Math. Proc. Camb. Phil. Soc. (2020), **169**, 103–140 © *Cambridge Philosophical Society* 2019
doi:10.1017/S0305004119000100
First published online 22 March 2019

Chebyshev's bias for analytic L-functions

BY LUCILE DEVIN
Department of Mathematics and Statistics, University of Ottawa, 150 *Louis Pasteur pvt, Ottawa, Ontario K*1*N* 6*N*5*, Canada.*
email: ldevin2@uottawa.ca

(*Received* 10 *May* 2018; *revised* 14 *January* 2019)

Abstract

We discuss the generalizations of the concept of Chebyshev's bias from two perspectives. First, we give a general framework for the study of prime number races and Chebyshev's bias attached to general L-functions satisfying natural analytic hypotheses. This extends the cases previously considered by several authors and involving, among others, Dirichlet L-functions and Hasse–Weil L-functions of elliptic curves over $\mathbf{Q}$. This also applies to new Chebyshev's bias phenomena that were beyond the reach of the previously known cases. In addition, we weaken the required hypotheses such as GRH or linear independence properties of zeros of L-functions. In particular, we establish the existence of the logarithmic density of the set $\{x \geq 2 : \sum_{p \leq x} \lambda_f(p) \geq 0\}$ for coefficients $(\lambda_f(p))$ of general L-functions conditionally on a much weaker hypothesis than was previously known.

2010 Mathematics Subject Classification: 11N45, 11F30, 11S40 (Primary); 11G40 (Secondary)

1. *Introduction*

1·1. *Context*

In 1853 Chebyshev noticed in a letter to Fuss that there is a bias in the distribution of primes modulo 4. In initial intervals of the integers, there seems to be more primes congruent to 3 [mod 4] than congruent to 1 [mod 4]. Over the years the synonymous expression "prime number races" has emerged to describe problems of Chebyshev's type. Since Chebyshev's original work, many authors have investigated this problem and provided various generalizations for it.

In [**RS94**] Rubinstein and Sarnak gave a framework for the quantification of Chebyshev's bias in prime number races in arithmetic progressions. Following an observation of Wintner [**Win41**] used for the race between $\pi(x)$ and $\mathrm{Li}(x)$, they studied the logarithmic density of the set $\{x \geq 2 : \pi(x; q, a) > \pi(x; q, b)\}$ where $\pi(x; q, a)$ is the number of primes $\leq x$ that are congruent to a [mod q]. In particular, under the Generalized Riemann Hypothesis (GRH) and assuming the Linear Independence (LI for short) of the zeros, Rubinstein and Sarnak showed that the logarithmic density $\delta(4; 3, 1)$ of the set of $x \geq 2$ for which $\pi(x; 4, 3) > \pi(x; 4, 1)$ exists and is about 0.9959. In their paper about the Shanks–Renyi prime number race [**FM13**], Fiorilli and Martin gave a more precise estimation for $\delta(q; a, b)$

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

in general under the same hypotheses. For more details on prime number races, we refer to the expository article of Granville and Martin [**GM06**], see also [**FK02**].

The method used by Rubinstein and Sarnak is to prove conditionally the existence of a limiting logarithmic distribution for the vector valued function encoding the prime number race. This logarithmic distribution plays a crucial role in the analysis of the bias. It has been generalized greatly to study variants of Chebyshev's bias coming from a broad variety of arithmetic contexts. Let us briefly review some of them.

In his thesis [**Ng00**], Ng generalized this question to that of biases in the distribution of Frobenius substitutions in conjugacy classes of Galois groups of number fields. Here the underlying equidistribution property is Chebotarev's density Theorem. Ng's results are conditional on GRH, LI and Artin's Conjecture for Artin L-functions.

In an expository paper about error terms in arithmetic, Mazur [**Maz08**] sparked interest in prime number races for elliptic curves, or more generally for the Fourier coefficients of a modular form. For example, he plotted graphs of functions:

$$x \longmapsto |\{p \leq x : a_p(E) > 0\}| - |\{p \leq x : a_p(E) < 0\}|,$$

where $a_p(E) = p + 1 - |E(\mathbf{F}_p)|$, for some elliptic curve E defined over $\mathbf{Q}$. He observed that the race between the primes such that $a_p(E) > 0$ and the primes such that $a_p(E) < 0$ tends to be biased towards negative values when the algebraic rank of the elliptic curve is large. In a letter to Mazur [**Sar07**], Sarnak gave an effective framework to answer Mazur's question. Under GRH and LI, he explained the prime number race using the zeros of all the symmetric powers $L(\mathrm{Sym}^n E, s)$ of the Hasse–Weil L-function of $E/\mathbf{Q}$. Sarnak also introduced a related (simpler) race: study the sign of the function $\sum_{p \leq x} a_p(E)/\sqrt{p}$. In [**Fio14a**], Fiorilli developed Sarnak's idea and gave sufficient conditions to get highly biased prime number races in the context of elliptic curves conditionally on weaker versions of GRH and LI.

More recently, Akbary, Ng and Shahabi [**ANS14**] used the theory of almost periodic functions to study the limiting distribution associated with a very wide range of L-functions.

In this paper, we generalize the questions above to prime number races for the coefficients of *analytic L-functions*. We prove unconditionally (Theorem 2·1) the existence of the limiting logarithmic distribution associated with the prime number race for a wide variety of usual L-functions including Dirichlet L-functions and Hasse–Weil L-functions. In particular, we obtain unconditional proofs of some of the results of [**RS94**]. Our general framework is also applicable to new instances of Chebyshev's bias phenomena. For example, we prove unconditionally that, after suitable scaling, the functions

$$x \longmapsto \sum_{p=a^2+4b^2 \leq x} \frac{a^2-4b^2}{a^2+4b^2}, \quad x \longmapsto \sum_{p=a^2+b^2 \leq x} \frac{a^4+b^4-6a^2b^2}{(a^2+b^2)^2}, \quad x \longmapsto \sum_{p \leq x} \frac{a_p(E_i)a_p(E_j)}{p}$$

admit limiting logarithmic distributions with negative average value (see Theorem 3·6, 3·7 and 3·8 for precise statements). For the first function, the fact that the average value is negative gives evidence (see Corollary 2·4) that when writing $p = a^2 + 4b^2$ there is a bias: the even square tends to be more often larger than the odd square.

We also study minimal conditions to ensure that the distribution has nice properties such as regularity, symmetry, and concentration. We obtain results comparable to [**RS94**] under weaker hypotheses. In particular, we prove the existence of the logarithmic density $\delta(4; 3, 1)$

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

conditionally on a weaker version of LI related to the notion of self-sufficient zeros introduced by Martin and Ng [**MN17**], (see Theorem 2·2). We also highlight a relation between the support of the distribution and Riemann Hypothesis (see Theorem 2·5).

1·2. *Setting*

In this paper, we use a custom-made definition of "analytic" L-function inspired by [**IK04**, chapter 5] and Selberg's class. We will only use analytic properties of the function to study the associated prime number race.

Definition 1·1 (Analytic L-function). Let $L(f, s)$ be a complex-valued function of the variable $s \in \mathbf{C}$ attached to an auxiliary parameter f to which one can attach an integer $q(f)$ (usually f is of arithmetic origin and $q(f)$ is its conductor). We say that $L(f, s)$ is an analytic L-function if we have the following data and conditions:

(i) a Dirichlet series factorising as an Euler product of degree $d \geq 1$ that coincides with $L(f, s)$ for $\mathrm{Re}(s) > 1$:

$$L(f, s) = \sum_{n \geq 1} \lambda_f(n) n^{-s} = \prod_p \prod_{j=1}^{d} \left(1 - \alpha_{f,j}(p) p^{-s}\right)^{-1}$$

with $\lambda_f(1) = 1$ and $\alpha_{f,j}(p) \in \mathbf{C}$, satisfying $|\alpha_{f,j}(p)| = 1$ for all j and $p \nmid q(f)$. In particular the series and Euler product are absolutely convergent for $\mathrm{Re}(s) > 1$;

(ii) a gamma factor with local parameters $\kappa_j \in \mathbf{C}$, $\mathrm{Re}(\kappa_j) > -1$:

$$\gamma(f, s) = \pi^{-ds/2} \prod_{j=1}^{d} \Gamma\left(\frac{s + \kappa_j}{2}\right).$$

The *analytic conductor* of f is then defined by:

$$\mathfrak{q}(f) = q(f) \prod_{j=1}^{d} \left(|\kappa_j| + 3\right),$$

and we can define the completed L-function

$$\Lambda(f, s) = q(f)^{s/2} \gamma(f, s) L(f, s).$$

It admits an analytic continuation to a meromorphic function of order 1, with at most poles at $s = 0$ and $s = 1$. Moreover it satisfies a functional equation $\Lambda(f, s) = \epsilon(f)\Lambda(\overline{f}, 1 - s)$, with $|\epsilon(f)| = 1$. Here $\Lambda(\overline{f}, 1 - s)$ is the completed L-function associated to $L(\overline{f}, s) := \sum_{n \geq 1} \overline{\lambda_f(n)} n^{-s}$;

(iii) the *second moment L-function*

$$L(f^{(2)}, s) = \prod_p \prod_{j=1}^{d} \left(1 - \alpha_j(p)^2 p^{-s}\right)^{-1}$$

is defined for $\mathrm{Re}(s) > 1$. We assume that there exists $c > 0$ and an open subset $U \supset \{\mathrm{Re}(s) \geq 1 - c/(\log(|\mathrm{Im}(s)| + 3))\}$, such that $L(f^{(2)}, \cdot)$ can be continued to a meromorphic function for $s \in U$, and on $U - \{1\}$ there is neither a zero nor a

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

pole of $L(f^{(2)}, \cdot)$. Moreover, we assume that there exists $C > 0$ such that for all $s \in U$, $\mathrm{Re}(s) \leq 2$ one has

$$\left|\frac{L'(f^{(2)}, s)}{L(f^{(2)}, s)}\right| \leq C \log^2(|\mathrm{Im}(s)| + 3).$$

Remark 1.

(i) [Order of a meromorphic function]. A meromorphic function on $\mathbf{C}$ is said to be of order 1 if it can be written as the quotient of two entire functions of order 1 (that is functions Λ such that for every $\beta > 1$ and no $\beta < 1$ one has $|\Lambda(s)| \ll \exp(|s|^\beta)$). Usually this property is obtained by proving that the function is bounded in vertical strips. This occurs naturally in the proof of the functional equation using the method of zeta-integrals (see e.g. [**GJ72**, Corollary 13·8, Proposition 13·9] for the general case of automorphic L-functions).
In particular, by Jensen's formula (see [**Rud80**, 15·20]) one deduce that the sum over the zeros $\sum_{\Lambda(\rho)=0} 1/|\rho|^{1+\epsilon}$ converges for every $\epsilon > 0$.

(ii) [Second moment]. In [**Con05**], the function $L(f^{(2)}, s)$ used in Definition 1·1(iii) is called the *second moment* of $L(f, s)$ over $\mathbf{Q}$. We note that it is determined by the local roots $\alpha_j(p)$ over $\mathbf{Q}$ (rather than by the meromorphic function $L(f, s)$) and it is related to the Rankin–Selberg product (see Example 1(iv)). The assumption on the function $L(f^{(2)}, s)$ is a stronger version of the *second moment hypothesis* ([**Con05**, definition 4·4]), we assume a classical zero-free region.

Example 1.

(i) Riemman's zeta function is a classical example of L-function, it satisfies all the conditions of Definition 1·1, see Theorem 3·1.

(ii) Dirichlet L-functions are analytic L-functions. Hence some of the results of [**RS94**] can be given unconditionally, see Theorem 3·2.

(iii) Modular L-functions of degree 2 are analytic L-functions. It is a consequence of results of Deligne and Serre [**Del74**, Theorem 8·2] and [**DS74**]. In particular, following results on modularity ([**Wil95**, **TW95**, **BCDT01**]), if $E/\mathbf{Q}$ is an elliptic curve, $L(E, s)$ is an analytic L-function. This property was already used by Fiorilli in [**Fio14a**], see Proposition 3·5.

(iv) Under the Ramanujan–Petersson Conjecture, general automorphic L-functions associated to cusp forms on $GL(m)$ for $m \geq 1$, are analytic L-functions in the sense of Definition 1·1. Indeed (i) precisely says that the Ramanujan–Petersson Conjecture is satisfied and (ii) is known for such L-functions ([**GJ72**], [**Cog04**]). One has

$$L(f^{(2)}, s) = L(\mathrm{Sym}^2 f, s)L(\wedge^2 f, s)^{-1} = L(f \otimes f, s)L(\wedge^2 f, s)^{-2}.$$

By [**BG92**, Theorem 6·1, Theorem 7·5] for $L(\mathrm{Sym}^2 f, s)$ and [**BF90**, Theorems 1-3], [**JS90**, Theorems 1-2] for $L(\wedge^2 f, s)$, (iii) is satisfied. As F. Brumley pointed out to us one could ask for the third hypothesis above to be about any two of the three functions $L(f \otimes f, s)$, $L(\mathrm{Sym}^2 f, s)$ and $L(\wedge^2 f, s)$.
In [**ANS14**, Corollary 1·5], the existence of the limiting logarithmic distribution of the function $\psi(f, x)/\sqrt{x}$ associated to an automorphic L-function $L(f, s)$ is

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

proved under GRH and does not depend on the Ramanujan–Petersson Conjecture. In the present paper, we need to assume the Ramanujan–Petersson Conjecture but not GRH to prove that the function $\pi(f,x)\log x/x^{\beta_{f,0}}$ has a limiting logarithmic distribution (see Theorem 3·4).

(v) If $L(f,s)$ and $L(g,s)$ are two modular L-functions of degree 2, such that $g \neq \overline{f}$, then the Rankin–Selberg product $L(f\otimes g,s)$ is an analytic L-function. The conditions (i) and (ii) are satisfied (see e.g. [**CPS04**, Theorem 2·3]). One has

$$L\big((f\otimes g)^{(2)},s\big)=\frac{L\big(\mathrm{Sym}^2 f\otimes \mathrm{Sym}^2 g,s\big)L(\chi_f\chi_g,s)}{L\big((\mathrm{Sym}^2 f)\otimes\chi_g,s\big)L\big((\mathrm{Sym}^2 g)\otimes\chi_f,s\big)},$$

where χ_f, χ_g are the nebentypus respectively associated to the modular forms f and g. We deduce that Definition 1·1(iii) is satisfied. We use this property in Theorem 3·8 in the case f and g are associated to two elliptic curves over $\mathbf{Q}$ that are non-isogeneous in a strong sense.

The parameters in Definition 1·1 satisfy for every prime p, $\lambda_f(p)=\sum_{j=1}^{d}\alpha_{f,j}(p)$. In case $L(f,\cdot)=L(\overline{f},\cdot)$ is real, one has $\lambda_f(p)\in\mathbf{R}$. The Generalized Sato–Tate conjecture states that in this case the $(\lambda_f(p))_p$ should equidistribute according to a certain probability measure on $[-d,d]$. In case the L-function is entire and does not vanish on the line $\mathrm{Re}(s)=1$, a general Prime Number Theorem for the L-function implies that the Sato–Tate law has a mean value equal to 0. We expect this to hold more generally.

CONJECTURE 1·2. *Let $\mathcal{S}$ be a finite set of entire analytic L-functions. Suppose $\mathcal{S}$ is stable by conjugation (i.e. $\overline{\mathcal{S}}=\mathcal{S}$), and let $(a_f)_{f\in\mathcal{S}}$ be a set of complex numbers satisfying $a_{\overline{f}}=\overline{a_f}$. Then the sequence $\big(\sum_{f\in\mathcal{S}}a_f\lambda_f(p)\big)_p$ is equidistributed in an interval of $\mathbf{R}$ according to a Sato–Tate law with mean value* 0.

In particular for L-functions associated with an elliptic curve over $\mathbf{Q}$, Conjecture 1·2 is a (very) weak version of the Sato–Tate conjecture, and is known to hold ([**CHT08**], [**HSBT10**]).

1·3. *Chebyshev type questions for analytic L-functions*

The Chebyshev type question this paper primarily focuses on is the following. Let $\mathcal{S}$ be a finite set of entire analytic L-functions such that $\overline{\mathcal{S}}=\mathcal{S}$, and $(a_f)_{f\in\mathcal{S}}$ a set of complex numbers satisfying $a_{\overline{f}}=\overline{a_f}$. Under Conjecture 1·2 one has

$$\frac{1}{\pi(x)}\sum_{p\leq x}\sum_{f\in\mathcal{S}}a_f\lambda_f(p)\longrightarrow 0$$

as $x\to+\infty$. It is natural to study the sign of the summatory function

$$x\longmapsto\sum_{p\leq x}\sum_{f\in\mathcal{S}}a_f\lambda_f(p) \tag{1·1}$$

for $x>0$.

Remark 2. In the case of elliptic curve L-functions (where the Sato–Tate law is known to be symmetric), Mazur ([**Maz08**]) was first interested in studying the function

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

$$x \mapsto \left|\left\{p \leq x : \sum_{f\in\mathcal{S}} a_f \lambda_f(p) > 0\right\}\right| - \left|\left\{p \leq x : \sum_{f\in\mathcal{S}} a_f \lambda_f(p) < 0\right\}\right|$$

for $x > 0$. But as Sarnak showed in [**Sar07**], for this study, we need information about all symmetric powers of the L-functions involved. It may yield non-converging infinite sums. Hence, following Sarnak's idea, we focus on the summatory function as in (1·1).

More generally, if $L(f, \cdot)$ has a pole of order r_f at $s = 1$, we study the sign of the summatory function

$$S : x \longmapsto \sum_{p\leq x} \sum_{f\in\mathcal{S}} a_f \lambda_f(p) - \sum_{f\in\mathcal{S}} a_f r_f \operatorname{Li}(x).$$

In [**Sar07**], Sarnak presents a method to deal with this question in the case $\mathcal{S}$ is a singleton. Precisely, he considers the cases of the L-function associated with Ramanujan's τ function and of L-functions associated with an elliptic curve over $\mathbf{Q}$.

Building on this method, we wish to understand the set of x for which $S(x) \geq 0$. In [**Kac95**], Kaczorowski showed that in certain situations, under the Generalized Riemann Hypothesis, the natural density does not exist. We use the logarithmic density to measure this set.

Definition 1·3.

(i) Define

$$\overline{\delta}(\mathcal{S}) = \limsup \frac{1}{Y}\int_2^Y \mathbf{1}_{\geq 0}(S(e^y))\,\mathrm{d}y \quad \text{and} \quad \underline{\delta}(\mathcal{S}) = \liminf \frac{1}{Y}\int_2^Y \mathbf{1}_{\geq 0}(S(e^y))\,\mathrm{d}y,$$

where $\mathbf{1}_{\geq 0}$ is the Heaviside step function, indicator function of the set of non-negative real numbers. If these two densities are equal, we denote $\delta(\mathcal{S})$ their common value. These quantities measure the bias of $S(x)$ towards positive values.

(ii) If $\delta(\mathcal{S})$ exists and is $> 1/2$ we say that there is a bias towards positive values. If it is $< 1/2$ we say that there is a bias towards negative values.

Under GRH and LI, Sarnak ([**Sar07**]) showed that for an L-function of degree 2, the bias exists and always differs from $1/2$. One of the main results of this article is that the bias exists without assuming GRH and under a hypothesis weaker than LI on the independence of the zeros of the L-functions involved (see Theorem 2·2). To state the existence of the bias we prove that a suitable normalization of the function $S(x)$ admits a limiting logarithmic distribution.

Definition 1·4. Let $F : \mathbf{R} \to \mathbf{R}$ be a real function, we say that F admits a limiting logarithmic distribution if there exists a probability measure μ on the Borel sets in $\mathbf{R}$ such that for any bounded Lipschitz continuous function g, we have

$$\lim_{Y\to\infty} \frac{1}{Y}\int_2^Y g(F(e^y))\,\mathrm{d}y = \int_{\mathbf{R}} g(t)\,\mathrm{d}\mu(t).$$

Before stating our main result, let us set some notations. Since we do not assume the Riemann Hypothesis for our L-functions, we denote

$$\beta_{\mathcal{S},0} = \sup\{\operatorname{Re}(\rho) : \exists f \in \mathcal{S},\, L(f, \rho) = 0\}.$$

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

One has $\beta_{\mathcal{S},0} \geq 1/2$ and equality is equivalent to the Riemann Hypothesis for all the L-functions $L(f, \cdot)$, $f \in \mathcal{S}$. We use the Dirac delta $\delta_{\beta_{\mathcal{S},0}=1/2}$ as an indicator of the truth of the Riemann Hypothesis for all the L-functions $L(f, \cdot)$, $f \in \mathcal{S}$:

$$\delta_{\beta_{\mathcal{S},0}=1/2} = \begin{cases} 1 \text{ if } \beta_{\mathcal{S},0} = 1/2, \\ 0 \text{ otherwise.} \end{cases}$$

For L a meromorphic function in a neighbourhood of a point $\rho \in \mathbf{C}$, let $m(L, \rho)$ be the multiplicity of the zero of L at $s = \rho$. (One has $m(L, \rho) = 0$ if $L(\rho) \neq 0$, $m(L, \rho) > 0$ if $L(\rho) = 0$ and $m(L, \rho) < 0$ if L has a pole at $s = \rho$.) We denote

$$M_{\mathcal{S}}(\gamma) := \sum_{f \in \mathcal{S}} a_f m\big(L(f, \cdot), \beta_{\mathcal{S},0} + i\gamma\big).$$

Define

$$\mathcal{Z}_{\mathcal{S}} = \big\{\gamma > 0 : M_{\mathcal{S}}(\gamma) \neq 0\big\}, \qquad \mathcal{Z}_{\mathcal{S}}(T) = \mathcal{Z}_{\mathcal{S}} \cap (0, T],$$

seen as multi-sets of zeros of largest real part (i.e. we count the zeros with multiplicities). We denote $\mathcal{Z}_{\mathcal{S}}^*$ and $\mathcal{Z}_{\mathcal{S}}^*(T)$ the corresponding sets (i.e. repetitions are not allowed). Note that these sets can be empty if $\beta_{\mathcal{S},0} > 1/2$, this will not be a problem. Finally, we denote $N_{\mathcal{S}}(T) = |\mathcal{Z}_{\mathcal{S}}^*(T)|$ the number of distinct zeros of maximal real part and imaginary part bounded by T.

If it does not lead to confusion, we may omit the subscript $\mathcal{S}$. In the case the set $\mathcal{S}$ is a singleton $\{f\}$ and $a_f = 1$, we will write f in subscript instead of $\{f\}$.

2. *Statement of the theoretical results*

2·1. *Limiting distribution*

Our first result is the existence of the limiting logarithmic distribution for the prime number races associated to a finite set of analytic L-functions.

THEOREM 2·1. *Let $\{L(f, \cdot) : f \in \mathcal{S}\}$ be a finite set of analytic L-functions such that $\overline{\mathcal{S}} = \mathcal{S}$, and $(a_f)_{f \in \mathcal{S}}$ a set of complex numbers satisfying $a_{\overline{f}} = \overline{a_f}$. Define*

$$E_{\mathcal{S}}(x) = \frac{\log x}{x^{\beta_{\mathcal{S},0}}} \left(\sum_{p \leq x} \sum_{f \in \mathcal{S}} a_f \lambda_f(p) + \sum_{f \in \mathcal{S}} a_f m\big(L(f, \cdot), 1\big) \operatorname{Li}(x) \right).$$

The function $E_{\mathcal{S}}(x)$ admits a limiting logarithmic distribution $\mu_{\mathcal{S}}$. There exists a positive constant C (depending on $\mathcal{S}$) such that one has

$$\mu_{\mathcal{S}}(\mathbf{R} - [-A, A]) \ll \exp(-C\sqrt{A}).$$

Moreover let $X_{\mathcal{S}}$ be a random variable of law $\mu_{\mathcal{S}}$, then the expected value of $X_{\mathcal{S}}$ is

$$\mathbb{E}(X_{\mathcal{S}}) = m_{\mathcal{S}} := \sum_{f \in \mathcal{S}} a_f \big(m(L(f^{(2)}, \cdot), 1)\delta_{\beta_{\mathcal{S},0}=1/2} - \beta_{\mathcal{S},0}^{-1} m\big(L(f, \cdot), \beta_{\mathcal{S},0}\big)\big),$$

and its variance is

$$Var(X_{\mathcal{S}}) = 2 \sum_{\gamma \in \mathcal{Z}_{\mathcal{S}}^*} \frac{|M(\gamma)|^2}{\beta_{\mathcal{S},0}^2 + \gamma^2},$$

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

where

$$\delta_{\beta_{\mathcal{S},0}=1/2} = \begin{cases} 1 \text{ if } \beta_{\mathcal{S},0} = 1/2, \\ 0 \text{ otherwise,} \end{cases}$$

and for γ *in* $\mathcal{Z}_{\mathcal{S}}^*$, $M(\gamma) = \sum_{f\in\mathcal{S}} a_f m(L(f,\cdot), \beta_{\mathcal{S},0} + i\gamma)$.

Remark 3.

(i) This result generalizes [**RS94**, Theorem 1·1] and [**Ng00**, Theorem 5·1·2] which are conditional on GRH and respectively deal with the cases of sets of Dirichlet L-functions and sets of Artin L-functions under Artin's Conjecture. Similar results are obtained under GRH in [**ANS14**] for the prime number race corresponding to the function ψ associated to general L-functions. An unconditional proof is given in [**Fio14a**] in the case $\mathcal{S}$ is a singleton composed of one Hasse–Weil L-function. The proof of Theorem 2·1, in section 4, is essentially an adaptation of Fiorilli's proof to more general L-functions.

(ii) Note that there is no assumption made about the set $\mathcal{Z}_{\mathcal{S}}$. In particular the limiting logarithmic distribution exists even if the set $\mathcal{Z}_{\mathcal{S}}$ is empty (see Remark 7(i)).

(iii) The sign of the expected value $\mathbb{E}(X_{\mathcal{S}})$ gives an idea of the kind of bias one should expect. When it is non-zero, we conjecture that the bias is imposed by the sign of the expected value. We prove this conjecture conditionally on additional hypotheses (see Corollary 2·4).

(iv) More precisely Theorem 2·1 states that "on average" the coefficients $\lambda_f(p)$ of an entire analytic L-function are equal to $(\beta_{f,0} m\big(L(f^{(2)},\cdot), 1\big)\delta_{\beta_{f,0}=1/2} - m(L(f,\cdot),\beta_{f,0}))/p^{1-\beta_{f,0}}$. Under GRH the bias is due to the second moment function $L(f^{(2)}, s)$. A sum over squares of primes appears and cannot be considered as an error term if the function $L(f^{(2)}, s)$ admits a zero or a pole at $s = 1$ (see Section 4·3). As A. Granville pointed out to us, this phenomenon is related to another kind of bias in the Birch and Swinnerton-Dyer conjecture. Using this second moment function in the case of elliptic curves over $\mathbf{Q}$, Goldfeld [**Gol82**] showed that there is an unexpected factor $\sqrt{2}$ appearing in the asymptotics for partial Euler products for the L-function at the central point. This is due to the fact that in the case of a non-CM elliptic curve over $\mathbf{Q}$, the corresponding second moment function has a zero of order 1 at $s = 1$ (see also Proposition 3·5). Goldfeld's result has been generalized by Conrad [**Con05**, Theorem 1·2] to general L-functions (quite similar to our analytic L-functions of Definition 1·1).

2·2. *Further properties under extra hypotheses*

Under additional hypotheses over the zeros of the L-functions, we deduce properties of $\mu_{\mathcal{S}}$, and in turn results on the bias. This idea is developed in the following results. A standard hypothesis about the set $\mathcal{Z}_{\mathcal{S}}$ is the Linear Independence hypothesis (LI), we show under a weaker hypothesis about linear independence that the logarithmic density $\delta(\mathcal{S})$ (see Definition 1·3) exists.

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

THEOREM 2·2. *Suppose that there exists $n \geq 1$ and $\lambda_1, \dots, \lambda_n \in \mathcal{Z}_{\mathcal{S}}^*$ such that*

$$\langle \lambda_1, \dots, \lambda_n \rangle_{\mathbf{Q}} \cap \left\langle \mathcal{Z}_{\mathcal{S}}^* - \{\lambda_1, \dots, \lambda_n\} \right\rangle_{\mathbf{Q}} = \{0\},$$

where $\langle \cdot \rangle_{\mathbf{Q}}$ denotes the $\mathbf{Q}$-span of a set of real numbers. Then the distribution $\mu_{\mathcal{S}}$ is continuous (i.e. $\mu_{\mathcal{S}}$ assigns zero mass to finite sets), and $\delta(\mathcal{S})$ exists.

This theorem is proved in Section 5·1. In [**RS94**], [**Ng00**], [**ANS14**] and [**Fio14a**] the corresponding result is obtained under LI i.e. assuming that $\mathcal{Z}_{\mathcal{S}}$ is linearly independent over $\mathbf{Q}$. Using the theory of almost periodic functions, Kaczorowski and Ramaré [**KR03**, Theorem 3] prove the existence of the logarithmic density of a comparable set in the general setting of the Selberg class, under the Riemann Hypothesis only. We prove other results related to the smoothness of the limiting distribution $\mu_{\mathcal{S}}$ in Section 5·1 where we use the concept of self-sufficient zeros introduced by Martin and Ng in [**MN17**] (see Definition 5·3).

We are also interested in the symmetry of the distribution $\mu_{\mathcal{S}}$. We prove the following result conditionally on a weak conjecture of linear independence of the zeros.

THEOREM 2·3. *Suppose that the set $\mathcal{S}$ satisfies the conditions of Theorem 2·1. Suppose that for every $(k_\gamma)_\gamma \in \mathbf{Z}^{(\mathcal{Z}_{\mathcal{S}})}$ one has*

$$\sum_{\gamma \in \mathcal{Z}_{\mathcal{S}}} k_\gamma \gamma = 0 \Rightarrow \sum_{\gamma \in \mathcal{Z}_{\mathcal{S}}} k_\gamma \equiv 0 \,[\mathrm{mod}\ 2].$$

Then the distribution $\mu_{\mathcal{S}}$ is symmetric with respect to $m_{\mathcal{S}}$.

We prove Theorem 2·3 in Section 5·2. This theorem improves again a result obtained in [**RS94**] under LI.

We can now come back to Remark 3(iii). If the bias exists, it should be imposed by the sign of the average value of the limiting distribution. We get the following result as a Corollary of Theorems 2·1, 2·2 and Theorem 2·3 or Chebyshev's inequality (Lemma 5·7).

COROLLARY 2·4. *Let $\{L(f, \cdot) : f \in \mathcal{S}\}$ be a finite set of analytic L-functions such that $\overline{\mathcal{S}} = \mathcal{S}$, and $(a_f)_{f \in \mathcal{S}}$ a set of complex numbers satisfying $a_{\overline{f}} = \overline{a_f}$. Suppose that*:

(i) *there exists $n \geq 1$ and $\lambda_1, \dots, \lambda_n \in \mathcal{Z}_{\mathcal{S}}^*$ such that*

$$\langle \lambda_1, \dots, \lambda_n \rangle_{\mathbf{Q}} \cap \left\langle \mathcal{Z}_{\mathcal{S}}^* - \{\lambda_1, \dots, \lambda_n\} \right\rangle_{\mathbf{Q}} = \{0\},$$

and

(ii)(a) *for every $(k_\gamma)_\gamma \in \mathbf{Z}^{(\mathcal{Z}_{\mathcal{S}})}$ one has $\sum_{\gamma \in \mathcal{Z}_{\mathcal{S}}} k_\gamma \gamma = 0 \Rightarrow \sum_{\gamma \in \mathcal{Z}_{\mathcal{S}}} k_\gamma \equiv 0 \,[\mathrm{mod}\ 2]$,*

or

(ii)(b) *one has $(2/m_{\mathcal{S}}^2) \sum_{\gamma \in \mathcal{Z}_{\mathcal{S}}^*} |M(\gamma)|^2/(\beta_{\mathcal{S},0}^2 + \gamma^2) < 1/2$.*

Then $\delta(\mathcal{S})$ exists and $(\delta(\mathcal{S}) - 1/2)\, m_{\mathcal{S}} \geq 0$. In particular, the direction of the bias is given by the sign of the mean value $m_{\mathcal{S}}$.

We have avoided so far the use of the Riemann Hypothesis, weakening the hypotheses made in previous works. We generalize Rubinstein and Sarnak result [**RS94**, Theorem 1·2] by stating a dichotomy depending on the validity of the Riemann Hypothesis.

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

THEOREM 2·5. *Suppose that the set $\mathcal{S}$ satisfies the conditions of Theorem 2·1.*

(i) Suppose the Riemann Hypothesis is satisfied for every $L(f,s)$, $f \in \mathcal{S}$ (i.e. $\beta_{\mathcal{S},0} = 1/2$). Suppose also that for every $f \in \mathcal{S}$, one has $\mathrm{Re}(a_f) \geq 0$, and that there exists $f \in \mathcal{S}$ such that $\mathrm{Re}(a_f) > 0$. Then there exists a constant c depending on $\mathcal{S}$ such that

$$\mu_{\mathcal{S}}(\mathbf{R} - [-A, A]) \gg \exp(-\exp(cA)).$$

In particular $0 < \underline{\delta}(\mathcal{S}) \leq \overline{\delta}(\mathcal{S}) < 1$.

(ii) Suppose $\beta_{\mathcal{S},0} > 1/2$, and that the sum $\sum_{\gamma \in \mathcal{Z}_{\mathcal{S}}} 1/|\beta_{\mathcal{S},0} + i\gamma|$ converges. Then $\mu_{\mathcal{S}}$ has compact support.

This result is proved in Section 5·3.

Remark 4.

(i) The proof of Theorem 2·5(i) is an adaptation to a more general case of the proof given by Rubinstein and Sarnak. The hypothesis about $\mathrm{Re}(a_f) \geq 0$ does not seem very natural, but is necessary in our analysis. It is satisfied in the case the set $\mathcal{S}$ is a singleton i.e. in most of the examples presented in Section 3. In the case of the prime number race between congruence classes (Theorem 3·2), the condition holds when one studies the race between 1 [mod q] and another invertible class modulo q.

(ii) One can also show that the race is inclusive — i.e. that each contestant leads the race infinitely many times (this is implied by $0 < \underline{\delta}(\mathcal{S}) \leq \overline{\delta}(\mathcal{S}) < 1$) — assuming GRH and LI and nothing on the a_f's (see [**RS94**]). In [**MN17**], Martin and Ng prove that the race is inclusive assuming GRH and a weaker hypothesis than LI based on self-sufficient zeros (see Definition 5·3).

(iii) The hypothesis in Theorem 2·5(ii) is a weak Zero Density hypothesis, we assume that there are not too many zeros off the critical line. There are results supporting this hypothesis, see for example [**IK04**, chapter 10] for the Riemann zeta function and Dirichlet L-functions. More generally, Kaczorowski and Perelli ([**KP03**, lemma 3]) proved a stronger version of this hypothesis in the case of a Selberg class L-function of degree d with $\beta_{f,0} \geq 1 - 1/4(d+3)$.

3. *Applications to old and new instances of prime number races*

In this section we present two kinds of applications. First we find some results of the literature as special cases of our general result. Most of them were conditional on GRH, they are now unconditional. In a second part, we use the fact that analytic L-functions describe a wide range of L-functions to present new applications of Chebyshev's races.

3·1. *Proofs of old results under weaker assumptions*

3·1·1. *Sign of the second term in the Prime Number Theorem*

The first example of analytic L-function is Riemann's zeta function (see Example 1(i)). Adapting Theorem 2·1 to ζ yields an unconditional proof of the existence of the logarithmic limiting distribution for the race $\pi(x)$ versus $\mathrm{Li}(x)$ (see e.g. [**Win41**], [**RS94**, p. 175] for previous results under RH).

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

THEOREM 3·1. *With the notations as in Section 1·2, the function*

$$E_\zeta(x) = \frac{\log x}{x^{\beta_{\zeta,0}}}\big(\pi(x) - \mathrm{Li}(x)\big)$$

has a limiting logarithmic distribution μ_ζ *on* $\mathbf{R}$. *There exists a positive constant* C *such that one has*

$$\mu_\zeta\big(\mathbf{R} - [-A, A]\big) \ll \exp\big(-C\sqrt{A}\big).$$

Moreover, the expected value of μ_ζ *is* $m_\zeta = -\delta_{\beta_{\zeta,0}=\frac{1}{2}}$.

Proof. This follows from the fact that the function with squared local roots associated with ζ is ζ itself, moreover ζ has a pole of order 1 at $s = 1$ and does not vanish over $(0, 1)$.

We deduce the already known idea (e.g. under the conditions of Corollary 2·4) that assuming RH, the race $\pi(x) - \mathrm{Li}(x)$ should be biased towards negative values. Conversely, a bias towards negative values in the race $\pi(x) - \mathrm{Li}(x)$ gives evidence that RH should hold.

3·1·2. *Prime number races between congruence classes modulo an integer*

The results of Rubinstein and Sarnak [**RS94**, Theorem 1·1, Theorem 1·2] in the case of a prime number race with only two contestants a and b modulo q is a particular case of Theorem 2·1. Indeed take

$$\mathcal{S} = \big\{L(\chi, \cdot) : \chi \bmod q, \chi \neq \chi_0\big\}$$

and $a_\chi = \overline{\chi}(a) - \overline{\chi}(b)$. We obtain an uncondtional (i.e. without assuming GRH) proof of [**RS94**, Theorem 1·1, Theorem 1·2].

THEOREM 3·2. *Let* q *be an integer,* $a \neq b$ [mod q] *two invertible residue classes. Define*

$$\beta_{q,0} = \sup\big\{\mathrm{Re}(\rho) : \exists \chi \bmod q, \chi \neq \chi_0, L(\chi, \rho) = 0\big\}.$$

The function

$$E_{q;a,b}(x) = \frac{\log x}{x^{\beta_{q,0}}}\big(\pi(x, q, a) - \pi(x, q, b)\big)$$

has a limiting logarithmic distribution $\mu_{q;a,b}$ *on* $\mathbf{R}$. *There exists a positive constant* C *(depending on* q*) such that one has*

$$\mu_{q;a,b}\big(\mathbf{R} - [-A, A]\big) \ll \exp\big(-C\sqrt{A}\big).$$

Moreover, suppose GRH is satisfied (i.e. $\beta_{q,0} = 1/2$*) and* $L(\chi, 1/2) \neq 0$ *for every* $\chi \bmod q$, $\chi \neq \chi_0$. *Then the mean value of* $\mu_{q;a,b}$ *is*

$$m_{q;a,b} = \sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0}} \big(\chi(b) - \chi(a)\big).$$

Remark 5. In particular under these hypotheses and the conditions of Corollary 2·4.

(i) if ab^{-1} is a square, then there is no bias.
(ii) otherwise, the bias is in the direction of the non quadratic residue.

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

Following the idea of [RS94] and [Fio14b], we can also study the prime number race between the subsets of quadratic residues and non-quadratic residues modulo an integer q. For this, take

$$\mathcal{S} = \{L(\chi, \cdot) : \chi \bmod q, \chi \neq \chi_0, \chi^2 = \chi_0\},$$

and for each real character χ modulo q, take $a_\chi = 1/\rho(q) := [(\mathbf{Z}/q\mathbf{Z})^\times : (\mathbf{Z}/q\mathbf{Z})^{\times(2)}]^{-1}$. We apply Theorem 2·1 to this setting and get the following result.

THEOREM 3·3. *Let $q \geq 3$ be an integer. Define*

$$\beta_{q,0}^{(2)} = \sup\{\mathrm{Re}(\rho) : \exists \chi \bmod q, \chi \neq \chi_0, \chi^2 = \chi_0, L(\chi, \rho) = 0\}.$$

The function

$$E_{q;\mathrm{R,NR}}(x) = \frac{\log(x)}{\rho(q)x^{\beta_{q,0}^{(2)}}}\big((\rho(q) - 1)\pi(x; q, \mathrm{R}) - \pi(x; q, \mathrm{NR})\big)$$

has a limiting logarithmic distribution $\mu_{q;\mathrm{R,NR}}$ on $\mathbf{R}$.

Moreover, suppose GRH is satisfied for all Dirichlet L-function of real characters modulo q and that $L(\chi, 1/2) \neq 0$ for every $\chi \bmod q$, $\chi \neq \chi_0$, $\chi^2 = \chi_0$. Then the average value of $\mu_{q;\mathrm{R,NR}}$ is

$$m_{q;a,b} = \frac{1 - \rho(q)}{\rho(q)}.$$

Thus we have obtained an unconditional proof (without GRH) of [RS94, Theorem 1·1] in the case of the prime number race between the subsets of quadratic residues and non-residues modulo an integer q (see also [Fio14b, Lemma 2·2]). Under GRH, the mean of the logarithmic limiting distribution is negative, hence we expect to find a bias towards negative values (i.e. towards non-quadratic residues). This result has already been used by Fiorilli in [Fio14b] to find arbitrarily biased races between quadratic residues and non-quadratic residues modulo integers having a lot of prime factors (so that the mean value is as far from 0 as possible).

3·1·3. *L-functions of automorphic forms on $GL(m)$*

As announced in Example 1(iv), by results on automorphic L-functions ([BF90, BG92, GJ72, JS90]), we only miss the Ramanujan–Petersson conjecture to ensure that L-functions associated to irreducible cuspidal automorphic forms on $GL(m)$ are analytic L-functions in the sense of Definition 1·1. We get a version of Theorem 2·1 for automorphic L-functions conditional on the Ramanujan–Petersson conjecture.

THEOREM 3·4. *Let $L(\pi, s)$ be a real L-function associated to an irreducible unitary cuspidal automorphic representation of $GL(m)$ with $m \geq 2$. Suppose the Ramanujan–Petersson conjecture holds for $L(\pi, s)$. Then, following the notations of Section 1·2, the function $E_\pi(x) = (\log x / x^{\beta_{\pi,0}}) \sum_{p \leq x} \lambda_\pi(p)$ has a limiting logarithmic distribution μ_π.*

Moreover under GRH for $L(\pi, s)$, the mean of μ_π is $m_\pi = \pm 1 - 2m(L(\pi, \cdot), 1/2) \neq 0$.

This result should be compared to [ANS14, Corollary 1·5] where GRH is assumed but not the Ramanujan–Petersson Conjecture. Under GRH, since the mean value is non zero, we expect that the prime number race associated to such an L-function has always a bias.

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

Proof. Under GRH for $L(\pi, s)$, we study the behaviour of the function

$$L(\pi^{(2)}, s) = L(\mathrm{Sym}^2 \pi, s) L(\wedge^2 \pi, s)^{-1}$$

around $s = 1$. By [**Sha97**, Theorem 1·1], in the case π is an irreducible non trivial representation, the functions $L(\mathrm{Sym}^2 \pi, s)$, $L(\wedge^2 \pi, s)$ do not vanish at $s = 1$. Moreover one has $L(\pi \otimes \pi, s) = L(\mathrm{Sym}^2 \pi, s) L(\wedge^2 \pi, s)$, and ([**MW89**, Appendix]) this function has a simple pole at $s = 1$ when π is self-dual (i.e. $L(\pi, s)$ is real). Hence there are only two possibilities:

(i) either $L(\mathrm{Sym}^2 \pi, s)$ has a simple pole at $s = 1$ and $m\big(L(f^{(2)}, \cdot), 1\big) = -1$;
(ii) or $L(\wedge^2 \pi, s)$ has a simple pole at $s = 1$ and $m\big(L(f^{(2)}, \cdot), 1\big) = 1$.

Theorem 3·4 follows.

In the case $m = 2$, the Ramanujan–Petersson conjecture has been proved by works of Deligne and Deligne–Serre [**Del74**, **DS74**]. In particular the normalized Hasse–Weil L-function associated to an elliptic curve defined over $\mathbf{Q}$ is an analytic L-function (see Example 1(iii)). Hence we deduce [**Fio14a**, Lemma 2·3, Lemma 2·6, Lemma 3·4] from Theorem 2·1.

PROPOSITION 3·5. *Let $E/\mathbf{Q}$ be an elliptic curve, and $L(E, s)$ its normalized Hasse L-function. The function*

$$E_E(x) = \frac{\log x}{x^{\beta_{E,0}}} \sum_{p \leq x} \frac{a_p(E)}{\sqrt{p}}$$

has a limiting logarithmic distribution μ_E on $\mathbf{R}$. Moreover, suppose GRH is satisfied for $L(E, s)$. Then the mean of μ_E is

$$m_E = -2 r_{an}(E) + 1$$

where $r_{an}(E)$ is the analytic rank of E.

Proof. In the case of a Hasse–Weil L-function attached to an elliptic curve $E/\mathbf{Q}$, one has $L(\wedge^2, E, s) = \zeta(s)$. Proposition 3·5 follows.

We observe the two distinct cases pointed out by Mazur: either $r_{an}(E) = 0$ and we should expect a bias towards positive values, or $r_{an}(E) > 0$ and we should expect a bias towards negative values. As Fiorilli noticed in [**Fio14a**] we can expect an arbitrarily large bias in the case the rank of the elliptic curve is arbitrarily large compared to the variance of the distribution.

3·2. *New applications*

3·2·1. *Chebyshev's bias and prime numbers of the form $a^2 + Db^2$*

In [**SB85**], Beukers and Stienstra give several examples of L-functions of degree 2 related to $K3$ surfaces. Precisely they define the three following functions:

$$L_D(s) = \prod_{p \nmid 2D} \left(1 - a_p p^{-s} + \left(\frac{-D}{p}\right) p^{-2s}\right)^{-1}$$

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

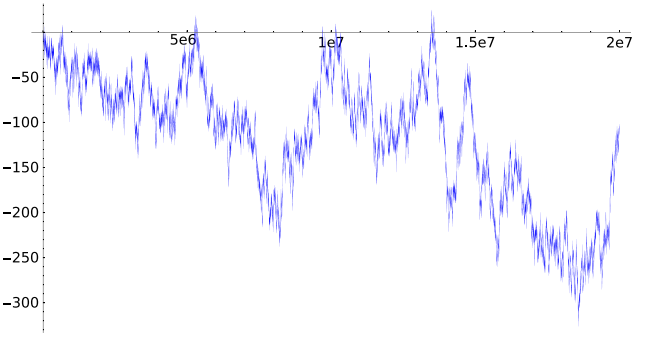


Fig. 1. Values of $S_2(x)$ in the range $(0, 2 \cdot 10^7)$.

for $D = 2$, 3 and 4, where

$$a_p = \begin{cases} 0 & \text{if } \left(\frac{-D}{p}\right) = 0 \text{ or } -1, \\ 2\frac{a^2 - Db^2}{p} & \text{if one can write } p = a^2 + Db^2 \text{ with } a, b \in \mathbf{N}. \end{cases}$$

By [**SB85**, Theorem 14·2] (see also [**Sch53**] in the case $D = 4$), those L-functions are associated to cusp forms of weight 3 and level $4D$. In particular they satisfy Definition 1·1.

To these functions L_D one can associate the prime number race that consists in understanding the sign of

$$E_D(x) = \frac{\log x}{x^{\beta_{0,D}}} \sum_{p = a^2 + Db^2 \leq x} 2\frac{a^2 - Db^2}{a^2 + Db^2}. \tag{3·1}$$

The adaptation of Theorem 2·1 to this context is the following result.

THEOREM 3·6. *For $D = 4$, 2 and 3, E_D has a limiting logarithmic distribution whose average value is*

$$-\frac{m(L_D, \beta_0)}{\beta_0} - \delta_{\beta_0 = \frac{1}{2}} \leq 0.$$

Proof. By [**SB85**, Theorem 14·2], we are in the situation of Theorem 3·4. In particular the limiting logarithmic distribution exists. One can compute that, for each of the three cases $D = 2$, 3 and 4, and for every p, the products of the local roots are $\alpha_1(p)\alpha_2(p) = (-D/p)$. We deduce

$$L\left(\wedge^2 f_D, s\right) = \prod \left(1 - \left(\frac{-D}{p}\right) p^{-s}\right)^{-1},$$

and in particular this function is entire. Hence (by [**MW89**, Appendix]) the L-function $L(\mathrm{Sym}^2 f_D, s)$ has a pole of multiplicity 1 at $s = 1$. In conclusion the function $L(f_D^{(2)}, s)$ has a pole of multiplicity 1 at $s = 1$.

We can interpret this result by saying that in the decomposition $p = a^2 + Db^2$, the term Db^2 is often larger than a^2. The Figures 1, 2 and 3 represent respectively the races between a^2 and $-Db^2$ for $D = 2, 3, 4$. We used `sage` and Cornacchia's algorithm to obtain the

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

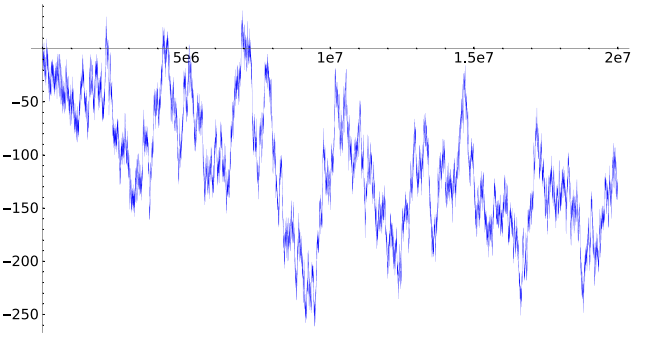


Fig. 2. Values of $S_3(x)$ in the range $(0, 2 \cdot 10^7)$.

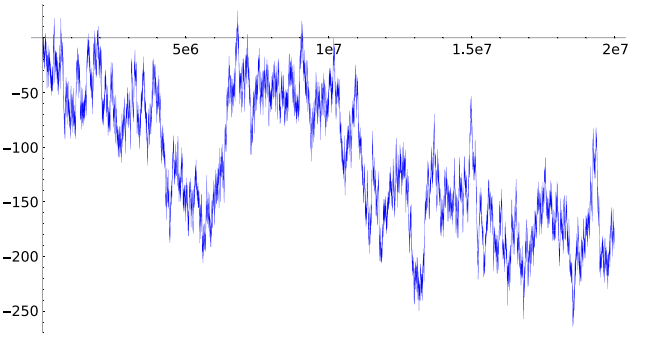


Fig. 3. Values of $S_4(x)$ in the range $(0, 2 \cdot 10^7)$.

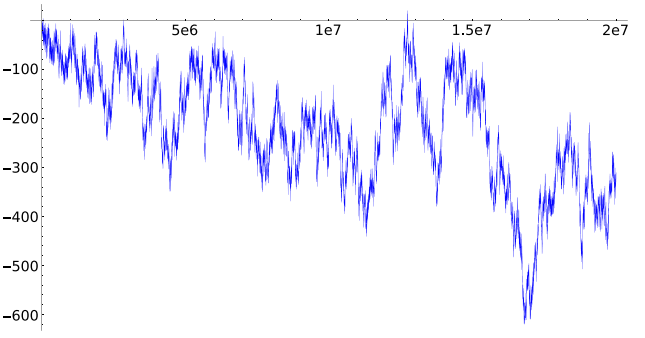


Fig. 4. Values of $S_\lambda(x)$ in the range $(0, 2 \cdot 10^7)$.

values of the functions $S_D(x) := \sum_{p=a^2+Db^2 \le x} (a^2 - Db^2)/(a^2 + Db^2)$ for x between 0 and $2 \cdot 10^7$. We see on these figures that it is natural to expect a bias towards the negative values.

3·2·2. *Prime number races for angles of Gaussian primes*

As Z. Rudnick pointed out to us, in the case $D = 4$, the prime number race (3·1) is related to the question of the bias in the distribution of the angles of the Gaussian

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

primes. Let λ be the Hecke character on $\mathbf{Z}[i]$ defined by $\lambda(z) = (z/\bar{z})^2$. The L-function $L(\lambda, s) = \prod_p \prod_{\mathfrak{p}|p}(1 - \lambda(\mathfrak{p})N\mathfrak{p}^{-s})^{-1}$ (seen as a Euler product over rational primes) is an analytic L-function in the sense of Definition 1·1. In the case $p \equiv 1 \ [\mathrm{mod}\ 4]$ the local factor is $(1 - \cos(4\theta_p)p^{-s} + p^{-2s})^{-1}$ where $\pm\theta_p$ are the angle of the Gaussian primes dividing p (they are defined modulo $\pi/2$). The prime number race associated with this situation consists in understanding the sign of the function

$$E_\lambda(x) = \frac{\log x}{x^{\beta_{0,\lambda}}} \sum_{\substack{p \leq x \\ p \equiv 1 \ [\mathrm{mod}\ 4]}} \cos(4\theta_p).$$

THEOREM 3·7. *The function E_λ admits a limiting logarithmic distribution whose average value is negative.*

Proof. See for example [**RV99**, Theorems 7-19] to verify the hypotheses of Definition 1·1. In the case $p \equiv 3 \ [\mathrm{mod}\ 4]$ the local factor of $L(\lambda, s)$ is $(1 - p^{-2s})^{-1}$, so the local roots as a L-function of degree 2 over $\mathbf{Q}$ are ± 1. Hence for every p, the product of the local roots is $(-1/p)$. Thus $L\left(\wedge^2\lambda, s\right) = \prod \left(1 - (-1/p)\ p^{-s}\right)^{-1}$ is entire. Similarly the function $L(\mathrm{Sym}^2\,\lambda, s) = \zeta(s)L(\lambda^2, s)$ has a pole of multiplicity 1 at $s = 1$. In conclusion the function $L(\lambda^{(2)}, s)$ has a pole of multiplicity 1 at $s = 1$.

This result implies that the corresponding prime number race should be biased towards negative values. This can be guessed from computations. In Figure 4 we used `sage` to compute the values of the function $S_\lambda(x) := \sum_{\substack{p \leq x \\ p \equiv 1 \ [\mathrm{mod}\ 4]}} \cos(4\theta_p)$ for x between 0 and $2 \cdot 10^7$. The consequence of such a bias towards negative values is that we expect that the Gaussian primes are more often closer to the line $y = x$ than to the coordinate axes.

3·2·3. *Correlations for two elliptic curves*

As advertised in Example 1(v), we can study the prime number race associated to a Rankin–Selberg product of L-functions. Let E_1 and E_2 be two non-isogenous non-CM elliptic curves defined over $\mathbf{Q}$. By works of Wiles, Taylor–Wiles, Breuil–Conrad–Diamond–Taylor ([**Wil95**, **TW95**, **BCDT01**]), there exists cuspidal modular forms $f_1 \neq f_2$ associated to E_1 and E_2 respectively (i.e. the corresponding normalized L-functions are the same). One has

$$\lambda_{f_i}(p) = \frac{a(E_i, p)}{\sqrt{p}} = \frac{p + 1 - |E_i(\mathbf{F}_p)|}{\sqrt{p}}.$$

In the case E_1 and E_2 do not become isogenous over a quadratic extension of $\mathbf{Q}$, by [**Ram00**] the Rankin–Selberg convolution $L(f_{E_1} \otimes f_{E_2}, \cdot)$ is a real analytic L-function in the sense of Definition 1·1. Its coefficients are $\lambda(p) = a_p(E_1)a_p(E_2)/p$. Moreover, if we assume that the curves E_1 and E_2 do not become isogenous over any abelian extension of $\mathbf{Q}$, a strong version of Conjecture 1·2 holds for these coefficients (see [**Har09**, Theorem 5·4]).

Hence we can apply Theorem 2·1. The function $E(x) = (\log x/x^{\beta_0}) \sum_{p \leq x} a_p(E_1)$ $a_p(E_2)/p$ admits a limiting logarithmic distribution, and we can give its mean explicitly. The term $m(L(f_{E_1} \otimes f_{E_2}, \cdot), 1/2)$ may not be easy to evaluate, but $m(L((f_{E_1} \otimes f_{E_2})^{(2)}, \cdot), 1)$ can be computed. From these considerations we obtain the following result.

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

THEOREM 3·8. *Let E_1 and E_2 be two non-CM elliptic curves defined over* $\mathbf{Q}$. *Assume E_1 and E_2 do not become isogenous over a quadratic extension of* $\mathbf{Q}$. *The function*

$$E(x) = \frac{\log x}{x^{\beta_0}} \sum_{p \leq x} \frac{a_p(E_1)a_p(E_2)}{p}$$

admits a limiting logarithmic distribution. Assuming also that the Riemann Hypothesis holds, the logarithmic distribution has negative mean value.

This result is a consequence of the following Lemma.

LEMMA 3·9. *Let E_1 and E_2 be two non-CM elliptic curves defined over* $\mathbf{Q}$. *Suppose that E_1 and E_2 do not become isogenous over any quadratic extension of* $\mathbf{Q}$, *then*

$$m\Big(L\big((f_{E_1} \otimes f_{E_2})^{(2)}, \cdot\big), 1\Big) = -1.$$

Proof. To fix the notation we write for $i = 1, 2$,

$$L(f_{E_i}, s) = \prod_p \big(1 - \pi_i p^{-s}\big)^{-1}\big(1 - \overline{\pi_i} p^{-s}\big)^{-1}.$$

The local roots of $L(f_{E_1} \otimes f_{E_2}, \cdot)$ at p are $\pi_1\pi_2$, $\overline{\pi_1}\pi_2$, $\pi_1\overline{\pi_2}$ and $\overline{\pi_1\pi_2}$. Hence

$$\begin{aligned} L\big(\wedge^2 (f_{E_1} \otimes f_{E_2}), s\big) &= \prod_p \prod_{i=1,2} \big(1 - \pi_i^2 p^{-s}\big)^{-1}\big(1 - \overline{\pi_i}^2 p^{-s}\big)^{-1}(1 - p^{-s})^{-1} \\ &= L\big(\operatorname{Sym}^2 f_{E_1}, s\big)L\big(\operatorname{Sym}^2 f_{E_2}, s\big) \end{aligned}$$

and

$$L\big(\operatorname{Sym}^2(f_{E_1} \otimes f_{E_2}), s\big) = L\big(\operatorname{Sym}^2 f_{E_1} \otimes \operatorname{Sym}^2 f_{E_2}, s\big)\zeta(s).$$

For $i = 1, 2$, the function $L(\operatorname{Sym}^2 f_{E_i}, \cdot)$ is holomorphic and does not vanish at $s = 1$. By [**Ram00**] one can associate to $f_{E_1} \otimes f_{E_2}$ a cuspidal irreducible representation of $GL(4)$ with the same L-function, hence by [**MW89**, Appendix] the function

$$L\big((f_{E_1} \otimes f_{E_2}) \otimes (f_{E_1} \otimes f_{E_2}), s\big) = L\big(\operatorname{Sym}^2(f_{E_1} \otimes f_{E_2}), s\big)L\big(\wedge^2 (f_{E_1} \otimes f_{E_2}), s\big)$$

has a pole of multiplicity 1 at $s = 1$. As a consequence:

$$L\big((f_{E_1} \otimes f_{E_2})^{(2)}, s\big) = L\big(\operatorname{Sym}^2(f_{E_1} \otimes f_{E_2}), s\big)L\big(\wedge^2 (f_{E_1} \otimes f_{E_2}), s\big)^{-1}$$

has a pole of multiplicity 1 at $s = 1$.

The proof of Theorem 3·8 then follows from Theorem 2·1 and Lemma 3·9. Under the Riemann Hypothesis, the average value of the limiting logarithmic distribution is

$$-2m\left(L(f_{E_1} \otimes f_{E_2}, \cdot), \frac{1}{2}\right) - 1 < 0.$$

We may interpret this result by saying that given two non-isogenous elliptic curves (in the strong sense used above), the coefficients $a_p(E_1)$ and $a_p(E_2)$ often have opposite signs. The Figures 5, 6, 7 and 8 represent various prime number races for the correlations of the signs

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

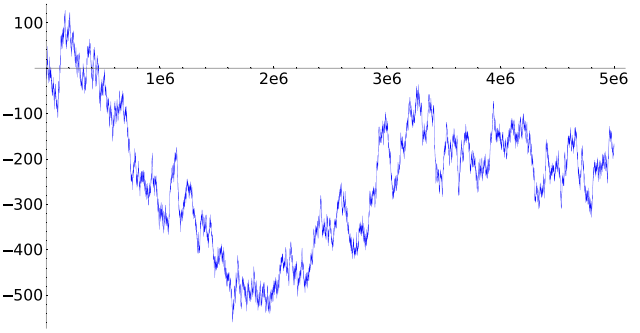


Fig. 5. Values of $S_{E_1,E_2}(x)$ in the range $(0, 5 \cdot 10^6)$.

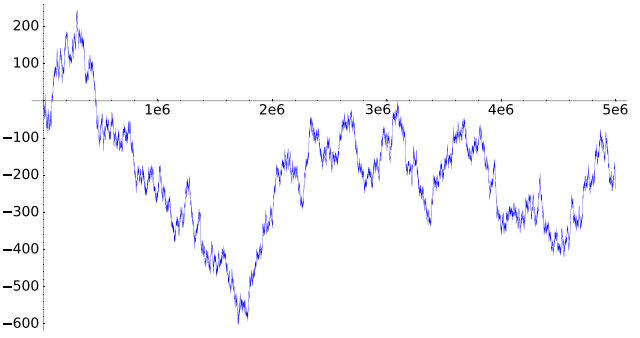


Fig. 6. Values of $S_{E_0,E_2}(x)$ in the range $(0, 5 \cdot 10^6)$.

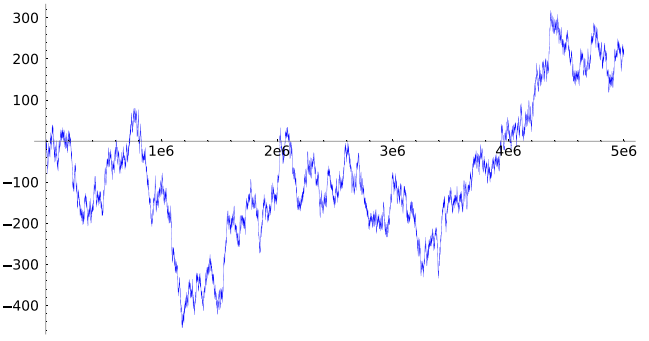


Fig. 7. Values of $S_{E_0,E_1}(x)$ in the range $(0, 5 \cdot 10^6)$.

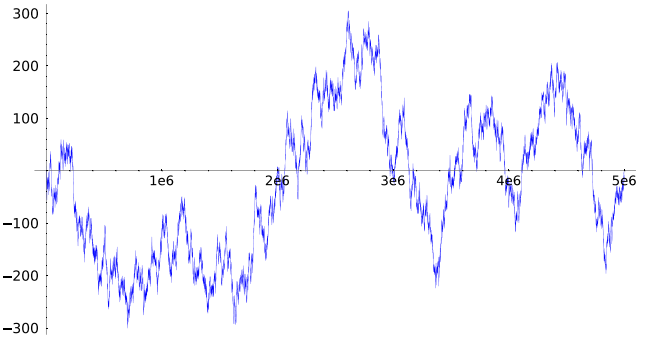


Fig. 8. Values of $S_{E_0,E_0'}(x)$ in the range $(0, 5 \cdot 10^6)$.

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

of the $a_p(E)$ for two elliptic curves. We used four elliptic curves that we can define by an affine model as follows:

$$E_1 : y^2 + y = x^3 - x, \quad E_2 : y^2 + y = x^3 + x^2 - 2x,$$
$$E_0 : y^2 + y = x^3 - x^2, \quad E_0' : y^2 + y = x^3 + x^2 + x.$$

The elliptic curves have algebraic rank respectively equal to 1, 2 and 0. We used `sage` and the counting points algorithm for elliptic curves implemented in `pari` to obtain the values of the functions $S_{E_i,E_j}(x) := \sum_{p \leq x} a_p(E_i)a_p(E_j)/p$ for x between 0 and $5 \cdot 10^6$. The bias towards negative values can be guessed from Figures 5 and 6, it is less clear on Figures 7 and 8. The bias may be smaller in the last two cases and appear only on a larger scale.

3·2·4. *Jacobian of modular curves.*

Our last example is the prime number race for the L-functions of the modular curves. Let q be a prime number. We study the prime number race for the sum of the coefficients of all L-functions of primitive weight two cusp forms of level q. The L-function associated to this race is the finite product

$$\prod_{f \in S_2(q)^*} L(f, s) = L(J_0(q), s),$$

where $J_0(q)$ is the Jacobian of the modular curve $X_0(q)$ (this factorisation is due to Shimura [**Shi94**]). The function $L(J_0(q), \cdot)$ is an analytic L-function in the sense of Definition 1·1 since it is a product of analytic L-functions.

Assuming the Riemann Hypothesis for $L(J_0(q), s)$, Theorem 2·1 applies to the function

$$E_{J_0(q)}(x) = \frac{\log x}{\sqrt{x}} \sum_{p \leq x} \sum_{f \in S_2(q)^*} \lambda_f(p).$$

One can conjecture a value for the mean of the limiting logarithmic distribution.

CONJECTURE 3·10. *One has:*

$$m\left(L(J_0(q), \cdot), \frac{1}{2}\right) \sim \frac{1}{2}|S_2(q)^*|$$

as $q \to \infty$.

In the articles [**KM00a**] and [**KM00b**], Kowalski and Michel showed that there exist two explicit constants $c < 1/2 < C$ such that

$$c|S_2(q)^*| \leq m\left(L(J_0(q), \cdot), \frac{1}{2}\right) \leq C|S_2(q)^*|,$$

for all sufficiently large q.

The large multiplicity given by Conjecture 3·10 may lead us to think that we could get a large bias, but considering all the primitive weight two forms of level q at once, the biases towards positive or negative values should in fact cancel each other. Precisely:

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

THEOREM 3·11. *Assume the Riemann Hypothesis for $L(J_0(q), \cdot)$ (for all q) and assume Conjecture 3·10 holds. Then the function*

$$E_{J_0(q)}(x) = \frac{\log x}{\sqrt{x}} \sum_{p \leq x} \sum_{f \in S_2(q)^*} \lambda_f(p)$$

admits a limiting logarithmic distribution with mean $o_{q\to\infty}(|S_2(q)^*|)$ and variance $\gg |S_2(q)^*| \log q$.

Remark 6. In this situation Chebyshev's inequality is not conclusive (see Section 5·4). As is the case for the original work of [**RS94**] the bias probably dissipates as $q \to \infty$. If we want to show this, we need a better error term in Conjecture 3·10: we need that $m_q/\sqrt{\mathrm{Var}_q} \to 0$ as $q \to \infty$.

For the proof of Theorem 3·11, we compute $m(L(J_0(q)^2, \cdot), 1)$.

LEMMA 3·12. *Let q be an integer. One has $m(L(J_0(q)^{(2)}, \cdot), 1) = |S_2(q)^*|$.*

For the record, one has $|S_2(q)^*| \sim q/12$.

Proof. As in the proof of Lemma 3·9, we use local roots to determine the multiplicities of the zero at $s = 1$ of $L(\wedge^2(J_0(q)), \cdot)$ and $L(\mathrm{Sym}^2(J_0(q)), \cdot)$. For $f \in S_2(q)^*$, denote by $\alpha_f(p)$ and $\overline{\alpha_f(p)}$ its local roots. They satisfy $\alpha_f(p)\overline{\alpha_f(p)} = 1$ if $p \nmid q$. One has

$$L\big(\wedge^2(J_0(q)), s\big) = \zeta_q(s)^{|S_2(q)^*|} \prod_{f \neq f'} L(f \otimes f', s)$$

and

$$L\big(\mathrm{Sym}^2(J_0(q)), s\big) = \prod_f L(\mathrm{Sym}^2 f, s) \prod_{f \neq f'} L(f \otimes f', s).$$

Hence $L(\wedge^2(J_0(q)), \cdot)$ has a pole of multiplicity $|S_2(q)^*|$ at $s = 1$, and $L(\mathrm{Sym}^2(J_0(q)), \cdot)$ is holomorphic and does not vanish at $s = 1$. We conclude that $L(J_0(q)^{(2)}, \cdot)$ has a zero of multiplicity $|S_2(q)^*|$ at $s = 1$.

Proof of Theorem 3·11. It follows from Theorem 2·1. and Lemma 3·12, under the Riemann Hypothesis that the mean of the limiting logarithmic distribution is

$$2m\left(L(J_0(q), \cdot), \frac{1}{2}\right) - |S_2(q)^*|.$$

If we assume Conjecture 3·10 is satisfied, then the mean is $= o(|S_2(q)^*|)$. The variance is

$$\sum_{\substack{L(J_0(q), \frac{1}{2}+i\gamma)=0 \\ \gamma \neq 0}}^{*} \frac{m(L(J_0(q), \cdot), \frac{1}{2}+i\gamma)^2}{(\frac{1}{4}+\gamma^2)} \gg \sum_{\substack{L(J_0(q), \frac{1}{2}+i\gamma)=0 \\ \gamma \neq 0}} \frac{1}{(\frac{1}{4}+\gamma^2)}$$

$$\gg \log\big(\mathfrak{q}(J_0(q))\big) \asymp |S_2(q)^*| \log q.$$

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

4. *Proof of Theorem 2·1*

In this section we prove Theorem 2·1 as a consequence of the following result relating $\mu_{\mathcal{S}}$ with the zeros of the L-functions.

PROPOSITION 4·1. *Let $\{L(f,\cdot) : f \in \mathcal{S}\}$ be a finite set of analytic L-functions such that $\overline{\mathcal{S}} = \mathcal{S}$, and $(a_f)_{f\in\mathcal{S}}$ a set of complex numbers satisfying $a_{\overline{f}} = \overline{a_f}$. Let $T > 2$ and*

$$G_{\mathcal{S},T}(x) = m_{\mathcal{S}} - \sum_{\gamma\in\mathcal{Z}^*_{\mathcal{S}}(T)} 2\,\mathrm{Re}\left(M(\gamma)\frac{x^{i\gamma}}{\beta_{\mathcal{S},0}+i\gamma}\right),$$

where as in Theorem 2·1, for γ in $\mathcal{Z}^*_{\mathcal{S}}$, one has $M(\gamma) = \sum_{f\in\mathcal{S}} a_f m(L(f,\cdot), \beta_{\mathcal{S},0}+i\gamma)$.

The function $G_{\mathcal{S},T}(x)$ admits a limiting logarithmic distribution $\mu_{\mathcal{S},T}$. Moreover, for any bounded Lipschitz continuous function g, one has

$$\lim_{T\to\infty}\int_{\mathbf{R}} g(t)\,\mathrm{d}\mu_{\mathcal{S},T}(t) = \int_{\mathbf{R}} g(t)\,\mathrm{d}\mu_{\mathcal{S}}(t).$$

Remark 7.

(i) In the case $\mathcal{Z}_{\mathcal{S}}$ is empty (it may happen if the Riemann Hypothesis is not satisfied), the functions $G_{\mathcal{S},T}(x)$ are constant, and do not depend on T. Hence the limiting logarithmic distributions $\mu_{\mathcal{S},T}$ and $\mu_{\mathcal{S}}$ exist and are equal to the Dirac delta function $\delta_{m_{\mathcal{S}}}$. In particular in the case $\beta_{\mathcal{S},0} = 1$, the set $\mathcal{Z}_{\mathcal{S}}$ is empty, and the limiting logarithmic distribution is δ_0. Hence the only information we get from Theorem 2·1 is that $S(x) = o(x/\log x)$.

(ii) Another approach for this result can be found in [**ANS14**]. The function $G_{\mathcal{S},T}(e^y)$ is a *trigonometric polynomial*, and as $T\to\infty$ it approximates the function $E_{\mathcal{S}}(e^y)$. The improvement in our result is that we do not need to assume that the Generalized Riemann Hypothesis holds.

To obtain this result (except for the statement about $\mathrm{Var}(X_{\mathcal{S}})$) it is enough to consider the case where $\mathcal{S}$ is a singleton $\{f\}$ and $a_f = 1$ (by linearity). The proof follows ideas from [**Fio14a**, Lemma 3·4] and [**ANS14**], hence we only give the necessary extra details. The proof is decomposed in the following way. Subsections 4·2 and 4·3 are dedicated to the proof that the functions $G_{f,T}(x)$ are good approximations for $E_f(x)$. The existence part of the proposition is proved in Subsection 4·4 as a consequence of the Kronecker–Weyl Theorem (of which we sketch the proof in Subsection 4·1) and Helly's selection Theorem. We conclude the proof of Theorem 2·1 in Subsection 4·5 by computing the mean and variance of the limiting logarithmic distribution $\mu_{\mathcal{S}}$.

4·1. *Preliminary result on Kronecker–Weyl Theorem*

We prove a generalization of Kronecker–Weyl equidistribution theorem without assuming linear independence following the idea given by Humphries in [**Hum10**].

THEOREM 4·2. *Let $\gamma = (\gamma_1,\dots,\gamma_N) \in \mathbf{R}^N$ be an N-uple of real numbers. Let $A(\gamma)$ be the closure of the 1-parameter group $\{y(\gamma_1,\dots,\gamma_N) : y \in \mathbf{R}\}/\mathbf{Z}^N$ in the N-dimensional torus $\mathbf{T}^N := (\mathbf{R}/\mathbf{Z})^N$. Let $h : \mathbf{T}^N \to \mathbf{C}$ be a continuous function. Then $A(\gamma)$ is a sub-torus of $\mathbf{T}^N$ and we have*

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

$$\lim_{Y\to\infty}\frac{1}{Y}\int_0^Y h(y\gamma_1,\dots,y\gamma_N)\,\mathrm{d}y=\int_{A(\gamma)}h(a)\,\mathrm{d}\omega_{A(\gamma)}(a), \tag{4·1}$$

where ω_A is the normalized Haar measure on A satisfying $\omega_A(A)=1$.

Let us first note that by the continuity of the group operations, $A(\gamma)$ is a closed subgroup of $\mathbf{T}^N$ hence a sub-torus (in particular it is compact). Thus the Haar measure used in the theorem is well defined.

To understand the group $A(\gamma)$ we will use its annihilator. We define the dual of a sub-torus A as the group $\widehat{A}$ of continuous group homomorphisms (characters) from A to $\mathbf{T}$. Note that one has $\widehat{\mathbf{T}^N}\simeq\mathbf{Z}^N$ given by the pairing $\langle x,k\rangle=x_1k_1+\cdots+x_Nk_N\in\mathbf{T}$. Then the annihilator of a subgroup $A<\mathbf{T}^N$ is the closed subgroup of $\mathbf{Z}^N$ ([**Rud90**, 2·1·1]) defined by

$$A^\perp=\left\{k\in\mathbf{Z}^N:\forall x\in A,\ \langle x,k\rangle=0\right\}.$$

In the other direction, using the same pairing, we see that the set of characters of $\mathbf{Z}^N$ is in correspondence with $\mathbf{T}^N$. The annihilator of a subgroup $B<\mathbf{Z}^N$ is the closed subgroup of $\mathbf{T}^N$ defined by

$$B_\perp=\left\{x\in\mathbf{T}^N:\forall k\in B,\ \langle x,k\rangle=0\right\}.$$

The following result gives a precise description of $A(\gamma)$ using its annihilator.

LEMMA 4·3. *Let A be a subgroup in $\mathbf{T}^N$, one has*

$$(A^\perp)_\perp=\overline{A}.$$

In particular for $\gamma=(\gamma_1,\dots,\gamma_N)\in\mathbf{R}^N$, the annihilator of the group $A(\gamma)$ is

$$A(\gamma)^\perp=\left\{(k_1,\dots,k_N)\in\mathbf{Z}^N:k_1\gamma_1+\cdots k_N\gamma_N=0\right\}.$$

We note that if $\gamma_1,\dots,\gamma_N$ are linearly independent over $\mathbf{Q}$ then $A(\gamma)^\perp=\{0\}$, hence $A(\gamma)=\mathbf{T}^N$.

Proof. First, from the fact that the characters are continuous we have $A^\perp=\overline{A}^\perp$. Thus, it suffices to obtain the result for closed subgroups. This follows from the fact that a non-trivial character admits a non-trivial value (see also [**Bou47**, Chapter VII Section 1, Proposition 4]).

In our particular case, the annihilator of the 1-parameter group $\{y(\gamma_1,\dots,\gamma_N):y\in\mathbf{R}\}/\mathbf{Z}^N$ is

$$\begin{aligned}\left\{(k_1,\dots,k_N)\in\mathbf{Z}^N:\forall y\in\mathbf{R},\left\langle y(\gamma_1,\dots,\gamma_N),(k_1,\dots,k_N)\right\rangle=0_{\mathbf{T}}\right\}&\\ =\left\{(k_1,\dots,k_N)\in\mathbf{Z}^N:\forall y\in\mathbf{R},\ y(\gamma_1k_1+\cdots+\gamma_Nk_N)\in\mathbf{Z}\right\}&\\ =\left\{(k_1,\dots,k_N)\in\mathbf{Z}^N:\gamma_1k_1+\cdots+\gamma_Nk_N=0_{\mathbf{R}}\right\}&\end{aligned}$$

which gives the conclusion.

Let us finally give the proof of this version of the Kronecker–Weyl equidistribution Theorem.

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

Proof of Theorem 4·2. This proof follows [**Hum10**, Lemma 4·3]. Since $\mathbf{T}^N$ is compact, continuous functions on $\mathbf{T}^N$ are uniformly continuous, in particular they are limits of polynomials in the uniform convergence topology. Thus, it is enough to show that the two terms in (4·1) are equal when h is a trigonometric polynomial. Then, by linearity, it is enough to show it is true for monomials (i.e. characters of $\mathbf{T}^N$). Let $k = (k_1, \dots, k_N) \in \mathbf{Z}^N$, and χ_k be the associated character (with image in $\mathbf{C}$) in $\widehat{\mathbf{T}^N}$: $\chi_k(x) = e^{2\pi i(k_1x_1+\dots+k_Nx_N)}$.

Then, using $\widehat{A} \simeq \mathbf{Z}^N/A^\perp$ (see e.g. [**Bou47**, Chapter VII, Section 1·5]) and orthogonality of the characters, we have that the left-hand side of the equality is

$$\int_{A(\gamma)} \chi_k(a)\, \mathrm{d}\omega_{A(\gamma)} = \begin{cases} 1 \text{ if } k \in A(\gamma)^\perp \\ 0 \text{ otherwise.} \end{cases}$$

For the right-hand side of the equality, one has

$$\int_0^Y \chi_k(y\gamma_1, \dots, y\gamma_N)\, \mathrm{d}y = \int_0^Y e^{2\pi i y(k_1\gamma_1+\dots+k_N\gamma_N)}\, \mathrm{d}y = \begin{cases} Y \text{ if } k_1\gamma_1 + \dots + k_N\gamma_N = 0 \\ O(1) \text{ if } k_1\gamma_1 + \dots + k_N\gamma_N \neq 0. \end{cases}$$

This concludes the proof.

4·2. *Approximation of* $\psi(f, x)$

It is a standard step in proofs of theorems reminiscent of the Prime Number Theorem to start by studying the associated ψ-function:

$$\psi(f, x) = \sum_{k=1}^{\infty} \sum_{p^k \leq x} \left(\sum_{j=1}^{d} \alpha_j(p)^k \right) \log p.$$

Note that for $\mathrm{Re}(s) > 1$, one has

$$-\frac{L'(f, s)}{L(f, s)} = \sum_{k=1}^{\infty} \sum_{p} \left(\sum_{j=1}^{d} \alpha_j(p)^k \right) p^{-ks} \log p =: \sum_{n=1}^{\infty} \Lambda_f(n) n^{-s}.$$

Then Perron's Formula and integration around the zeros yields an explicit formula for $\psi(f, x)$.

PROPOSITION 4·4. *Let* $L(f, \cdot)$ *be an analytic L-function. One has*

$$\psi(f, x) + m\big(L(f, \cdot), 1\big)x = -\sum_{\substack{L(f,\rho)=0 \\ |\mathrm{Im}(\rho)| \leq T}} \frac{x^\rho}{\rho} - x^{\beta_{f,0}} \epsilon_f(x, T) + O\left(\log(\mathfrak{q}(f)x^d) \log x\right), \tag{4·2}$$

where the function $\epsilon_f(x, T)$ *satisfies*

$$\int_2^Y |\epsilon_f(e^y, T)|^2\, \mathrm{d}y \ll Y \frac{d^2\big(\log(\mathfrak{q}(f)T)\big)^2}{T} + \frac{d^2\big(\log(\mathfrak{q}(f)T)\big)^2 \log T}{T}$$

with an absolute implicit constant.

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

Proof. Using Perron's Formula as in [**MV07**, Corollary 5·3]. we obtain a main term

$$\frac{1}{2i\pi}\int_{c-iX}^{c+iX} -\frac{L'(f,s)}{L(f,s)}x^s\frac{\mathrm{d}s}{s},$$

where we choose $c = 1 + 1/\log x$. Using Cauchy's residue Theorem, we write this integral as a sum over the zeros and poles of $L(f,s)$ and an integral that goes on the left of the critical strip (that can be bounded using bounds on the logarithmic derivative of the L-function close to the critical strip, see [**IK04**, Proposition 5·7(2)]) we obtain

$$\psi(f,x) + m\big(L(f,\cdot),1\big)x = -\sum_{\substack{L(f,\rho)=0\\ |\mathrm{Im}(\rho)|\leq X}} \frac{x^\rho}{\rho} \tag{4·3}$$
$$+ O\left(d\log x + \frac{x}{X}\left(d(\log x)^2 + \log\left(\mathfrak{q}(f)X^d\right)\right) + \log\left(\mathfrak{q}(f)X^d\right)\log X\right)$$

with an absolute implicit constant, see also [**IK04**, Chapter 5, Example 7].

Taking $X = x$ and cutting the sum at $T \leq x$, we obtain

$$\psi(f,x) + m\big(L(f,\cdot),1\big)x = -\left\{\sum_{\substack{L(f,\rho)=0\\ |\mathrm{Im}(\rho)|\leq T}} + \sum_{\substack{L(f,\rho)=0\\ T<|\mathrm{Im}(\rho)|\leq x}}\right\}\frac{x^\rho}{\rho} + O\big(d\log(\mathfrak{q}(f)x)\log x\big)$$

with an absolute implicit constant.

The first sum is our main term, we bound the second moment of the second sum. Define

$$\epsilon_f(x,T) := x^{-\beta_{f,0}}\sum_{\substack{\rho=\beta+i\gamma\\ L(f,\rho)=0\\ T<|\gamma|\leq x}}\frac{x^{\beta+i\gamma}}{\beta+i\gamma}.$$

The bound given in Proposition 4·4 follows from a generalization of [**Fio14a**, Lemma 3·3] and [**RS94**, Lemma 2·2].

As we do not assume that the Riemann Hypothesis holds, the sum in (4·2) is not obviously an almost periodic function in the sense of [**ANS14**]. We now decompose this sum to highlight the main term and bound the error.

LEMMA 4·5. *Let $L(f,s)$ be an analytic L-function, and let $T > 2$ be fixed. Define*

$$\beta_{f,T} = \sup\left\{\mathrm{Re}(\rho) : L(f,\rho) = 0,\ |\mathrm{Im}(\rho)| \leq T,\ \mathrm{Re}(\rho) < \beta_{f,0}\right\}.$$

One has

$$x^{-\beta_{f,0}}\sum_{\substack{\rho\\ L(f,\rho)=0\\ \mathrm{Im}(\rho)\leq T}}\frac{x^\rho}{\rho} = \sum_{\substack{\gamma\leq T\\ L(f,\beta_{f,0}+i\gamma)=0}}\frac{x^{i\gamma}}{\beta_{f,0}+i\gamma} + O\left(x^{\beta_{f,T}-\beta_{f,0}}d\log(\mathfrak{q}(f)T)\log T\right).$$

Remark 8. We use the conventions: $\sup\emptyset = -\infty$ and for $x > 0$ one has $x^{-\infty} = 0$.

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

Proof. Using [**IK04**, Proposition 5·7.(1)], we write

$$x^{-\beta_{f,0}} \sum_{\substack{\mathrm{Re}(\rho)<\beta_{f,0}\\ |\mathrm{Im}(\rho)|\le T}} \frac{x^\rho}{\rho} \ll x^{\beta_{f,T}-\beta_{f,0}} \sum_{\substack{\mathrm{Re}(\rho)<\beta_{f,0}\\ |\mathrm{Im}(\rho)|\le T}} \frac{1}{|\rho|} \ll x^{\beta_{f,T}-\beta_{f,0}} \log\left(\mathfrak{q}(f)T^d\right)\log T.$$

The implicit constant is absolute.

4·3. *Back to $E_f(x)$*

The study for $\psi(f,x)$ is now almost settled. However, $E_f(x)$ contains another term of potential equal interest.

LEMMA 4·6. *Let $L(f,s)$ be an analytic L-function, one has*

$$\theta(f,x) := \sum_{p\le x} \lambda_f(p)\log p = \psi(f,x) + m\big(L(f^{(2)},\cdot),1\big)x^{\frac12} + o_f(x^{\frac12}).$$

Proof. The Ramanujan–Petersson Conjecture and the Prime Number Theorem yield

$$\sum_{p\le x} \lambda_f(p)\log p = \psi(f,x) - \sum_{p^2\le x}\left(\sum_{j=1}^{d}\alpha_j(p)^2\right)\log p + O(dx^{1/3}).$$

To evaluate the second term, we use the Prime Number Theorem for the function $L(f^{(2)},s)$ (see e.g. [**IK04**, Theorem 5·13]). According to Definition 1·1(iii), its logarithmic derivative $L'(f^{(2)},s)/L(f^{(2)},s)$ extends meromorphically to the region $\{\mathrm{Re}(s)\ge 1-c/\log(|\mathrm{Im}(s)|+3)\}$, with no poles except a simple pole at $s=1$ with residue $-m\big(L(f^{(2)},\cdot),1\big)$. We obtain

$$\sum_{p^2\le x}\left(\sum_{j=1}^{d}\alpha_j(p)^2\right)\log p = -m\big(L(f^{(2)},\cdot),1\big)\sqrt{x} + o_f(\sqrt{x}).$$

Finally, using Stieltjes integral, we write $E_f(x) = \log x/x^{\beta_{f,0}}\int_2^x \mathrm{d}(\theta(f,t)+m\big(L(f,\cdot),1\big)t)/\log t$. Using integration by parts we obtain

$$x^{\beta_{f,0}}E_f(x) = \psi(f,x) + m\big(L(f,\cdot),1\big)x + m\big(L(f^{(2)},\cdot),1\big)x^{\frac12} + O\left(\log x\int_2^x \frac{\psi(f,t)+m\big(L(f,\cdot),1\big)t}{t(\log t)^2}\,\mathrm{d}t\right) + o_f(x^{\frac12}).$$

We use again an integration by parts to evaluate the O term. From (4·3), after integrating and taking $X=x$, we have

$$\int_2^x \big(\psi(f,t)+m\big(L(f,\cdot),1\big)t\big)\,\mathrm{d}t = -\sum_{\substack{L(f,\rho)=0\\ |\mathrm{Im}(\rho)|\le x}}\frac{x^{\rho+1}}{\rho(\rho+1)} + O_f\big(x(\log x)^2\big).$$

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

This series converges absolutely, so we can permute the limits. We deduce that the O term is $O_f(x^{\beta_{f,0}}/\log x)$. Hence we have

$$E_f(x) = \frac{1}{x^{\beta_{f,0}}}\left(\psi(f,x) + m\big(L(f,\cdot),1\big)x\right) + m\big(L(f^{(2)},\cdot),1\big)x^{\frac{1}{2}-\beta_{f,0}} + o_f(1). \tag{4·4}$$

4·4. *Existence of the limiting distribution*

We can now prove Proposition 4·1. In particular we prove the first point of Theorem 2·1: the existence of the limiting distribution for the function E_f.

Define (see Proposition 4·1),

$$G_{f,T}(x) = \frac{m\big(L(f,\cdot),\beta_{f,0}\big)}{\beta_{f,0}} + m\big(L(f^{(2)},\cdot),1\big)\delta_{\beta_{f,0}=\frac{1}{2}} - \sum_{\gamma\in\mathcal{Z}_f^*(T)} 2\,\mathrm{Re}\left(m\big(L(f,\cdot),\beta_{f,0}+i\gamma\big)\frac{x^{i\gamma}}{\beta_{f,0}+i\gamma}\right).$$

We use (4·4) where we evaluate $\psi(f,x)x^{-\beta_{f,0}}$ using Proposition 4·4 and Lemma 4·5. We can now write

$$E_f(x) = G_{f,T}(x) + O_f\left(x^{\beta_{f,T}-\beta_{f,0}}(\log T)^2\right) - \epsilon_f(x,T) + o_f(1), \tag{4·5}$$

where the second term vanishes if the Riemann Hypothesis is satisfied. We first prove that the real function $G_{f,T}$ admits a logarithmic distribution.

LEMMA 4·7. *Let $T > 2$ fixed. Then $G_{f,T}$ admits a limiting logarithmic distribution $\mu_{f,T}$.*

Proof. This follows from the generalized Kronecker–Weyl Theorem (see [**RS94**, Lemma 2·3], [**ANS14**, Proposition 2·4], or Theorem 4·2). We write $\mathcal{Z}_f^*(T) = \{\gamma_1,\dots,\gamma_{N(T)}\}$. Let $g:\mathbf{R}\to\mathbf{R}$ be a bounded Lipschitz continuous function, one can associate to g the continuous function on $\mathbf{T}^{N(T)}$ defined by

$$\tilde{g}(t) = g\left(m_f - 2\,\mathrm{Re}\left(\sum_{k=1}^{N(T)}\big(m(L(f,\cdot),\beta_{f,0}+i\gamma_k)\frac{e^{2i\pi t_k}}{\beta_{f,0}+i\gamma_k}\right)\right). \tag{4·6}$$

One has

$$\int_2^Y g(G_{f,T}(e^y))\,\mathrm{d}y = \int_2^Y \tilde{g}\left(\frac{\gamma_1}{2\pi}y,\dots,\frac{\gamma_{N(T)}}{2\pi}y\right)\mathrm{d}y.$$

Then we see that the measure $\mu_{f,T}$ is the push-forward measure of the normalized Haar measure on the closure of $\{(\gamma_1 y/2\pi,\dots,\gamma_{N(T)}y/2\pi) : y\in\mathbf{R}\}/\mathbf{Z}^{N(T)}$ in $\mathbf{T}^{N(T)}$.

Next, using (4·5) we see that E_f is a B^2-almost periodic function, hence by [**ANS14**, Theorem 2·9] it admits a limiting logarithmic distribution (see also [**KR03**, Theorem 1(e)]). In particular the proof uses the following inequality: for g a continuous C_g-Lipschitz bounded function, and $T>0$ fixed,

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

$$\int_{\mathbf{R}} g\,\mathrm{d}\mu_{f,T} + O_f\left(C_g\frac{\log T}{\sqrt{T}}\right) \le \liminf_{Y\to\infty}\frac{1}{Y}\int_2^Y g\big(E_f(e^y)\big)\,\mathrm{d}y$$

$$\le \limsup_{Y\to\infty}\frac{1}{Y}\int_2^Y g\big(E_f(e^y)\big)\,\mathrm{d}y \le \int_{\mathbf{R}} g\,\mathrm{d}\mu_{f,T} + O_f\left(C_g\frac{\log T}{\sqrt{T}}\right). \quad (4{\cdot}7)$$

The proof of Proposition 4·1 follows from Helly's Selection Theorem.

4·5. *Mean and variance*

We complete the proof of Theorem 2·1 by studying the decay of μ_f at infinity and computing its mean and variance.

Using (4·7) and information on the support of $\mu_{f,T}$ we can show that μ_f has exponential decay.

LEMMA 4·8. *There exists a positive constant $c(f)$ depending only on f such that*

$$\mu_f\Big(\mathbf{R} - \big[m_f - R, m_f + R\big]\Big) \ll_f e^{-c(f)\sqrt{R}}.$$

Proof. One has

$$\left|\sum_{\gamma\in\mathcal{Z}_f^*(T)} 2\,\mathrm{Re}\left(m\big(L(f,\cdot),\beta_{f,0}+i\gamma\big)\frac{x^{i\gamma}}{\beta_{f,0}+i\gamma}\right)\right| \le \sum_{\gamma\in\mathcal{Z}_f(T)} 2\frac{1}{|\beta_{f,0}+i\gamma|} \ll_f (\log T)^2$$

thanks to the fact that the completed L-function is of order 1 (see Definition 1·1(ii)). Therefore, the function $G_{f,T}(e^y)$ is bounded. We deduce that the measure $\mu_{f,T}$ has compact support included in $[m_f - c_1(\log T)^2, m_f + c_1(\log T)^2]$ for some constant c_1 depending on f. Using (4·7), we have

$$\mu_f\Big(\mathbf{R} - \big[m_f - c_1(\log T)^2, m_f + c_1(\log T)^2\big]\Big) = O_f\left(\frac{\log T}{\sqrt{T}}\right).$$

For $R = c_1(\log T)^2$ the result follows.

The measure μ_f has exponential decay at infinity, hence it has finite moments. The values for the mean and variance given in Theorem 2·1 follow from computations for $\mu_{f,T}$, letting T be arbitrarily large as we now explain, the proofs follow the ideas of [**Fio14a**]. Let $T \ge 2$ be fixed, one has

$$\int_{\mathbf{R}} t\,\mathrm{d}\mu_{f,T} = \lim_{Y\to\infty}\frac{1}{Y}\int_2^Y \left(m_f - \sum_{\gamma\in\mathcal{Z}_f(T)} 2\,\mathrm{Re}\left(\frac{e^{iy\gamma}}{\beta_{f,0}+i\gamma}\right)\right)\mathrm{d}y$$

$$= m_f - \lim_{Y\to\infty}\frac{1}{Y}O\left(\sum_{\gamma\in\mathcal{Z}_f(T)}\frac{1}{|\beta_{f,0}+i\gamma||\gamma|}\right) = m_f$$

because the sum over $\mathcal{Z}_f(T)$ is finite. Taking the limit as $T\to\infty$ the assertion on the average value of μ_f is proved.

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

For the computation of the variance, we cannot use the linearity anymore, we go back to the general case. Set

$$G_{\mathcal{S},T}(x) = m_{\mathcal{S}} - \sum_{\gamma \in \mathcal{Z}^*_{\mathcal{S}}(T)} 2\,\mathrm{Re}\left(M(\gamma)\frac{x^{i\gamma}}{\beta_{\mathcal{S},0}+i\gamma}\right),$$

where for γ in $\mathcal{Z}^*_{\mathcal{S}}$, we denote $M(\gamma) = \sum_{f\in\mathcal{S}} a_f m(L(f,\cdot), \beta_{\mathcal{S},0}+i\gamma)$. Then

$$\begin{aligned}\int_{\mathbf{R}} |t-m_{\mathcal{S}}|^2\,\mathrm{d}\mu_{\mathcal{S},T} &= \lim_{Y\to\infty}\frac{1}{Y}\int_2^Y \left|\sum_{\gamma\in\mathcal{Z}^*_{\mathcal{S}}(T)}\left(\frac{M(\gamma)e^{iy\gamma}}{\beta_0+i\gamma}+\frac{M(-\gamma)e^{-iy\gamma}}{\beta_0-i\gamma}\right)\right|^2 \mathrm{d}y\\ &= \lim_{Y\to\infty}\frac{1}{Y}\sum_{\gamma,\lambda}^{*}\frac{M(\gamma)\overline{M(\lambda)}}{(\beta_0+i\gamma)(\beta_0-i\lambda)}\int_2^Y e^{i(\gamma-\lambda)y}\,\mathrm{d}y,\end{aligned}$$

where the γ, λ are in the index set $\mathcal{Z}^*_{\mathcal{S}}(T)\cup(-\mathcal{Z}^*_{\mathcal{S}}(T))$ (counted without multiplicities). The diagonal term $\lambda=\gamma$ is the main term. The off-diagonal term vanishes as $Y\to\infty$ when T is fixed. One has

$$\sum_{\gamma\neq\lambda}^{*}\frac{M(\gamma)\overline{M(\lambda)}}{(\beta_{\mathcal{S},0}+i\gamma)(\beta_{\mathcal{S},0}-i\lambda)}\int_2^Y e^{i(\gamma-\lambda)y}\,\mathrm{d}y = O\left(\sum_{\gamma\neq\lambda}^{*}\frac{|M(\gamma)||M(\lambda)|}{|\gamma||\lambda|}\min\left(Y,|\gamma-\lambda|^{-1}\right)\right).$$

Using [**Fio14a**, Lemma 2·6], we deduce that

$$\int_{\mathbf{R}}|t-m_{\mathcal{S}}|^2\,\mathrm{d}\mu_{\mathcal{S}} = \sum_{\gamma\in\mathcal{Z}^*_{\mathcal{S}}}\frac{2|M(\gamma)|^2}{|\beta_{\mathcal{S},0}+i\gamma|^2}.$$

This concludes the proof of Theorem 2·1.

5. *Results under additional hypotheses*

It is clear from the proof of Theorem 2·1 that the properties of the set of zeros of largest real part of our L-functions in $\mathcal{S}$ are related to the properties of $\mu_{\mathcal{S}}$. In this section we investigate in more details what can be inferred from additional hypotheses on the zeros.

5·1. *Existence of the bias and regularity of the distribution*

We show that the existence of self-sufficient zeros in $\mathcal{Z}^*_{\mathcal{S}}$ gives properties of smoothness for $\mu_{\mathcal{S}}$. Such results were previously obtained (e.g. in [**RS94**]) conditionally on LI.

The first point we need to address is to understand more precisely the measures $\mu_{\mathcal{S},T}$ and for this we need to give more precisions on the sub-tori $A(\gamma_1,\dots,\gamma_{N(T)})$. We prove that such a sub-torus can be decomposed into products of sub-tori if there is some linear independence between the zeros.

PROPOSITION 5·1. *Let $\gamma_1,\dots,\gamma_M,\lambda_1,\dots,\lambda_N\in\mathbf{R}$ be real numbers satisfying*

$$\langle\gamma_1,\dots,\gamma_M\rangle_{\mathbf{Q}}\cap\langle\lambda_1,\dots,\lambda_N\rangle_{\mathbf{Q}} = \{0\}.$$

Then the closure of the 1-parameter group $\{y(\gamma_1,\dots,\gamma_M,\lambda_1,\dots,\lambda_N): y\in\mathbf{R}\}/\mathbf{Z}^{M+N}$ in the $M+N$-dimensional torus $\mathbf{T}^{M+N}$ is the Cartesian product of the closure of the 1-parameter group $\{y(\gamma_1,\dots,\gamma_M): y\in\mathbf{R}\}/\mathbf{Z}^M$ in the M-dimensional torus $\mathbf{T}^M$ and the closure of the 1-parameter group $\{y(\lambda_1,\dots,\lambda_N): y\in\mathbf{R}\}/\mathbf{Z}^N$ in the N-dimensional torus $\mathbf{T}^N$:

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

$$A(\gamma, \lambda) = A(\gamma) \times A(\lambda).$$

In particular, the normalized Haar measure over $A(\gamma, \lambda)$ is the product of the normalized Haar measures over the sub-tori:

$$\mathrm{d}\omega_{A(\gamma,\lambda)} = \mathrm{d}\omega_{A(\gamma)}\, \mathrm{d}\omega_{A(\lambda)}.$$

A corollary of this result is that we can deal with independent sets independently when taking the limiting distribution (for example we can use Fubini Theorem).

Proof. Using Lemma 4·3, we write that $A(\gamma, \lambda) = A(\gamma_1, \ldots, \gamma_M, \lambda_1, \ldots, \lambda_N)$ is the annihilator of the annihilator of the 1-parameter group $\{y(\gamma_1, \ldots, \gamma_M, \lambda_1, \ldots, \lambda_N) : y \in \mathbf{R}\}/\mathbf{Z}^{M+N}$ in the group $\mathbf{T}^{M+N}$.

So we first need to determine the annihilator:

$$\begin{aligned}
&\Big(\big\{y(\gamma_1, \ldots, \gamma_M, \lambda_1, \ldots, \lambda_N) : y \in \mathbf{R}\big\}/\mathbf{Z}^{M+N}\Big)^{\perp} \\
&\quad = \big\{(k_1, \ldots, k_M, \ell_1, \ldots, \ell_N) \in \mathbf{Z}^{M+N} : \gamma_1 k_1 + \cdots + \gamma_N k_M + \lambda_1 \ell_1 + \ldots \lambda_N \ell_N = 0\big\} \\
&= \big\{(k_1, \ldots, k_M, \ell_1, \ldots, \ell_N) \in \mathbf{Z}^{M+N} : \gamma_1 k_1 + \cdots + \gamma_N k_M = 0 \text{ and } \lambda_1 \ell_1 + \ldots \lambda_N \ell_N = 0\big\} \\
&\quad = A(\gamma_1, \ldots, \gamma_M)^{\perp} \times A(\lambda_1, \ldots, \lambda_N)^{\perp}.
\end{aligned}$$

The second equality follows from the linear independence. If for some $(k_1, \ldots, k_M, \ell_1, \ldots, \ell_N) \in \mathbf{Z}^{M+N}$ one has $\gamma_1 k_1 + \cdots + \gamma_N k_M + \lambda_1 \ell_1 + \cdots \lambda_N \ell_N = 0$ then $\gamma_1 k_1 + \cdots + \gamma_N k_M = \lambda_1 \ell_1 + \cdots \lambda_N \ell_N = 0$. The third equality is Lemma 4·3.

We conclude by taking the annihilator (thanks to Lemma 4·3 once again) and the fact that the annihilator of a product is the product of annihilators (see Lemma 5·2).

LEMMA 5·2. *Let $N_1, N_2 \geq 1$, let A_1 be a sub-torus of $\mathbf{T}^{N_1}$ and A_2 be a subgroup of $\mathbf{T}^{N_2}$. Then $A_1 \times A_2$ is a sub-torus of $\mathbf{T}^{N_1+N_2}$ and its annihilator is a subgroup of $\mathbf{Z}^{N_1+N_2}$ given by the product*

$$(A_1 \times A_2)^{\perp} = A_1^{\perp} \times A_2^{\perp}.$$

Proof. The proof follows essentially from [**Bou47**, Chapter VII Section 1, Proposition 8]. For any $(k_1, k_2) \in A_1^{\perp} \times A_2^{\perp}$ and any $(x_1, x_2) \in A_1 \times A_2$, one has

$$\big\langle (x_1, x_2), (k_1, k_2) \big\rangle = \langle x_1, k_1 \rangle + \langle x_2, k_2 \rangle = 0$$

hence $A_1^{\perp} \times A_2^{\perp} \subset (A_1 \times A_2)^{\perp}$. Now take $(k_1, k_2) \in (A_1 \times A_2)^{\perp}$, for all $(x_1, x_2) \in A_1 \times A_2$ one has

$$\langle x_1, k_1 \rangle + \langle x_2, k_2 \rangle = \big\langle (x_1, x_2), (k_1, k_2) \big\rangle = 0.$$

In particular ($0 \in A_1$), for all $x_2 \in A_2$,

$$\langle x_2, k_2 \rangle = \langle 0, k_1 \rangle + \langle x_2, k_2 \rangle = \big\langle (0, x_2), (k_1, k_2) \big\rangle = 0,$$

hence $k_2 \in A_2^{\perp}$. Similarly, $k_1 \in A_1^{\perp}$ and the proof is complete.

Our main contribution in the following result is that we get the existence of the logarithmic density $\delta(\mathcal{S})$ (as defined in Definition 1·3) under weaker hypotheses than LI. We use the concept of self-sufficient zeros introduced by Martin and Ng in [**MN17**].

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

Definition 5·3.

(i) An ordinate $\gamma \in \mathcal{Z}_{\mathcal{S}}^*$ is self-sufficient if it is not in the $\mathbf{Q}$-span of $\mathcal{Z}_{\mathcal{S}}^* - \{\gamma\}$.
(ii) For $U, V > 0$, we say that an ordinate $\gamma \in \mathcal{Z}_{\mathcal{S}}^*(U)$ is (U, V)-self-sufficient if it is not in the $\mathbf{Q}$-span of $\mathcal{Z}_{\mathcal{S}}^*(V) - \{\gamma\}$.

Using this concept we prove conditional results for the regularity of the distribution $\mu_{\mathcal{S}}$.

THEOREM 5·4. *Suppose that the set $\mathcal{S}$ satisfies the conditions of Theorem 2·1.*

(i) Suppose that there exists $\epsilon > 0$ and $T_\epsilon > 0$ such that for every $T \geq T_\epsilon$ the set $\mathcal{Z}_{\mathcal{S}}^*(T^{\frac{1}{2}-\epsilon})$ contains a $(T^{\frac{1}{2}-\epsilon}, T)$-self-sufficient zero γ_T. Then $\delta(\mathcal{S})$ exists.
(ii) Suppose $\mathcal{Z}_{\mathcal{S}}^*$ contains at least one self-sufficient element, then the distribution $\mu_{\mathcal{S}}$ is continuous (i.e. $\mu_{\mathcal{S}}$ assigns zero mass to finite sets).
(iii) Suppose $\mathcal{Z}_{\mathcal{S}}^*$ contains three or more self-sufficient elements, then the distribution $\mu_{\mathcal{S}}$ admits a density $\phi \in L^1$ (i.e. $\mathrm{d}\mu_{\mathcal{S}}(x) = \phi(x)\,\mathrm{d}x$).
(iv) Suppose that the set $\{\gamma \in \mathcal{Z}_{\mathcal{S}}^* : \gamma$ self-sufficient$\}$ is infinite, then the distribution $\mu_{\mathcal{S}}$ admits a density ϕ which is in the Schwartz space of indefinitely differentiable and rapidly decreasing functions.

Remark 9. This improves some results of [**RS94**] which are obtained under the Grand Simplicity Hypothesis (also called LI). In loc. cit. Rubinstein and Sarnak obtain Theorem 5·4 under LI, i.e. assuming that $\mathcal{Z}_{\mathcal{S}}$ is linearly independent over $\mathbf{Q}$. In Theorem 5·4 the Riemann Hypothesis is not needed, and hypotheses in (i)–(iv) are ordered by increasing strength but are all weaker than LI. In particular Theorem 5·4(i) gives a condition for the logarithmic density of the set $\{x \geq 2 : E_{\mathcal{S}}(x) \geq 0\}$ to exist, where the function $E_{\mathcal{S}}$ is as defined in Theorem 2·1.

Proof of Theorem 5·4(i). Fix $T \geq T_\epsilon$. Following [**RS94**, part 3·1], we compute the Fourier transform of $\mu_{\mathcal{S},T}$. We obtain

$$\hat{\mu}_{\mathcal{S},T}(\xi) = \int_{A_T} \exp\left(-i\xi\left(m_{\mathcal{S}} - 2\,\mathrm{Re}\left(\sum_{\gamma \in \mathcal{Z}_{\mathcal{S}}^*(T)} \frac{M(\gamma)e^{2i\pi t_\gamma}}{\beta_{\mathcal{S},0} + i\gamma}\right)\right)\right)\mathrm{d}t$$
$$= e^{-im_{\mathcal{S}}\xi} \int_{A_T} \prod_{\gamma \in \mathcal{Z}_{\mathcal{S}}^*(T)} \exp\left(i\xi 2\,\mathrm{Re}\left(M(\gamma)\frac{e^{2i\pi t_\gamma}}{\beta_{\mathcal{S},0} + i\gamma}\right)\right)\mathrm{d}t,$$

where A_T is the closure of $\{(\gamma_1 y, \ldots, \gamma_{N(T)} y) : y \in \mathbf{R}\}/\mathbf{Z}^{N(T)}$ in $\mathbf{T}^{N(T)}$. The ordinate γ_T is self-sufficient in $\mathcal{Z}_{\mathcal{S}}^*(T)$, hence by Proposition 5·1, one can write $A_T = \mathbf{T} \times A_T'$ and separate the integral:

$$\hat{\mu}_{\mathcal{S},T}(\xi) = e^{-im_{\mathcal{S}}\xi} \int_{A_T'} \prod_{\gamma \in \mathcal{Z}_{\mathcal{S}}^*(T) - \{\gamma_0\}} \exp\left(i\xi 2\,\mathrm{Re}\left(M(\gamma)\frac{e^{2i\pi t_\gamma}}{\beta_{\mathcal{S},0} + i\gamma}\right)\right)\mathrm{d}t$$
$$\int_{\mathbf{T}} \exp\left(i\xi 2\,\mathrm{Re}\left(M(\gamma_T)\frac{e^{2i\pi\theta}}{\beta_{\mathcal{S},0} + i\gamma_T}\right)\right)\mathrm{d}\theta.$$

The integral over $\mathbf{T}$ is a 0th Bessel function of the first kind equal to $J_0(|2\xi M(\gamma_T)|/|\beta_{\mathcal{S},0} + i\gamma_T|)$ Using properties of the Bessel function (see e.g. [**Wat95**]) and the fact that

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

the first integral on the right-hand side is bounded from above by 1, one can bound the Fourier transform of $\mu_{\mathcal{S},T}$:

$$|\hat{\mu}_{\mathcal{S},T}(\xi)| \leq \left| J_0\left(\left| \frac{2\xi M(\gamma_T)}{\beta_{\mathcal{S},0} + i\gamma_T} \right| \right) \right| \leq \min\left(1, \sqrt{\left| \frac{\beta_{\mathcal{S},0} + i\gamma_T}{\pi \xi M(\gamma_T)} \right|}\right). \tag{5·1}$$

Let us come back to the existence of $\delta(\mathcal{S})$. We want to prove that the limits

$$\limsup \frac{1}{Y} \int_2^Y \mathbf{1}_{\geq 0}(E_{\mathcal{S}}(e^y))\, \mathrm{d}y \qquad \text{and} \qquad \liminf \frac{1}{Y} \int_2^Y \mathbf{1}_{\geq 0}(E_{\mathcal{S}}(e^y))\, \mathrm{d}y$$

coincide. We write $\mathbf{1}_{\geq 0} = g_n + (\mathbf{1}_{\geq 0} - g_n)$ where g_n is the n-Lipschitz function satisfying

$$g_n(x) = \begin{cases} 0 \text{ if } x \leq -1/2n, \\ 1 \text{ if } x \geq 1/2n, \\ nx + 1/2 \text{ otherwise.} \end{cases}$$

The functions g_n and $|\mathbf{1}_{\geq 0} - g_n|$ are bounded, continuous, n-Lipschitz. Hence

$$\lim_{Y\to\infty} \frac{1}{Y} \int_2^Y g_n\big(G_{\mathcal{S},T}(e^y)\big)\, \mathrm{d}y = \int_{\mathbf{R}} g_n(t)\, \mathrm{d}\mu_{\mathcal{S},T}(t)$$

and we have the same result if we replace g_n by $|\mathbf{1}_{\geq 0} - g_n|$. Taking T arbitrarily large, we approximate the limiting distribution $\mu_{\mathcal{S}}$. Precisely one has

$$\begin{aligned} \int_{\mathbf{R}} g_n(t)\, \mathrm{d}\mu_{\mathcal{S},T} - \int_{\mathbf{R}} |\mathbf{1}_{\geq 0} - g_n|(t)\, \mathrm{d}\mu_{\mathcal{S},T} + O_{\mathcal{S}}\left(n \frac{\log T}{\sqrt{T}}\right) \\ \leq \liminf_{Y\to\infty} \frac{1}{Y} \int_2^Y \mathbf{1}_{\geq 0}(E_{\mathcal{S}}(e^y))\, \mathrm{d}y \leq \limsup_{Y\to\infty} \frac{1}{Y} \int_2^Y \mathbf{1}_{\geq 0}(E_{\mathcal{S}}(e^y))\, \mathrm{d}y \\ \leq \int_{\mathbf{R}} g_n(t)\, \mathrm{d}\mu_{\mathcal{S},T} + \int_{\mathbf{R}} |\mathbf{1}_{\geq 0} - g_n|(t)\, \mathrm{d}\mu_{\mathcal{S},T} + O_{\mathcal{S}}\left(n \frac{\log T}{\sqrt{T}}\right). \end{aligned} \tag{5·2}$$

Moreover, we bound $\mu_{\mathcal{S},T}(|\mathbf{1}_{\geq 0} - g_n|)$. Using Parseval's formula (see e.g. [**Kat04**, Theorem VI·2·2]) we have:

$$\begin{aligned} \int_{\mathbf{R}} |\mathbf{1}_{\geq 0} - g_n|\, \mathrm{d}\mu_{\mathcal{S},T} &= \int_{\mathbf{R}} 2n \frac{1 - \cos(\xi/2n)}{\xi^2} \hat{\mu}_{\mathcal{S},T}(\xi)\, \mathrm{d}\xi \\ &\ll \int_{|\xi|\leq\alpha(n)} \frac{1}{2n} |\hat{\mu}_{\mathcal{S},T}(\xi)|\, \mathrm{d}\xi + \int_{|\xi|\geq\alpha(n)} \frac{4n}{\xi^2} |\hat{\mu}_{\mathcal{S},T}(\xi)|\, \mathrm{d}\xi \end{aligned}$$

for $\alpha(n) < 2n$. Using (5·1), we get:

$$\int_{\mathbf{R}} |\mathbf{1}_{\geq 0} - g_n|\, \mathrm{d}\mu_{\mathcal{S},T} \ll \frac{2\alpha(n)}{2n} + \int_{|\xi|\geq\alpha(n)} \frac{4n\sqrt{\gamma_T}}{|\xi|^{5/2}}\, \mathrm{d}\xi \ll \frac{\alpha(n)}{n} + \frac{n\sqrt{\gamma_T}}{\alpha(n)^{3/2}}.$$

Choose $n = \sqrt{T^{1-\epsilon}}$ and $\alpha(n) = n^{1-\frac{\epsilon}{3}}$. Since $\gamma_T \leq T^{\frac{1}{2}-\epsilon}$, the terms of rest in (5·2) vanish as $T \to +\infty$. It ensures that the inferior and superior limits coincide.

To prove the other points of Theorem 5·4 we follow the same idea without dependence on T.

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

Proof of Theorem 5·4(ii). The fact that $\mu_{\mathcal{S}}$ is continuous is a consequence of a theorem of Wiener (see e.g. [**Kat04**, Theorem VI·2·11]). A necessary and sufficient condition for $\mu_{\mathcal{S}}$ to be continuous is:

$$\lim_{Y\to\infty} \frac{1}{2Y} \int_{-Y}^{Y} |\hat{\mu}_{\mathcal{S}}(\xi)|^2 \, \mathrm{d}\xi = 0. \tag{5·3}$$

In the case $\gamma_T = \gamma_0$ does not depend on T, the bound (5·1) becomes, for all $T > \gamma_0$,

$$|\hat{\mu}_{\mathcal{S},T}(\xi)| \leq \min\left(1, \sqrt{\left|\frac{\beta_{\mathcal{S},0} + i\gamma_0}{\pi\xi M(\gamma_0)}\right|}\right).$$

Letting $T \to \infty$, the same bound holds for $\hat{\mu}_{\mathcal{S}}$. In particular condition (5·3) holds.

Proof of Theorem 5·4(iii). Let $\gamma_1 < \gamma_2 < \gamma_3$ be three self-sufficient elements of $\mathcal{Z}_{\mathcal{S}}^*$. Following the lines of the previous proofs, we get that for all $T > \gamma_3$ one has

$$|\hat{\mu}_{\mathcal{S},T}(\xi)| \leq \prod_{j=1}^{3} \min\left(1, \sqrt{\left|\frac{\beta_{\mathcal{S},0} + i\gamma_j}{\pi\xi M(\gamma_j)}\right|}\right).$$

Letting $T \to \infty$, the same bound holds for $\hat{\mu}_{\mathcal{S}}$. In particular one has $\hat{\mu}_{\mathcal{S}} \in L^1 \cap L^2$, Theorem 5·4(iii) follows by Fourier inversion.

Proof of Theorem 5·4(iv). As in the previous proofs we can write for all $T > 0$, for all ξ,

$$|\hat{\mu}_{\mathcal{S},T}(\xi)| \leq \prod_{\substack{\gamma\in\mathcal{Z}_{\mathcal{S}}^*(T)\\ \text{self-sufficient}}} \min\left(1, \sqrt{\left|\frac{\beta_{\mathcal{S},0} + i\gamma}{\pi\xi M(\gamma)}\right|}\right).$$

We assume that there are infinitely many self-sufficient elements in $\mathcal{Z}_{\mathcal{S}}^*$. For each $n \in \mathbf{N}$ there exists $T_n > 0$ such that

$$|\{\gamma \in \mathcal{Z}_{\mathcal{S}}^*(T_n) : \gamma \text{ self-sufficient }\}| \geq 2n + 1.$$

Hence, there exists a constant C_n depending only on $\mathcal{Z}_{\mathcal{S}}^*(T_n)$ such that for every $T \geq T_n$, for every ξ large enough (in terms of T_n) one has

$$|\hat{\mu}_{\mathcal{S},T}(\xi)| \leq \frac{C_n}{|\xi|^n \sqrt{|\xi|}}.$$

Letting $T \to \infty$, the same bound holds for $\hat{\mu}_{\mathcal{S}}$. In particular one has $|\xi^n \hat{\mu}_{\mathcal{S}}(\xi)| \to 0$ as $|\xi| \to \infty$. By Fourier inversion, we obtain that the density of $\mu_{\mathcal{S}}$ is n times differentiable for all $n \geq 0$.

The statement about fast decay is a consequence of the exponential decrease obtained in Theorem 2·1.

In the previous proofs we have used the decay at infinity of the Bessel 0th function J_0 to obtain the bounds for $\hat{\mu}_{\mathcal{S}}$. Using the theory of oscillatory integrals we can deduce the decay of other functions. Actually, we can verify condition (5·3) under a weaker hypothesis, that leads us to Theorem 2·2.

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

Proof of Theorem 2·2. The hypothesis and Proposition 5·1 imply that for every $T \geq \max\{\lambda_j : 1 \leq j \leq n\}$, the sub-torus A_T given by the closure of the 1-parameter group $\left\{y\left(\gamma_1, \ldots, \gamma_{N(T)}\right) : y \in \mathbf{R}\right\}/\mathbf{Z}^{N(T)}$ can be written as a Cartesian product $A(\lambda_1, \ldots, \lambda_n) \times A'_T$ where the two components are respectively the sub-torus associated with the set $\{\lambda_1, \ldots, \lambda_n\}$ and with the set $\mathcal{Z}^*_{\mathcal{S}}(T) - \{\lambda_1, \ldots, \lambda_n\}$ (see Theorem 4·2). In particular for $T \geq \max\{\lambda_j : 1 \leq j \leq n\}$, the Fourier Transform of $\mu_{\mathcal{S},T}$ is

$$\hat{\mu}_{\mathcal{S},T}(\xi) = e^{-im_{\mathcal{S}}\xi} \int_{A'_T} \exp\left(2i\xi \sum_{\gamma \in \mathcal{Z}^*_{\mathcal{S}}(T) - \{\lambda_1, \ldots, \lambda_n\}} \mathrm{Re}\left(M(\gamma)\frac{e^{2i\pi t_\gamma}}{\beta_{\mathcal{S},0} + i\gamma}\right)\right) \mathrm{d}\omega_{A'_T}(t)$$
$$\times \int_{A(\lambda_1,\ldots,\lambda_n)} \exp\left(2i\xi \sum_{j=1}^{n} \mathrm{Re}\left(M(\lambda_j)\frac{e^{2i\pi\theta_j}}{\beta_{\mathcal{S},0} + i\lambda_j}\right)\right) \mathrm{d}\omega_{A(\lambda_1,\ldots,\lambda_n)}(\theta).$$

We deduce that

$$|\hat{\mu}_{\mathcal{S},T}(\xi)| \leq \left| \int_{A(\lambda_1,\ldots,\lambda_n)} \exp\left(2i\xi \sum_{j=1}^{n} \left|\frac{M(\lambda_j)}{\beta_{\mathcal{S},0} + i\lambda_j}\right| \cos(2\pi\theta_j + w_j)\right) \mathrm{d}\omega_{A(\lambda_1,\ldots,\lambda_n)}(\theta) \right|, \tag{5·4}$$

where w_j is defined by $M(\lambda_j)/(\beta_{\mathcal{S},0} + i\lambda_j) = |M(\lambda_j)/(\beta_{\mathcal{S},0} + i\lambda_j)|e^{iw_j}$. The right-hand side of inequality (5·4) does not depend on T, hence for ξ fixed, we can let $T \to \infty$ and obtain the same inequality for $\hat{\mu}_{\mathcal{S}}(\xi)$.

Thus, we only need to prove that the right-hand side of (5·4) approaches 0 when $|\xi| \to \infty$ to ensure condition (5·3) and thus the continuity of $\mu_{\mathcal{S}}$. The function

$$\phi : \begin{cases} A(\lambda_1, \ldots, \lambda_n) \longrightarrow \mathbf{R} \\ (\theta_1, \ldots, \theta_n) \quad \longmapsto 2\sum_{j=1}^{n} \left|\frac{M(\lambda_j)}{\beta_{\mathcal{S},0} + i\lambda_j}\right| \cos(2\pi\theta_j + w_j) \end{cases}$$

is a non-constant analytic function on a non-discrete compact set (this is a particular case of Lemma 5·5 below). Hence, there exists $K \geq 1$ such that for each $\theta \in A(\lambda_1, \ldots, \lambda_n)$ there exists a multi-index $\underline{k}$ of length $1 \leq |\underline{k}| \leq K$ satisfying $\partial^{\underline{k}}\phi(\theta) \neq 0$. By [**Ste93**, VIII 2·2 Proposition 5], one has

$$|\hat{\mu}_{\mathcal{S}}(\xi)| \leq \left| \int_{\mathbf{T}(\lambda_1,\ldots,\lambda_n)} \exp\left(i\xi 2 \sum_{j=1}^{n} \left|\frac{M(\lambda_j)}{\beta_{\mathcal{S},0} + i\lambda_j}\right| \cos(2\pi\theta_j + w_j)\right) \mathrm{d}\omega_{\mathbf{T}(\lambda_1,\ldots,\lambda_n)}(\theta) \right| \ll |\xi|^{-1/K}$$

with an implicit constant depending on K, ϕ, and on a choice of partition of unity adapted to ϕ. In particular condition (5·3) holds, hence $\mu_{\mathcal{S}}$ is continuous. Theorem 2·2 is proved by the same argument as in Theorem 5·4.

We finish this part with the statement of the following general result that we used in the proof of Theorem 2·2.

LEMMA 5·5. *Let $n \geq 1$ be an integer, $\lambda_1, \ldots, \lambda_n > 0$ be distinct real numbers, and let $A(\lambda_1, \ldots, \lambda_n)$ be the the sub-torus of $\mathbf{T}^n$ given by the closure of the 1-parameter group $\{y\,(\lambda_1, \ldots, \lambda_n) : y \in \mathbf{R}\}/\mathbf{Z}^n$. Let $(M_j)_{1\leq j\leq n}, (w_j)_{1\leq j\leq n} \in \mathbf{C}^n$, such that for at least one $1 \leq j \leq n$, one has $M_j \neq 0$. Then the function*

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

$$\phi_{M,w} : \begin{cases} A(\lambda_1, \dots, \lambda_n) \longrightarrow \mathbf{R} \\ (\theta_1, \dots, \theta_n) \quad \longmapsto \sum_{j=1}^{n} M_j \cos(2\pi\theta_j + w_j) \end{cases}$$

is non-constant and analytic.

Proof. The function ϕ is analytic as a finite sum of analytic functions. It is defined as a linear combination of characters of $A(\lambda_1, \dots, \lambda_n)$ seen as a subset of $\mathbf{T}^n$. Precisely, for $1 \leq j \leq n$ we denote by χ_j (resp. χ_j') the character $A(\lambda_1, \dots, \lambda_n) \ni (\theta_1, \dots, \theta_n) \mapsto e^{2i\pi\theta_j}$ (resp. $\mapsto e^{-2i\pi\theta_j}$). Let us study the values on $\{y\,(\lambda_1, \dots, \lambda_n) : y \in \mathbf{R}\}/\mathbf{Z}^n$, and take the derivative at $y = 0$, we obtain distinct values $\lambda_j, -\lambda_j$, all of them distinct from 0. Hence, the $2n+1$ characters $1, \chi_1, \chi_1', \dots, \chi_n, \chi_n'$ are distinct. By a result of Dedekind–Artin [**Lan02**, VI, Theorem 4·1], the characters are linearly independent over $\mathbf{C}$. Given $(M_j)_{1\leq j\leq n}, (w_j)_{1\leq j\leq n} \in \mathbf{C}^n$, one has that

$$\phi_{M,w} = \sum_{j=1}^{n} \frac{M_j}{2} \left(e^{iw_j}\chi_j + e^{-iw_j}\chi_j'\right)$$

is linearly independent from 1, as soon as one M_j is non-zero. In particular the function $\phi_{M,w}$ is non constant, and the proof is complete.

5·2. *Symmetry*

We prove Theorem 2·3 by showing that under its hypotheses, for every T, the distribution $\mu_{\mathcal{S},T}$ is symmetric with respect to its mean value. For this we use the Kronecker–Weyl Theorem (Theorem 4·2), and the following result.

LEMMA 5·6. *The following assertions are equivalent*:

(i) *for all integral linear combination*

$$\sum_{\gamma\in\mathcal{Z}_\mathcal{S}} k_\gamma \gamma = 0, k_\gamma \in \mathbf{Z}, \text{ one has } \sum_{\gamma\in\mathcal{Z}_\mathcal{S}} k_\gamma \equiv 0 \,[\mathrm{mod}\ 2];$$

(ii) *for every finite subset* $\{\gamma_1, \dots \gamma_N\} \subset \mathcal{Z}_\mathcal{S}$, *the element* $\left(\frac{1}{2}, \dots, \frac{1}{2}\right)$ *is in the closure of the one parameter group* $\{(\gamma_1 y, \dots, \gamma_N y) : y \in \mathbf{R}\}/\mathbf{Z}^N$ *in* $\mathbf{T}^N$.

Proof. This is a consequence of Lemma 4·3, one has $(1/2, \dots, 1/2) \in A(\gamma_1, \dots, \gamma_N)$ if and only if for every $k \in A(\gamma)^\perp$ one has $\sum_{i=1}^{N} k_i = 0 \,[\mathrm{mod}\ 2]$

Remark 10. In the formulation of Lemma 5·6, LI is equivalent to the fact that the closure of the one parameter group generated by a finite number of ordinates is always the largest possible (i.e. the N-dimensional torus when there are N ordinates). The improvement in Theorem 2·3, is that we only need to know that the element $(1/2, \dots, 1/2)$ is in this group to obtain the symmetry.

Proof of Theorem 2·3. By Theorem 4·2 and Lemma 5·6, we deduce that for all T large enough, one has $A_T = A_T + (1/2, \dots, 1/2)$. This way we can change variables $a \to a + (1/2, \dots, 1/2)$ in the integral defining $\mu_{\mathcal{S},T}$. For every $T > 1$, and for every bounded Lipschitz continous function g one has

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

$$\int_{\mathbf{R}} g(t)\,\mathrm{d}\mu_{\mathcal{S},T} = \int_{A_T} \tilde{g}(a)\,\mathrm{d}\omega_{A_T}$$
$$= \int_{A_T} \tilde{g}\left(a + \left(\frac{1}{2}, \ldots, \frac{1}{2}\right)\right)\mathrm{d}\omega_{A_T} = \int_{\mathbf{R}} g(2m_f - t)\,\mathrm{d}\mu_{\mathcal{S},T},$$

where we use the definition of $\tilde{g}$ given in (4·6). One has $\tilde{g}(a + (1/2, \ldots, 1/2)) = \tilde{h}(a)$ where h is the function given by $h(x) = g(2m_f - x)$. Take T arbitrarily large: by (4·7) the property of symmetry is true for $\mu_{\mathcal{S}}$.

Remark 11. In particular, under the condition of Theorem 2·3 for the set of the zeros of maximal real part associated to a set $\mathcal{S}$ of L-functions, we deduce that if the prime number race associated with $\mathcal{S}$ is biased it implies that the average value is not 0. So if the prime number race is biased, either RH is satisfied, or at least one of the L-functions vanishes at a point of $[1/2, 1)$.

5·3. *Riemann Hypothesis and support*

We prove Theorem 2·5, the first point is a generalization of [**RS94**, Theorem 1·2].

Proof of Theorem 2·5(i). The proof follows from an adaptation of [**RS94**, 2·2]. The idea is to find a lower bound for the measure of the set $\{y \leq M : E_{\mathcal{S}}(e^y) \geq A\}$ as M varies. One can also see [**Dev17**, Theorem 3·40] for a very detailed proof.

In the case RH is not satisfied ($\beta_{\mathcal{S},0} > 1/2$), one can conjecture that $\mathcal{Z}_{\mathcal{S}}$ is not too large. In particular, it may have density equal to 0 (in the set of all zeros). This is the point of Zero Density Theorems, and in particular of the condition in Theorem 2·5(ii).

Proof of Theorem 2·5(ii). We assume that the sum $\sum_{\gamma\in\mathcal{Z}_{\mathcal{S}}} 1/|\beta_{\mathcal{S},0} + i\gamma|$ converges, hence for every T, the limiting logarithmic distribution $\mu_{\mathcal{S},T}$ has compact support included in the interval

$$\left[m_{\mathcal{S}} - \sum_{\gamma\in\mathcal{Z}_{\mathcal{S}}} \frac{1}{|\beta_{\mathcal{S},0} + i\gamma|}, m_{\mathcal{S}} + \sum_{\gamma\in\mathcal{Z}_{\mathcal{S}}} \frac{1}{|\beta_{\mathcal{S},0} + i\gamma|}\right].$$

By Proposition 4·1, $\mu_{\mathcal{S}}$ has compact support included in the same interval.

Remark 12. Theorem 2·5(ii) could indicate a way to find completely biased prime number races: in the case one has an L-function with a zero β_0 in the interval $(1/2, 1)$ (e.g. a Siegel zero), and such that it has no zeros with larger real part, we can imagine that there will not be many other zeros of maximal real part (if ever they exist). For example if

$$\sum_{\gamma\in\mathcal{Z}_L} \frac{1}{|\beta_0 + i\gamma|} \leq m(L, \beta_0)$$

then we would have $\delta(L) = 0$. But the existence of such an L-function seems very unlikely.

5·4. *Concentration of the distribution*

In the process of looking for large biases, a strategy is to ensure that the distribution is concentrated around its average value (the average value being itself far from zero). Such concentration results are usually consequences of Chebyshev's inequality (e.g. [**Bas95**, Proposition (1·2)]).

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

LEMMA 5·7 (Chebyshev's inequality). *Let X be a random variable with average $\mathbb{E}(X)$ and finite variance $Var(X)$. For any $\alpha > 0$, one has*:

$$\mathbf{P}(|X - \mathbb{E}(X)| \geq \alpha) \leq \frac{Var(X)}{\alpha^2}.$$

Using Theorem 2·1, we can compute the average and variance of a random variable that has distribution equal to $\mu_{\mathcal{S}}$. With upper bounds for the variance or lower bounds for the average, we deduce concentration results (hence bounds for the bias).

COROLLARY 5·8.

(i) *In the case $m_{\mathcal{S}} < 0$ one has:* $\overline{\delta}(\mathcal{S}) \leq (2/m_{\mathcal{S}}^2) \sum_{\gamma \in \mathcal{Z}_{\mathcal{S}}^*} |M(\gamma)|^2/(\beta_{\mathcal{S},0}^2 + \gamma^2)$.
(ii) *In the case $m_{\mathcal{S}} > 0$ one has:* $\underline{\delta}(\mathcal{S}) \geq 1 - (2/m_{\mathcal{S}}^2) \sum_{\gamma \in \mathcal{Z}_{\mathcal{S}}^*} |M(\gamma)|^2/(\beta_{\mathcal{S},0}^2 + \gamma^2)$.

In particular if the ratio $\frac{2}{m_{\mathcal{S}}^2} \sum_{\gamma \in \mathcal{Z}_{\mathcal{S}}^*} \frac{|M(\gamma)|^2}{(\beta_{\mathcal{S},0}^2 + \gamma^2)}$ is small ($< 1/2$), we obtain the second version of Corollary 2·4.

Proof. The proof is inspired from [**Fio14a**, Lemma 2·7]. Assume $m_{\mathcal{S}} < 0$, one has

$$\overline{\delta}(\mathcal{S}) = \limsup \frac{1}{Y} \int_2^Y \mathbf{1}_{\geq 0}(E_{\mathcal{S}}(e^y))\, \mathrm{d}y \leq \limsup \frac{1}{Y} \int_2^Y g_n(E_{\mathcal{S}}(e^y))\, \mathrm{d}y,$$

where for $n \in \mathbf{N}$, the function g_n is continuous n-Lipschitz, bounded and has support $[-1/n, +\infty)$. Therefore one has

$$\limsup \frac{1}{Y} \int_2^Y g_n(E_{\mathcal{S}}(e^y))\, \mathrm{d}y = \mu_{\mathcal{S}}(g_n) \leq \mu_{\mathcal{S}}\left[\frac{-1}{n}, +\infty\right) = \mathbf{P}\left(X \geq \frac{-1}{n}\right)$$

for X a random variable of law $\mu_{\mathcal{S}}$. For n large enough, so that $m_{\mathcal{S}} + 1/n < 0$, we apply Chebyshev's inequality

$$\mathbf{P}\left(X \geq \frac{-1}{n}\right) \leq \mathbf{P}\left(|\mathbb{E}(X) - X| \geq \left|m_{\mathcal{S}} + \frac{1}{n}\right|\right) \leq \frac{\mathrm{Var}(X)}{\left|m_{\mathcal{S}} + \frac{1}{n}\right|^2}.$$

Letting $n \to \infty$ and using the values obtained in Theorem 2·1 for the average value and the variance yields the result. The case $m_{\mathcal{S}} > 0$ follows from similar computations.

Acknowledgements. This paper contains some results of my doctoral dissertation [**Dev17**]. I am very grateful to my advisor Florent Jouve for suggesting the problem, for all his advice, help and time spent correcting the first drafts of this paper. I thank Daniel Fiorilli and Étienne Fouvry for their interest in this work and for the many conversations that have led to improvements in the results. This work has also benefited from conversations with Farrell Brumley, Gaëtan Chenevier, Andrew Granville, Emmanuel Kowalski, Greg Martin, Philippe Michel, Nathan Ng, Zeév Rudnick and Mikołaj Frączyk as well as the anonymous referees. I thank Jean–Pierre Serre for his interest in this work, for his advice and for providing the proof of Lemma 5·5 as well as a more natural approach to Kronecker–Weyl equidistribution theorem. The computations presented in this document have been performed with the `SageMath` [**Sag16**] software.

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

REFERENCES

[ANS14] A. AKBARY, N. NG and M. SHAHABI. Limiting distributions of the classical error terms of prime number theory. *Q. J. Math.* **65** (2014), no. 3, 743–780.

[Bas95] R. F. BASS. Probabilistic techniques in analysis. *Probability and its Applications* (New York), (Springer–Verlag, New York, 1995).

[BCDT01] C. BREUIL, B. CONRAD, F. DIAMOND and R. TAYLOR. On the modularity of elliptic curves over **Q**: wild 3-adic exercises. *J. Amer. Math. Soc.* **14** (2001), no. 4, 843–939.

[BF90] D. BUMP and S. FRIEDBERG. *The exterior square automorphic L-functions on* GL(n). Festschrift in honor of I. I. Piatetski–Shapiro on the occasion of his sixtieth birthday, Part II (Ramat Aviv, 1989), Israel Math. Conf. Proc., vol. 3 (Weizmann, Jerusalem, 1990), pp. 47–65.

[BG92] D. BUMP and D. GINZBURG. Symmetric square L-functions on GL(r). *Ann. of Math.* (2) **136** (1992), no. 1, 137–205.

[Bou47] N. BOURBAKI. Éléments de mathématique. Première partie: Les structures fondamentales de l'analyse. Livre III: Topologie générale. Chapitre V: Groupes à un paramètre. Chapitre VI: Espace numériques et espaces projectifs. Chapitre VII: Les groupes additifs R^n. Chapitre VIII: Nombres complexes. *Actualités Sci. Ind., no. 1029*, (Hermann et Cie., Paris, 1947).

[CHT08] L. CLOZEL, M. HARRIS and R. TAYLOR. Automorphy for some l-adic lifts of automorphic mod l Galois representations. *Publ. Math. Inst. Hautes Études Sci.* (2008), no. 108, 1–181. With Appendix A, summarising unpublished work of Russ Mann, and Appendix B by Marie–France Vignéras.

[Cog04] J. W. COGDELL. Lectures on L-functions, converse theorems and functoriality for GL_n, *Lectures on automorphic L-functions*. Fields Inst. Monogr., vol. 20, (Amer. Math. Soc., Providence, RI, 2004), pp. 1–96.

[Con05] K. CONRAD. Partial Euler products on the critical line. *Canad. J. Math.* **57** (2005), no. 2, 267–297.

[CPS04] J. W. COGDELL and I. I. PIATETSKI–SHAPIRO. Remarks on Rankin–Selberg Convolutions, *Contributions to automorphic forms, geometry, and number theory*. (University Johns Hopkins Univ. Press, Baltimore, MD, 2004), pp. 255–278.

[Del74] P. DELIGNE. La conjecture de Weil. I. *Inst. Hautes Études Sci. Publ. Math.* (1974), no. 43, 273–307.

[Dev17] L. DEVIN. Propriétés algébriques et analytiques de certaines suites indexées par les nombres premiers. PhD. thesis. Université Paris-Sud 11, Université Paris–Saclay, (2017).

[DS74] P. DELIGNE and J.–P. SERRE. Formes modulaires de poids 1. *Ann. Sci. École Norm. Sup.* (4) **7** (1974), 507–530 (1975).

[Fio14a] D. FIORILLI. Elliptic curves of unbounded rank and Chebyshev's bias. *Int. Math. Res. Not. IMRN* (2014), no. 18, 4997–5024.

[Fio14b] D. FIORILLI. Highly biased prime number races. *Algebra Number Theory* **8** (2014), no. 7, 1733–1767.

[FK02] K. FORD and S. KONYAGIN. *Chebyshev's conjecture and the prime number race*, IV International Conference "Modern Problems of Number Theory and its Applications": Current Problems, Part II (Russian) (Tula, 2001). Mosk. Gos. Univ. im. Lomonosova, Mekh.-Mat. Fak., (Moscow, 2002), pp. 67–91.

[FM13] D. FIORILLI and G. MARTIN. Inequities in the Shanks-Rényi prime number race: an asymptotic formula for the densities. *J. Reine Angew. Math.* **676** (2013), 121–212.

[GJ72] R. GODEMENT and H. JACQUET. Zeta functions of simple algebras. *Lecture Notes in Math. vol.* 260. (Springer-Verlag, Berlin-New York, 1972).

[GM06] A. GRANVILLE and G. MARTIN. Prime number races. *Amer. Math. Monthly* **113** (2006), no. 1, 1–33.

[Gol82] D. GOLDFELD. Sur les produits partiels eulériens attachés aux courbes elliptiques. *C. R. Acad. Sci. Paris Sér. I Math.* **294** (1982), no. 14, 471–474.

[Har09] M. HARRIS. Potential automorphy of odd-dimensional symmetric powers of elliptic curves and applications. *Algebra, arithmetic, and geometry*: in honor of Yu. I. Manin. Vol. II, Progr. Math., vol. 270. (Birkhäuser Boston, Inc., Boston, MA, 2009), pp. 1–21.

[HSBT10] M. HARRIS, N. SHEPHERD–BARRON and R. TAYLOR. A family of Calabi-Yau varieties and potential automorphy. *Ann. of Math.* (2) **171** (2010), no. 2, 779–813.

[Hum10] P. HUMPHRIES. *The summatory function of Liouville's function and Pólya's conjecture*. (October 2010).

[IK04] H. IWANIEC and E. KOWALSKI. Analytic number theory. *American Mathematical Society Colloquium Publications, vol. 53*. (American Mathematical Society, Providence, RI, 2004).

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press

[JS90] H. JACQUET and J. SHALIKA. Exterior square L-functions. *Automorphic forms, Shimura varieties and L-functions*, Vol. II (Ann Arbor, MI, 1988), Perspect. Math., vol. 11 (Academic Press, Boston, MA, 1990), pp. 143–226.

[Kac95] J. KACZOROWSKI. On the distribution of primes (mod 4). *Analysis* **15** (1995), no. 2, 159–171.

[Kat04] Y. KATZNELSON. *An introduction to harmonic analysis*, third ed. (Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2004).

[KM00a] E. KOWALSKI and P. MICHEL. Explicit upper bound for the (analytic) rank of $J_0(q)$. *Israel J. Math.* **120** (2000), no. part A, 179–204.

[KM00b] E. KOWALSKI. A lower bound for the rank of $J_0(q)$. *Acta Arith.* **94** (2000), no. 4, 303–343.

[KP03] J. KACZOROWSKI and A. PERELLI. On the prime number theorem for the Selberg class. *Archiv der Mathematik* **80** (2003), no. 3, 255–263.

[KR03] J. KACZOROWSKI and O. RAMARÉ. Almost periodicity of some error terms in prime number theory. *Acta Arith*. **106** (2003), no. 3, 277–297.

[Lan02] S. LANG. *Algebra*, third ed. Graduate Texts in Math. vol. 211 (Springer-Verlag, New York, 2002).

[Maz08] B. MAZUR. Finding meaning in error terms. *Bull. Amer. Math. Soc*. (N.S.) **45** (2008), no. 2, 185–228.

[MN17] G. MARTIN and N. NG. Inclusive prime number races. arXiv:1710.00088. (Oct. 2017).

[MV07] H. L. MONTGOMERY and R. C. VAUGHAN. Multiplicative number theory. I. Classical theory. Camb. Stud. Adv. Math., vol. 97. (Cambridge University Press, Cambridge, 2007).

[MW89] C. MŒGLIN and J.–L. WALDSPURGER. Le spectre résiduel de GL(n). *Ann. Sci. École Norm. Sup.* (4) **22** (1989), no. 4, 605–674.

[Ng00] N. NG. Limiting distributions and zeros of Artin L-functions. PhD. thesis University of British Colombia. (2000).

[Ram00] D. RAMAKRISHNAN. Modularity of the Rankin-Selberg L-series, and multiplicity one for SL(2). *Ann. of Math*. (2) **152** (2000), no. 1, 45–111.

[RS94] M. RUBINSTEIN and P. SARNAK. Chebyshev's bias. *Experiment. Math.* **3** (1994), no. 3, 173–197.

[Rud80] W. RUDIN. *Analyse Réelle et Complexe*. (Masson, Paris, 1980). Translated from the first English edition by N. Dhombres and F. Hoffman, Third printing.

[Rud90] W. RUDIN. *Fourier analysis on groups*. (Wiley Classics Library, John Wiley & Sons, Inc., New York, 1990). Reprint of the 1962 original, a Wiley-Interscience Publication.

[RV99] D. RAMAKRISHNAN and R. J. VALENZA. Fourier analysis on number fields. *Graduate Texts in Mathematics. vol. 186.* (Springer-Verlag, New York, 1999).

[Sag16] Sagemath, the Sage Mathematics Software System (Version 7.3). (2016), http://www.sagemath.org.

[Sar07] P. SARNAK. Letter to : Barry Mazur on "Chebychev's bias" for $\tau(p)$. *Publications.ias.* (2007), URL: https:// publications.ias.edu/ sites/default/files/MazurLtrMay08.PDF (version: 2008-05).

[SB85] J. STIENSTRA and F. BEUKERS. On the Picard-Fuchs equation and the formal Brauer group of certain elliptic $K3$-surfaces. *Math. Ann.* **271** (1985), no. 2, 269–304.

[Sch53] B. SCHOENEBERG. Über den Zusammenhang der Eisensteinschen Reihen und Thetareihen mit der Diskriminante der elliptischen Funktionen. *Math. Ann.* **126** (1953), 177–184.

[Sha97] F. SHAHIDI. On non-vanishing of twisted symmetric and exterior square L-functions for GL(n). *Pacific J. Math.* (1997), no. Special Issue, 311–322, Olga Taussky-Todd: in memoriam.

[Shi94] G. SHIMURA. Introduction to the arithmetic theory of automorphic functions. *Publications of the Mathematical Society of Japan, vol. 11*. (Princeton University Press, Princeton, NJ, 1994). Reprint of the 1971 original, Kanô Memorial Lectures, 1.

[Ste93] E. M. STEIN. Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals. *Princeton Mathematical Series, vol. 43*. (Princeton University Press, Princeton, NJ, 1993). With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

[TW95] R. TAYLOR and A. WILES. Ring-theoretic properties of certain Hecke algebras. *Ann. of Math.* (2) **141** (1995), no. 3, 553–572.

[Wat95] G. N. WATSON. *A treatise on the theory of Bessel functions*. (Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1995). Reprint of the second (1944) edition.

[Wil95] A. WILES. Modular elliptic curves and Fermat's last theorem. *Ann. of Math.* (2) **141** (1995), no. 3, 443–551.

[Win41] A. WINTNER. On the distribution function of the remainder term of the prime number theorem. *Amer. J. Math.* **63** (1941), 233–248.

https://doi.org/10.1017/S0305004119000100 Published online by Cambridge University Press