}_{Y,\mathbf{n},t,+}(m^2)}{m^2} + O_n\left(\sum_{\sqrt{Y} < m \leq Y^{n/2}} \frac{1}{m^{2-\varepsilon}} + \exp(-c_0 \sqrt{\log Q})\right) \\ & = \frac{1}{\pi^n} \sum_{m=1}^{\infty} \frac{\mathcal{B}_{\mathbf{n},t,+}(m^2)}{m^2} + O_n\left(\sum_{m > \sqrt{Y}} \frac{1}{m^{2-\varepsilon}} + \exp(-c_0 \sqrt{\log Q})\right) \\ & = M_{\mathbb{X},+}(\mathbf{n}) + O(Y^{-1/3}), \end{aligned}$$

by (2.7). Next, we bound the contribution of the terms a which are not of the form m^2 or $|q_1|m^2$. Since $Y^n \leq \exp(\sqrt{\log Q})$ by our assumptions on Y and Q , it follows from Lemma 3.3 that the contribution of these terms to the main term on the right hand side of (3.8) is

$$\begin{aligned} (3.10) \quad & \ll \exp\left(-\frac{c}{2}\sqrt{\log Q}\right) \sum_{a \leq Y^n} \frac{|\mathcal{B}_{Y,\mathbf{n},t,+}(a)|}{a} \\ & \ll Y^{n/2} \exp\left(-\frac{c}{2}\sqrt{\log Q}\right) \ll \exp\left(-\frac{c}{3}\sqrt{\log Q}\right), \end{aligned}$$

since $\mathcal{B}_{Y,\mathbf{n},t,+}(a) \ll \sqrt{a}$, and where $c > 0$ is the constant in Lemma 3.3.

Finally, we bound the contribution of the terms $a = |q_1|m^2$. Using the bound $\mathcal{B}_{Y,\mathbf{n},t,+}(a) \ll_n a^{1/4}$, and bounding the inner character sum in the right hand side of (3.8) trivially, we deduce that the contribution of these terms is

$$(3.11) \quad \ll \sum_{m \leq Y^{n/2}/\sqrt{|q_1|}} \frac{\mathcal{B}_{Y,\mathbf{n},t,+}(|q_1|m^2)}{|q_1|m^2} \ll_n |q_1|^{-3/4} \sum_{m \geq 1} \frac{1}{m^{3/2}} \ll_{A,n} (\log Q)^{-A},$$

by (3.3). Inserting the estimates (3.9), (3.10), and (3.11) in (3.8) completes the proof. \square

3.2. ESTABLISHING CONVERGENCE IN THE SENSE OF FINITE DISTRIBUTIONS FOR THE SEQUENCE OF “FOURIER COEFFICIENTS” OF THE PATH. — In this section, we describe a different approach to establishing convergence in the sense of finite distributions. Let $C_0([0, 1])$ be the subspace of $C([0, 1])$ consisting of functions vanishing at 0, and let $C_0(\mathbb{Z})$ be the Banach space of complex-valued functions on \mathbb{Z} converging to 0 at infinity with the sup norm. Then note that $f_p \in C_0([0, 1])$ for all primes p . We will follow

the notation in Chapter 6 of Kowalski's book [16]. Let $\text{FT}_p = (\text{FT}_p(h))_{h \in \mathbb{Z}}$ be defined by $\text{FT}_p(0) = 0$ and for $h \neq 0$

$$\text{FT}_p(h) := \int_0^1 f_p(t) e(-ht) dt.$$

By [16, Prop. B.11.8], one can replace Theorem 3.1 (convergence in the sense of finite distributions of the process) in the proof of Theorem 1.2, by the convergence in the sense of finite distributions of the sequence of the Fourier coefficients $(\text{FT}_p)_{Q \leq p \leq 2Q}$ in $C_0(\mathbb{Z})$. Here convergence in the sense of finite distributions of a sequence (X_n) of $C_0(\mathbb{Z})$ -valued random variables to X means that for any $H \geq 1$, the vectors $(X_{n,h})_{|h| \leq H}$ converge in law to $(X_h)_{|h| \leq H}$, in the sense of convergence in law in \mathbb{C}^{2H+1} . More specifically, in our case one needs to compute the moments

$$(3.12) \quad \frac{1}{\pi^*(Q)} \sum_{Q \leq p \leq 2Q} \prod_{0 < |h| \leq H} \text{FT}_p(h)^{j_h},$$

where $\{j_h\}_{0 < |h| \leq H}$ are non-negative integers, since $\text{FT}_p(0) = 0$ for all p .

Let p be a large prime number. It follows from formula (6.5) of [16, Chap. 6] that

$$\text{FT}_p(h) = \frac{1}{2\pi i h} \frac{\sin(\pi h/(p-1))}{\pi h/(p-1)} e\left(-\frac{-h}{2(p-1)}\right) \frac{1}{\sqrt{p}} \sum_{m=1}^{p-1} \left(\frac{m}{p}\right) e\left(-\frac{hm}{p-1}\right),$$

for $h \neq 0$. Since

$$e\left(-\frac{hm}{p-1}\right) = e\left(-\frac{hm}{p}\right) + O_h\left(\frac{1}{p}\right),$$

for fixed h and

$$\sum_{m=1}^{p-1} \left(\frac{m}{p}\right) e\left(-\frac{hm}{p}\right) = \tau\left(\left(\frac{\cdot}{p}\right)\right) \left(\frac{-h}{p}\right),$$

we deduce that

$$\text{FT}_p(h) = \frac{\varepsilon_p}{2\pi i h} \frac{\sin(\pi h/(p-1))}{\pi h/(p-1)} e\left(-\frac{-h}{2(p-1)}\right) \left(\frac{-h}{p}\right) + O_h\left(\frac{1}{\sqrt{p}}\right),$$

where ε_p is defined in (1.6). By using the Taylor expansions of $\sin(\pi h/(p-1))$ and $e(-h/(2(p-1)))$ we deduce that

$$\text{FT}_p(h) = \frac{\varepsilon_p}{2\pi i h} \left(\frac{-h}{p}\right) + O_h\left(\frac{1}{\sqrt{p}}\right).$$

Using this estimate and taking H and $\mathbf{j} = (j_h)_{0 < |h| \leq H}$ fixed we obtain

$$\begin{aligned} \frac{1}{\pi^*(Q)} \sum_{Q \leq p \leq 2Q} \prod_{0 < |h| \leq H} \text{FT}_p(h)^{j_h} &= \prod_{0 < |h| \leq H} \left(\frac{1}{2\pi i h}\right)^{j_h} \frac{1}{\pi^*(Q)} \sum_{\substack{Q \leq p \leq 2Q \\ p \equiv 1 \pmod{4}}} \left(\frac{\prod_{0 < |h| \leq H} (-h)^{j_h}}{p}\right) \\ &+ \prod_{0 < |h| \leq H} \left(\frac{1}{2\pi i h}\right)^{j_h} \frac{1}{\pi^*(Q)} \sum_{\substack{Q \leq p \leq 2Q \\ p \equiv 3 \pmod{4}}} \left(\frac{\prod_{0 < |h| \leq H} (-h)^{j_h}}{p}\right) + O_{H,\mathbf{j}}\left(\frac{1}{\sqrt{Q}}\right). \end{aligned}$$

Therefore, using Lemma 3.3 and the bound (3.3) together with the fact that H and j are fixed, we get

$$\frac{1}{\pi^*(Q)} \sum_{Q \leq p \leq 2Q} \prod_{0 < |h| \leq H} \text{FT}_p(h)^{j_h} = \mathbb{E} \left(\prod_{0 < |h| \leq H} \left(\frac{\mathbb{Y}}{2\pi i h} \mathbb{X}_{-h} \right)^{j_h} \right) + O_{H,j} \left(\exp \left(-\frac{c}{2} \sqrt{\log Q} \right) \right),$$

where c is the constant from Lemma 3.3. This shows that the Fourier coefficients of the Legendre path converge to those of $F_{\mathbb{X}}$ in the sense of finite distributions, as desired.

4. TIGHTNESS OF THE SEQUENCE OF PROCESSES \mathcal{F}_Q

To complete the proof of Theorem 1.2 we need to show that the sequence of processes $(\mathcal{F}_Q)_Q$ is tight. To this end we shall use the following variant of Kolmogorov's tightness criterion.

PROPOSITION 4.1 ([16, Prop. B.11.11]). — *Let $(\mathbb{Y}_n)_{n \geq 1}$ be a sequence of $C([0, 1])$ -valued random variables. Suppose that there exist positive real numbers $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \delta, C$, such that $\beta_2 < \beta_1$ and for any real numbers $0 \leq s < t \leq 1$ and any $n \geq 1$, we have*

$$\begin{cases} \mathbb{E}(|\mathbb{Y}_n(t) - \mathbb{Y}_n(s)|^{\alpha_1}) \leq C|t - s|^{1+\delta} & \text{if } 0 < |t - s| \leq n^{-\beta_1}, \\ \mathbb{E}(|\mathbb{Y}_n(t) - \mathbb{Y}_n(s)|^{\alpha_2}) \leq C|t - s|^{1+\delta} & \text{if } n^{-\beta_1} \leq |t - s| \leq n^{-\beta_2}, \\ \mathbb{E}(|\mathbb{Y}_n(t) - \mathbb{Y}_n(s)|^{\alpha_3}) \leq C|t - s|^{1+\delta} & \text{if } n^{-\beta_2} \leq |t - s| \leq 1. \end{cases}$$

Then $(\mathbb{Y}_n)_{n \geq 1}$ is tight.

Let $0 \leq s < t \leq 1$ be real numbers. To use this criterion for the sequence $(\mathcal{F}_Q)_Q$, we shall bound different moments of $|f_p(t) - f_p(s)|$ (as p varies in $[Q, 2Q]$) in the three ranges $0 < |t - s| \leq Q^{-1}$, $Q^{-1} \leq |t - s| \leq Q^{-2/5}$, and $Q^{-2/5} \leq |t - s| \leq 1$.

4.1. THE “TRIVIAL” RANGE: $0 < |t - s| \leq 1/Q$. — Recall the definition of f_p : the concatenation of points in the time graph of $\frac{1}{\sqrt{p}} \sum_{n \leq pt} \chi_p(n)$. This implies the trivial bound

$$|f_p(t) - f_p(s)| \leq \frac{1}{\sqrt{p}} |pt - ps| = \sqrt{p} |t - s|.$$

Therefore, if $0 < |t - s| \leq 1/Q$ then $|f_p(t) - f_p(s)| \leq |t - s|^{1/2}$ and hence

$$(4.1) \quad \frac{1}{\pi^*(Q)} \sum_{Q \leq p \leq 2Q} |f_p(t) - f_p(s)|^4 \leq |t - s|^2.$$

REMARK 4.2. — By taking a very large moment of $|f_p(t) - f_p(s)|$ we can cover the whole range $0 \leq |t - s| \leq Q^{-1/2-\varepsilon}$.

4.2. THE “BURGESS” RANGE: $Q^{-1} \leq |t - s| \leq Q^{-2/5}$. — In this range we shall bound a very large moment of $|f_p(t) - f_p(s)|$ using the following bound of Burgess for short character sums (see [5, Th. 1] with $r = 2$)

$$(4.2) \quad \sum_{M \leq n \leq M+N} \chi_p(n) \ll N^{1/2} p^{3/16} \log p,$$

which is valid uniformly in N .

LEMMA 4.3. — Let $0 \leq s < t \leq 1$ be real numbers such that $Q^{-1} \leq |t - s| \leq Q^{-2/5}$. Then we have

$$\frac{1}{\pi^*(Q)} \sum_{Q \leq p \leq 2Q} |f_p(t) - f_p(s)|^{1000} \ll |t - s|^2.$$

Proof. — It follows from (4.2) that

$$|f_p(t) - f_p(s)| \leq \frac{1}{\sqrt{p}} \left| \sum_{ps \leq n \leq pt} \chi_p(n) \right| + O\left(\frac{1}{\sqrt{p}}\right) \ll p^{3/16} (\log p) |t - s|^{1/2},$$

since $|t - s| \geq 1/Q$. Therefore, using that $|t - s| \leq Q^{-2/5}$ we obtain

$$\frac{1}{\pi^*(Q)} \sum_{Q \leq p \leq 2Q} |f_p(t) - f_p(s)|^{1000} \ll Q^{375/2} (\log Q)^{1000} |t - s|^{500} \ll |t - s|^2,$$

as desired. \square

REMARK 4.4. — By taking a very large moment of $|f_p(t) - f_p(s)|$ we can cover the whole range $Q^{-1} \leq |t - s| \leq Q^{-3/8-\varepsilon}$ using the Burgess bound (4.2).

4.3. THE RANGE $Q^{-2/5} \leq |t - s| \leq 1$. — This final range is the hardest to deal with. Indeed, a new feature occurs here (and more generally in the range $Q^{-1/2+\varepsilon} \leq |t - s| \leq 1$), caused by the fact that the random variables \mathbb{X}_n are not independent, and $F_{\mathbb{X},p}$ are not sub-Gaussian. In this case we shall use the quadratic large sieve to prove the following key proposition.

PROPOSITION 4.5. — Let $\beta = 1/1000$ and $0 \leq s < t \leq 1$ be real numbers such that $Q^{-2/5} \leq |t - s| \leq 1$. Then we have

$$\frac{1}{\pi^*(Q)} \sum_{Q \leq p \leq 2Q} |f_p(t) - f_p(s)|^4 \ll |t - s|^{1+\beta}.$$

We shall require the following important large sieve inequalities for quadratic characters. The first is a direct consequence of a classical result due to Heath-Brown [13].

LEMMA 4.6. — Let $Q, Y \geq 2$. Then for arbitrary complex numbers a_n and for any $\varepsilon > 0$ we have

$$\sum_{Q \leq p \leq 2Q} \left| \sum_{n \leq Y} a_n \chi_p(n) \right|^2 \ll_{\varepsilon} (QY)^{\varepsilon} (Q + Y) \sum_{\substack{m, n \leq Y \\ mn = \square}} |a_m a_n|.$$

Proof. — This follows from [13, Cor. 2] by embedding the set of primes p such that $Q \leq p \leq 2Q$ in the set of fundamental discriminants d with $|d| \leq 2Q$. \square

As a corollary we deduce the following lemma.

LEMMA 4.7. — *Let $\varepsilon > 0$ be a fixed small real number and $(a_n)_{n \geq 1}$ be an arbitrary bounded sequence of complex numbers. Let Q be large, $1 \leq Y \leq Q$ and $\delta > 0$ be real numbers. The number of primes p such that $Q \leq p \leq 2Q$ and*

$$(4.3) \quad \left| \sum_{Y \leq |n| \leq Q} \frac{a_n \chi_p(n)}{n} \right| > \delta,$$

is $\ll_\varepsilon Q^{1+\varepsilon}/\delta^2 Y$.

Proof. — By Lemma 4.6 we have

$$\sum_{Q \leq p \leq 2Q} \left| \sum_{Y \leq |n| \leq Q} \frac{a_n \chi_p(n)}{n} \right|^2 \ll_\varepsilon Q^{1+\varepsilon/2} \sum_{\substack{Y \leq |n_1|, |n_2| \leq Q \\ |n_1 n_2| = \square}} \left| \frac{a_{n_1} a_{n_2}}{n_1 n_2} \right| \ll_\varepsilon Q^{1+\varepsilon/2} \sum_{n > Y} \frac{d(n^2)}{n^2},$$

by writing $n^2 = |n_1 n_2|$ and using that $(a_n)_{n \geq 1}$ is bounded. Finally, using the trivial bound $d(n^2) \ll_\varepsilon n^{\varepsilon/2}$ we deduce that the number of primes $Q \leq p \leq 2Q$ such that (4.3) holds is

$$\ll_\varepsilon \frac{1}{\delta^2} Q^{1+\varepsilon/2} \sum_{n > Y} \frac{d(n^2)}{n^2} \ll_\varepsilon \frac{1}{\delta^2} Q^{1+\varepsilon/2} \sum_{n > Y} \frac{1}{n^{2-\varepsilon/2}} \ll_\varepsilon \frac{Q^{1+\varepsilon/2}}{\delta^2 Y^{1-\varepsilon/2}} \ll_\varepsilon \frac{Q^{1+\varepsilon}}{\delta^2 Y},$$

as desired. \square

The next lemma is a large sieve inequality for prime discriminants, which is a special case of [25, Lem. 9] (see also [2, Lem. 1]).

LEMMA 4.8 ([25, Lem. 9]). — *Let $\varepsilon > 0$ be a fixed small number. Let x, N be real numbers such that $x \geq 2$ and $2 \leq N \leq x^{1/2-\varepsilon}$. Then for arbitrary complex numbers a_1, \dots, a_N we have*

$$\sum_{p \leq x} \left| \sum_{n \leq N} a_n \chi_p(n) \right|^2 \ll \frac{x}{\log x} \sum_{\substack{m, n \leq N \\ mn = \square}} |a_m a_n|.$$

Proof of Proposition 4.5. — Taking $Z = Q$ in Pólya's Fourier expansion (1.4) gives for all primes p such that $Q \leq p \leq 2Q$ and all $t \in [0, 1]$

$$(4.4) \quad f_p(t) = \frac{\tau(\chi_p)}{2\pi i \sqrt{p}} \sum_{1 \leq |n| \leq Q} \frac{\chi_p(n)(1 - e(-nt))}{n} + O\left(\frac{\log Q}{\sqrt{Q}}\right).$$

Let $0 \leq s < t \leq 1$ be real numbers such that $Q^{-2/5} \leq |t - s| \leq 1$. Using (4.4) together with the easy inequality $|a + b|^4 \leq 2^4(|a|^4 + |b|^4)$, valid for all $a, b \in \mathbb{R}$, we get

$$(4.5) \quad \begin{aligned} & \sum_{Q \leq p \leq 2Q} |f_p(t) - f_p(s)|^4 \\ &= \sum_{Q \leq p \leq 2Q} \left| \frac{\tau(\chi_p)}{2\pi i \sqrt{p}} \sum_{1 \leq |n| \leq Q} \frac{\chi_p(n)(e(-ns) - e(-nt))}{n} + O\left(\frac{\log Q}{\sqrt{Q}}\right) \right|^4 \\ &\ll \sum_{Q \leq p \leq 2Q} \left| \sum_{1 \leq |n| \leq Q} \frac{\chi_p(n)(e(-ns) - e(-nt))}{n} \right|^4 + O\left(\frac{(\log Q)^3}{Q}\right). \end{aligned}$$

Therefore, our goal is to show that

$$\mathcal{J} := \sum_{Q \leq p \leq 2Q} \left| \sum_{1 \leq |n| \leq Q} \frac{\chi_p(n)(e(-ns) - e(-nt))}{n} \right|^4 \ll \pi^*(Q)|t-s|^{1+\beta}.$$

Hence, throughout the proof we will work under the assumption that

$$(4.6) \quad \mathcal{J} \geq \pi^*(Q)|t-s|^{1+\beta},$$

otherwise there is nothing to prove. Let $\varepsilon > 0$ be a small fixed real number. Let $0 < \delta < 1$ and $1 \leq Y \leq Q$ be parameters to be chosen, and define $\mathcal{A}(\delta, Y)$ to be the set of primes $p \in [Q, 2Q]$ such that

$$\left| \sum_{Y \leq |n| \leq Q} \frac{\chi_p(n)(e(-ns) - e(-nt))}{n} \right| > \delta.$$

Then it follows from Lemma 4.7 (with the choice $a_n = e(-ns) - e(-nt)$) that

$$(4.7) \quad |\mathcal{A}(\delta, Y)| \ll_{\varepsilon} \frac{Q^{1+\varepsilon/2}}{\delta^2 Y}.$$

We now write

$$\mathcal{J} := \mathcal{J}_1(\delta, Y) + \mathcal{J}_2(\delta, Y),$$

where $\mathcal{J}_1(\delta, Y)$ denotes the corresponding sum over primes $p \in [Q, 2Q] \setminus \mathcal{A}(\delta, Y)$ and $\mathcal{J}_2(\delta, Y)$ denotes the sum over those in $\mathcal{A}(\delta, Y)$. We start by bounding the latter term. By the Cauchy-Schwarz inequality and the bound (4.7) we derive

$$\begin{aligned} \mathcal{J}_2(\delta, Y) &\leq |\mathcal{A}(\delta, Y)|^{1/2} \left(\sum_{p \in \mathcal{A}(\delta, Y)} \left| \sum_{1 \leq |n| \leq Q} \frac{\chi_p(n)(e(-ns) - e(-nt))}{n} \right|^8 \right)^{1/2} \\ (4.8) \quad &\ll_{\varepsilon} \frac{Q^{1/2+\varepsilon/4}}{\delta Y^{1/2}} \left((\log Q)^4 \sum_{p \in \mathcal{A}(\delta, Y)} \left| \sum_{1 \leq |n| \leq Q} \frac{\chi_p(n)(e(-ns) - e(-nt))}{n} \right|^4 \right)^{1/2} \\ &\ll_{\varepsilon} \frac{Q^{1/2+\varepsilon/2}}{\delta Y^{1/2}} \sqrt{\mathcal{J}} \ll_{\varepsilon} \frac{Q^{\varepsilon}}{\delta Y^{1/2} |t-s|^{(1+\beta)/2}} \mathcal{J}. \end{aligned}$$

by (4.6) and since the inner sum over n is trivially $\ll \log Q$.

We now handle $\mathcal{J}_1(\delta, Y)$. Using our assumption on $\mathcal{A}(\delta, Y)$ together with the inequality $|a+b|^4 \leq 2^4(|a|^4 + |b|^4)$ we obtain

$$\begin{aligned} \mathcal{J}_1(\delta, Y) &= \sum_{p \in [Q, 2Q] \setminus \mathcal{A}(\delta, Y)} \left| \sum_{1 \leq |n| \leq Y} \frac{\chi_p(n)(e(-ns) - e(-nt))}{n} + O(\delta) \right|^4 \\ &\ll \sum_{Q \leq p \leq 2Q} \left| \sum_{1 \leq |n| \leq Y} \frac{\chi_p(n)(e(-ns) - e(-nt))}{n} \right|^4 + \pi^*(Q)\delta^4. \end{aligned}$$

We choose $\delta := |t-s|^{(1+\beta)/4}$. With this choice, the previous estimate and (4.8) become

$$(4.9) \quad \mathcal{J}_1(\delta, Y) \ll \sum_{Q \leq p \leq 2Q} \left| \sum_{1 \leq |n| \leq Y} \frac{\chi_p(n)(e(-ns) - e(-nt))}{n} \right|^4 + \pi^*(Q)|t-s|^{1+\beta},$$

and

$$(4.10) \quad \mathcal{J}_2(\delta, Y) \ll_{\varepsilon} \frac{Q^{\varepsilon}}{Y^{1/2}|t-s|^{3(1+\beta)/4}} \mathcal{J}.$$

To complete the proof we will choose Y in such a way to insure that $\mathcal{J}_2(\delta, Y) \leq \mathcal{J}/2$. This will imply $\mathcal{J} \leq 2\mathcal{J}_1(\delta, Y)$, and hence it will only remain to bound the fourth moment on the right hand side of (4.9), which we shall handle using the quadratic large sieve. The choice of the parameter Y as well as which large sieve inequality to use (Heath-Brown's large sieve (Lemma 4.6), or the Montgomery-Vaughan large sieve (Lemma 4.8)) will depend on the size of $|t-s|$. We will consider two cases:

CASE 1: $Q^{-2/5} \leq |t-s| \leq Q^{-1/6}$. — In this range we use Heath-Brown's large sieve. Indeed, by Lemma 4.6 and Equation (2.4) we obtain

$$(4.11) \quad \sum_{Q \leq p \leq 2Q} \left| \sum_{1 \leq |n| \leq Y} \frac{\chi_p(n)(e(-ns) - e(-nt))}{n} \right|^4 = \sum_{Q \leq p \leq 2Q} \left| \sum_{1 \leq |m| \leq Y^2} \frac{\chi_p(m)g(m)}{m} \right|^2$$

$$(4.12) \quad \ll_{\varepsilon} Q^{\varepsilon/2}(Q+Y^2) \sum_{\substack{1 \leq |n_1|, |n_2| \leq Y^2 \\ n_1 n_2 = \square}} \left| \frac{g(n_1)g(n_2)}{n_1 n_2} \right| \\ \ll_{\varepsilon} Q^{\varepsilon/2}(Q+Y^2)|t-s|^{2-2\varepsilon},$$

where g is defined by (2.1). We now choose $Y = Q^{1/2-\varepsilon}|t-s|^{(\beta-1)/2}$. Then by (4.10) we obtain

$$\mathcal{J}_2(\delta, Y) \ll_{\varepsilon} \frac{Q^{-1/4+2\varepsilon}}{|t-s|^{1/2+\beta}} \mathcal{J} \leq \frac{\mathcal{J}}{2},$$

if ε is suitably small and Q is large enough, by our assumption on $|t-s|$. Therefore, combining the estimates (4.9) and (4.11) we deduce that

$$(4.13) \quad \mathcal{J} \leq 2\mathcal{J}_1(\delta, Y) \ll_{\varepsilon} Q^{\varepsilon/2}(Q+Y^2)|t-s|^{2-2\varepsilon} + \pi^*(Q)|t-s|^{1+\beta} \ll \pi^*(Q)|t-s|^{1+\beta},$$

as desired.

CASE 2: $Q^{-1/6} \leq |t-s| \leq 1$. — In this range we can no longer use Lemma 4.6 (Heath-Brown's quadratic large sieve) to bound the left hand side of (4.11) since the extra factor $Q^{\varepsilon/2}$ is problematic if $|t-s|$ is large (for example if $|t-s| \asymp 1$). In this case we choose $Y = Q^{1/2-\varepsilon}$ and use the large sieve inequality of Montgomery and Vaughan (Lemma 4.8) instead. Indeed, it follows from Lemma 4.8 and Equation (2.4) that

$$(4.14) \quad \sum_{Q \leq p \leq 2Q} \left| \sum_{1 \leq |n| \leq Y} \frac{\chi_p(n)(e(-ns) - e(-nt))}{n} \right|^4 = \sum_{Q \leq p \leq 2Q} \left| \sum_{1 \leq |m| \leq Y^2} \frac{\chi_p(m)g(m)}{m} \right|^2 \\ \ll \pi^*(Q) \sum_{\substack{1 \leq |n_1|, |n_2| \leq Y^2 \\ n_1 n_2 = \square}} \left| \frac{g(n_1)g(n_2)}{n_1 n_2} \right| \\ \ll_{\varepsilon} \pi^*(Q)|t-s|^{2-2\varepsilon}.$$

Choosing ε to be suitably small and inserting this estimate in (4.9) gives

$$\mathcal{J}_1(\delta, Y) \ll \pi^*(Q)|t-s|^{1+\beta}.$$

Finally, using (4.10) and our assumption on $|t-s|$ we derive

$$\mathcal{J}_2(\delta, Y) \ll_{\varepsilon} \frac{Q^{-1/4+2\varepsilon}}{|t-s|^{3(1+\beta)/4}} \mathcal{J} \leq \frac{\mathcal{J}}{2},$$

if ε is suitably small and Q is large enough. Combining these estimates we deduce that

$$\mathcal{J} \leq 2\mathcal{J}_1(\delta, Y) \ll \pi^*(Q)|t-s|^{1+\beta},$$

which completes the proof. \square

REMARK 4.9. — By optimizing this method and allowing the exponent $\beta > 0$ to be very close to 0, we can cover the wider range $Q^{-1/2+\varepsilon} \leq |t-s| \leq 1$.

REFERENCES

- [1] N. C. ANKENY — “The least quadratic non residue”, *Ann. of Math. (2)* **55** (1952), p. 65–72.
- [2] R. C. BAKER & H. L. MONTGOMERY — “Oscillations of quadratic L -functions”, in *Analytic number theory (Allerton Park, IL, 1989)*, Progress in Math., vol. 85, Birkhäuser Boston, Boston, MA, 1990, p. 23–40.
- [3] P. BILLINGSLEY — *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., New York, 1999.
- [4] J. BOBER, L. GOLDMARKER, A. GRANVILLE & D. KOUKOULOPOULOS — “The frequency and the structure of large character sums”, *J. Eur. Math. Soc. (JEMS)* **20** (2018), no. 7, p. 1759–1818.
- [5] D. A. BURGESS — “On character sums and L -series. II”, *Proc. London Math. Soc. (3)* **13** (1963), p. 524–536.
- [6] H. DAVENPORT — *Multiplicative number theory*, third ed., Graduate Texts in Math., vol. 74, Springer-Verlag, New York, 2000.
- [7] H. DAVENPORT & P. ERDŐS — “The distribution of quadratic and higher residues”, *Publ. Math. Debrecen* **2** (1952), p. 252–265.
- [8] S. W. GRAHAM & C. J. RINGROSE — “Lower bounds for least quadratic nonresidues”, in *Analytic number theory (Allerton Park, IL, 1989)*, Progress in Math., vol. 85, Birkhäuser Boston, Boston, MA, 1990, p. 269–309.
- [9] A. GRANVILLE & A. P. MANGEREL — “Three conjectures about character sums”, *Math. Z.* **305** (2023), no. 3, article no. 49 (34 pages).
- [10] A. J. HARPER — “Moments of random multiplicative functions, II: High moments”, *Algebra Number Theory* **13** (2019), no. 10, p. 2277–2321.
- [11] ———, “Moments of random multiplicative functions, I: Low moments, better than square-root cancellation, and critical multiplicative chaos”, *Forum Math. Pi* **8** (2020), article no. e1 (95 pages).
- [12] ———, “A note on character sums over short moving intervals”, 2022, 30 p., [arXiv:2203.09448](https://arxiv.org/abs/2203.09448).
- [13] D. R. HEATH-BROWN — “A mean value estimate for real character sums”, *Acta Arith.* **72** (1995), no. 3, p. 235–275.
- [14] A. HUSSAIN — “The limiting distribution of character sums”, *Internat. Math. Res. Notices* (2022), no. 20, p. 16292–16326.
- [15] A. KALMYNIN — “Long nonnegative sums of Legendre symbols”, 2019, 23 p., [arXiv:1911.10634](https://arxiv.org/abs/1911.10634).
- [16] E. KOWALSKI — *An introduction to probabilistic number theory*, Cambridge Studies in Advanced Math., vol. 192, Cambridge University Press, Cambridge, 2021.
- [17] E. KOWALSKI & W. F. SAWIN — “Kloosterman paths and the shape of exponential sums”, *Compositio Math.* **152** (2016), no. 7, p. 1489–1516.
- [18] Y. LAMZOURI — “Extreme values of $\arg L(1, \chi)$ ”, *Acta Arith.* **146** (2011), no. 4, p. 335–354.

- [19] ———, “The distribution of large quadratic character sums and applications”, 2022, to appear in *Algebra Number Theory*, [arXiv:2212.03227](#).
- [20] D. H. LEHMER – “Incomplete Gauss sums”, *Mathematika* **23** (1976), no. 2, p. 125–135.
- [21] J. H. LOXTON – “The graphs of exponential sums”, *Mathematika* **30** (1983), no. 2, p. 153–163.
- [22] ———, “The distribution of exponential sums”, *Mathematika* **32** (1985), no. 1, p. 16–25.
- [23] D. MILIĆEVIĆ & S. ZHANG – “Distribution of Kloosterman paths to high prime power moduli”, *Trans. Amer. Math. Soc. Ser. B* **10** (2023), p. 636–669.
- [24] H. L. MONTGOMERY – *Topics in multiplicative number theory*, Lect. Notes in Math., vol. 227, Springer-Verlag, Berlin-New York, 1971.
- [25] H. L. MONTGOMERY & R. C. VAUGHAN – “Mean values of character sums”, *Canad. J. Math.* **31** (1979), no. 3, p. 476–487.
- [26] ———, *Multiplicative number theory. I. Classical theory*, Cambridge Studies in Advanced Math., vol. 97, Cambridge University Press, Cambridge, 2007.
- [27] G. RICOTTA & E. ROYER – “Kloosterman paths of prime powers moduli”, *Comment. Math. Helv.* **93** (2018), no. 3, p. 493–532.
- [28] G. RICOTTA, E. ROYER & I. SHPARLINSKI – “Kloosterman paths of prime powers moduli, II”, *Bull. Soc. math. France* **148** (2020), no. 1, p. 173–188.
- [29] K. SCHMÜDGEN – *The moment problem*, Graduate Texts in Math., vol. 277, Springer, Cham, 2017.

Manuscript received 26th April 2023

accepted 2nd May 2024

AYESHA HUSSAIN, Department of Mathematics, University of Exeter
North Park Road, Exeter, EX4 4QF, U.K.

E-mail : a.s.hussain@exeter.ac.uk

Url : <https://ayeshahussainphd.github.io/>

YOUNESS LAMZOURI, Université de Lorraine, CNRS, IECL, and Institut Universitaire de France
F-54000 Nancy, France

E-mail : youness.lamzouri@univ-lorraine.fr

Url : <https://iecl.univ-lorraine.fr/membre-iecl/lamzouri-youness/>