5

The Chebychev Bias

Probability tools	Arithmetic tools
Definition of convergence in law (§ B.3)	Primes in arithmetic progressions
Kronecker's Theorem (th. B.6.5)	Orthogonality of Dirichlet characters (prop. C.5.1)
Convergence in law using auxiliary parameters (prop. B.4.4)	Dirichlet L-functions (§ C.5)
Characteristic functions (§ B.5)	Generalized Riemann Hypothesis (conj. C.5.8)
Kolmogorov's Theorem for random series (th. B.10.1)	Explicit formula (th. C.5.6)
Method of moments (th. B.5.5)	Distribution of the zeros of L-functions (prop. C.5.3)
	Generalized Simplicity Hypothesis

5.1 Introduction

One of the most remarkable limit theorems in probabilistic number theory is related to a surprising feature of the distribution of prime numbers, which was first noticed by Chebychev [24] in 1853: there seemed to be many more primes p such that $p \equiv 3 \pmod 4$ than primes with $p \equiv 1 \pmod 4$ (any prime, except $p = 2$, must satisfy one of these two conditions). More precisely, he states:

> En cherchant l'expression limitative des fonctions qui déterminent la totalité des nombres premiers de la forme $4n + 1$ et de ceux de la forme $4n + 3$, pris au-dessous d'une limite très grande, je suis parvenu à reconnaître que ces deux

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

fonctions diffèrent notablement entre elles par leurs seconds termes, dont la valeur, pour les nombres $4n+3$, est plus grande que celle pour les nombres $4n+1$; ainsi, si de la totalité des nombres premiers de la forme $4n+3$, on retranche celle des nombres premiers de la forme $4n+1$, et que l'on divise ensuite cette différence par la quantité $\frac{\sqrt{x}}{\log x}$, on trouvera plusieurs valeurs de x telles, que ce quotient s'approchera de l'unité aussi près qu'on le voudra.[1]

It is unclear from Chebychev's very short note what exactly he had proved, or simply conjectured, and he did not publish anything more on this topic. It is definitely *not* the case that we have

$$\pi(x;4,3) > \pi(x;4,1)$$

for all $x \geqslant 2$, where (in general), for an integer $q \geqslant 1$ and an integer a, we write $\pi(x;q,a)$ for the number of primes $p \leqslant x$ such that $p \equiv a \pmod q$. Indeed, for $x = 26,861$, we have

$$\pi(x;4,3) = 1472 < 1473 = \pi(x;4,1)$$

(as discovered by Leech in 1957), and one can prove that there are infinitely many sign changes of the difference $\pi(x;4,3) - \pi(x;4,1)$.

In any case, by communicating his observations, Chebychev created a fascinating area of number theory. We will discuss some of the basic known results in this chapter, which put the question on a rigorous footing, and in particular confirm the existence of the *bias* toward the residue class of 3 modulo 4, in a precise sense (although this conclusion will depend on currently unproved conjectures). Because of this feature, the subject is called the study of the *Chebychev bias*.

5.2 The Rubinstein–Sarnak Distribution

In order to study the problem suggested by Chebychev, we consider for $\mathrm{X} \geqslant 1$ the probability space $\Omega_{\mathrm{X}} = [1,\mathrm{X}]$, with the probability measure

$$\mathbf{P}_{\mathrm{X}} = \frac{1}{\log \mathrm{X}} \frac{dx}{x}. \tag{5.1}$$

[1] English translation: "While searching for the limiting expression of the functions that determine the number of prime numbers of the form $4n+1$ and of those of the form $4n+3$, less than a very large limit, I have succeeded in recognizing that the second terms of these two functions differ notably from each other; its value [of this second term], for the numbers $4n+3$, is larger than that for the numbers $4n+1$; thus, if from the number of prime numbers of the form $4n+3$, we subtract that of the prime numbers of the form $4n+1$, and then divide this difference by the quantity $\frac{\sqrt{x}}{\log x}$, we will find several values of x such that this ratio will approach one as closely as we want."

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

Let $q \geqslant 1$ be an integer. We define a random variable on Ω_X, with values in the vector space $C_{\mathbf{R}}((\mathbf{Z}/q\mathbf{Z})^\times)$ of real-valued functions on the (fixed) finite group $(\mathbf{Z}/q\mathbf{Z})^\times$, by defining $\mathsf{N}_{X,q}(x)$, for $x \in \Omega_X$, to be the function such that

$$\mathsf{N}_{X,q}(x)(a) = \frac{\log x}{\sqrt{x}}\big(\varphi(q)\pi(x;q,a) - \pi(x)\big) \tag{5.2}$$

for $a \in (\mathbf{Z}/q\mathbf{Z})^\times$ (this could also, of course, be viewed as a random real vector with values in $\mathbf{R}^{|(\mathbf{Z}/q\mathbf{Z})^\times|}$, but the perspective of a function will be slightly more convenient).

We see that the knowledge of $\mathsf{N}_{X,q}$ allows us to compare the number of primes up to X in any family of invertible residue classes modulo q. It is therefore appropriate for the study of the questions suggested by Chebychev.

We observe that in the remainder of this chapter, we will consider q to be fixed (although there are interesting questions that one can ask about uniformity with respect to q). For this reason, we will often simplify the notation (especially during proofs) to write N_X instead of $\mathsf{N}_{X,q}$, and similarly dropping q in some other cases.

Remark 5.2.1 (1) If $q = 4$, then $(\mathbf{Z}/4\mathbf{Z})^\times = \{1,3\}$, and for $x \in \Omega_T$, the random function $\mathsf{N}_{X,4}(x)$ is given by

$$1 \mapsto \frac{\log x}{\sqrt{x}}(2\pi(x;4,1) - \pi(x)) \quad \text{and} \quad 3 \mapsto \frac{\log x}{\sqrt{x}}(2\pi(x;4,3) - \pi(x)).$$

(2) Recall that the fundamental theorem of Dirichlet, Hadamard and de la Vallée Poussin (Theorem C.3.7) shows that

$$\pi(x;q,a) \sim \frac{1}{\varphi(q)}\pi(x)$$

for all a coprime to q. Thus the random variables N_X are considering the correction term from the asymptotic behavior.

(3) The normalizing factor $(\log x)/\sqrt{x}$, which is the "correct one," is the same one that is suggested by Chebychev's quote.

The basic probabilistic result concerning these arithmetic quantities is the following:

Theorem 5.2.2 (Rubinstein–Sarnak) *Let $q \geqslant 1$. Assume the* Generalized Riemann Hypothesis *modulo q. Then the random functions $\mathsf{N}_{X,q}$ converge in law to a random function N_q. The support of N_q is contained in the hyperplane*

$$\mathrm{H}_q = \Big\{ f : (\mathbf{Z}/q\mathbf{Z})^\times \to \mathbf{R} \mid \sum_{a \in (\mathbf{Z}/q\mathbf{Z})^\times} f(a) = 0 \Big\}. \tag{5.3}$$

We call N_q the *Rubinstein–Sarnak distribution* modulo q.

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

Remark 5.2.3 One may wonder if the choice of the logarithmic weight in the probability measure $\mathbf{P}_X$ is necessary for such a statement of convergence in law: this is indeed the case, and we will say a few words to explain this in Remark 5.3.5.

The Generalized Riemann Hypothesis modulo q is originally a statement about the zeros of certain analytic functions, the *Dirichlet* L-*functions* modulo q. It has, however, a concrete formulation in terms of the distribution of prime numbers: it is equivalent to the statement that, for all integers a coprime with q and all $x \geqslant 2$, we have

$$\pi(x;q,a) = \frac{1}{\varphi(q)}\int_2^x \frac{dt}{\log t} + \mathrm{O}(x^{1/2}(\log qx)),$$

where the implied constant is absolute (see, e.g., [59, 5.14, 5.15] for this equivalence). The size of the (expected) error term, approximately $\sqrt{x}$, is related to the zeros of the Dirichlet L-functions, as we will see later; it explains that the normalization factor in (5.2) is the right one for the existence of a limit in law as in Theorem 5.2.2. Indeed, using the case $q = 1$, which is the formula

$$\pi(x) = \int_2^x \frac{dt}{\log t} + \mathrm{O}(x^{1/2}(\log x)),$$

we deduce that each value of the function N_X satisfies

$$\frac{\log x}{\sqrt{x}}(\varphi(q)\pi(x;q,a) - \pi(x)) = \mathrm{O}(\varphi(q)(\log qx)^2).$$

To see how Theorem 5.2.2 helps answer questions related to the Chebychev bias, we take $q = 4$. Then we expect that

$$\lim_{X\to+\infty} \mathbf{P}_X(\pi(x;4,3) > \pi(x;4,1)) = \mathbf{P}(\mathrm{N}_4 \in \mathrm{H}_4 \cap \mathrm{C}),$$

where $\mathrm{C} = \{(x_1,x_3) \mid x_3 > x_1\}$ (although whether this limit exists or not does not follow from Theorem 5.2.2, without further information concerning the properties of the limit N_4). Then Chebychev's basic observation could be considered to be confirmed if $\mathbf{P}(\mathrm{N}_4 \in \mathrm{H}_4 \cap \mathrm{C})$ is close to 1. But in the absence of any other information, it seems very hard to prove (or disprove) this last fact.

However, Rubinstein and Sarnak showed that one could go much further by making one extra assumption on the distribution of the zeros of Dirichlet L-functions. Indeed, one can then represent N_q explicitly as the sum of a series of independent random variables (and in particular compute explicitly the characteristic function of the random function N_q). We describe this random series in Section 5.4, since to do so at this point would lead to a statement that

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

would appear highly unmotivated. The proof of Theorem 5.2.2 will lead us naturally to this next step (see Theorem 5.4.4 for the details).

Below, we write

$$\sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q)} (\ldots) \quad \text{and} \quad \sideset{}{^*}\prod_{\chi \,(\mathrm{mod}\, q)} (\cdots)$$

for a sum or a product over nontrivial[2] Dirichlet characters modulo q; we recall that these are (completely) multiplicative functions on $\mathbf{Z}$ such that $\chi(n) = 0$ unless n is coprime to q, in which case we have $\chi(n) = \widetilde{\chi}(n)$ for some group homomorphism $\widetilde{\chi} \colon (\mathbf{Z}/q\mathbf{Z})^\times \to \mathbf{C}^\times$.

We define a function m_q on $(\mathbf{Z}/q\mathbf{Z})^\times$ by

$$m_q(a) = - \sideset{}{^*}\sum_{\substack{\chi \,(\mathrm{mod}\, q) \\ \chi^2 = 1}} \overline{\chi(a)} \tag{5.4}$$

for $a \in (\mathbf{Z}/q\mathbf{Z})^\times$. This can also, using orthogonality of characters modulo q (see Proposition C.5.1), be expressed in the form

$$m_q(a) = 1 - \sum_{\substack{b \in (\mathbf{Z}/q\mathbf{Z})^\times \\ b^2 = a \,(\mathrm{mod}\, q)}} 1,$$

from which we see that in fact we have simply two possible values, namely,

$$m_q(a) = \begin{cases} 1 & \text{if } a \text{ is not a square modulo } q, \\ 1 - \sigma_q & \text{otherwise,} \end{cases} \tag{5.5}$$

where

$$\sigma_q = |\{b \in (\mathbf{Z}/q\mathbf{Z})^\times \mid b^2 = 1\}| = |\{\chi \,(\mathrm{mod}\, q) \mid \chi^2 = 1\}|$$

is also the index of the subgroup of squares in $(\mathbf{Z}/q\mathbf{Z})^\times$.

In the remaining sections of this chapter, we will explain the proof of Theorem 5.2.2, following Rubinstein and Sarnak. We will assume some familiarity with Dirichlet L-functions (in Section C.5, we recall the relevant definitions and standard facts). Readers who have not yet been exposed to these functions will probably find it easier to assume in what follows that $q = 4$. In this case, there is only one nontrivial Dirichlet L-function modulo 4, which is defined by

[2] We emphasize, for readers already familiar with analytic number theory, that this does not mean *primitive* characters.

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

$$\mathrm{L}(\chi_4, s) = \sum_{k \geqslant 0} \frac{(-1)^k}{(2k+1)^s} = \sum_{n \geqslant 1} \chi_4(n) n^{-s},$$

corresponding to the character χ_4 such that

$$\chi_4(n) = \begin{cases} 0 & \text{if } n \text{ is even,} \\ (-1)^k & \text{if } n = 2k+1 \text{ is odd} \end{cases} \tag{5.6}$$

for $n \geqslant 1$. The arguments should then be reasonably transparent. In particular, any sum of the type

$$\sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, 4)} (\cdots)$$

means that one only considers the expression on the right-hand side for the character χ_4 defined in (5.6).

5.3 Existence of the Rubinstein–Sarnak Distribution

The proof of Theorem 5.2.2 depends roughly on two ingredients:

- on the arithmetic side, we can represent the arithmetic random functions N_X as combinations of $x \mapsto x^{i\gamma}$, where the γ are ordinates of zeros of the L-functions modulo q;
- once this is done, we observe that Kronecker's Equidistribution Theorem (Theorem B.6.5) implies convergence in law for any function of this type.

There are some intermediate approximation steps involved, but the ideas are quite intuitive.

In this section, we always assume the validity of the Generalized Riemann Hypothesis modulo q, unless otherwise noted.

For a Dirichlet character χ modulo q, we define random variables ψ_χ on Ω_X by

$$\psi_\chi(x) = \frac{1}{\sqrt{x}} \sum_{n \leqslant x} \Lambda(n)\chi(n)$$

for $x \in \Omega_X$, where Λ is the von Mangoldt function (see Section C.4, especially (C.6), for the definition of this function).

The next lemma is a key step to express N_X in terms of Dirichlet characters. It looks first like standard harmonic analysis, but there is a subtle point in the proof that is crucial for the rest of the argument, and for the very existence of the Chebychev bias.

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

Lemma 5.3.1 *We have*

$$\mathsf{N}_{\mathrm{X},q} = m_q + \sum_{\chi \pmod q}^{*} \psi_\chi\,\overline{\chi} + \mathsf{E}_{\mathrm{X},q},$$

where $\mathsf{E}_{\mathrm{X},q}$ *converges to* 0 *in probability as* $\mathrm{X} \to +\infty$.

Proof By orthogonality of the Dirichlet characters modulo q (see Proposition C.5.1), we have

$$\varphi(q)\pi(x;q,a) = \sum_{\chi \pmod q} \overline{\chi(a)} \sum_{p\leqslant x} \chi(p),$$

hence

$$\frac{\log x}{\sqrt{x}}\big(\varphi(q)\pi(x;q,a) - \pi(x)\big) = \sum_{\chi \pmod q}^{*} \overline{\chi(a)}\,\frac{\log x}{\sqrt{x}} \sum_{p\leqslant x} \chi(p) + \mathrm{O}\Big(\frac{\log x}{\sqrt{x}}\Big)$$

for $x \geqslant 2$, where the error term accounts for primes p dividing q (for which the trivial character takes the value 0 instead of 1); in particular, the implied constant depends on q.

We now need to connect the sum over primes, for a fixed character χ, to ψ_χ. Recall that the von Mangoldt functions differs little from the characteristic function of primes multiplied by the logarithm function. The sum of this simpler function is the random variable defined by

$$\theta_\chi(x) = \frac{1}{\sqrt{x}} \sum_{p\leqslant x} \chi(p)\log(p)$$

for $x \in \Omega_{\mathrm{X}}$. It is related to ψ_χ by

$$\theta_\chi(x) - \psi_\chi(x) = -\frac{1}{\sqrt{x}} \sum_{k\geqslant 2} \sum_{p^k\leqslant x} \chi(p^k)\log p = -\frac{1}{\sqrt{x}} \sum_{k\geqslant 2} \sum_{p^k\leqslant x} \chi(p)^k \log p.$$

We can immediately see that the contribution of $k \geqslant 3$ is very small: since the exponent k is at most of size $\log x$, and $|\chi(p)| \leqslant 1$ for all primes p, it is bounded by

$$\Bigg|\frac{1}{\sqrt{x}} \sum_{k\geqslant 2} \sum_{\substack{p^k\leqslant x\\ k\geqslant 3}} \chi(p)^k \log p\Bigg| \leqslant \frac{1}{\sqrt{x}} \sum_{3\leqslant k\leqslant \log x} (\log x) x^{1/k} \ll \frac{(\log x)^2}{x^{1/6}},$$

where the implied constant is absolute.

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

For $k = 2$, there are two cases. If χ^2 is the trivial character, then

$$\frac{1}{\sqrt{x}} \sum_{p \leqslant \sqrt{x}} \chi(p)^2 \log p = \frac{1}{\sqrt{x}} \sum_{\substack{p \leqslant \sqrt{x} \\ p \nmid q}} \log p = 1 + \mathrm{O}\left(\frac{1}{\log x}\right)$$

by a simple form of the Prime Number Theorem in arithmetic progressions (the Generalized Riemann Hypothesis would of course give a much better error term, but this is not needed here). If χ^2 is nontrivial, then we have

$$\frac{1}{\sqrt{x}} \sum_{p \leqslant \sqrt{x}} \chi(p)^2 \log p \ll \frac{1}{\log x}$$

for the same reason. Thus we have

$$\theta_\chi(x) = \psi_\chi(x) - \delta_{\chi^2} + \mathrm{O}\left(\frac{1}{\log x}\right), \tag{5.7}$$

where δ_{χ^2} is 1 if χ^2 is trivial, and is zero otherwise.

By summation by parts, we have

$$\sum_{p \leqslant x} \chi(p) = \frac{1}{\log x} \sum_{p \leqslant x} \chi(p) \log p + \int_2^x \left(\sum_{p \leqslant t} \chi(p) \log p\right) \frac{dt}{t(\log t)^2}$$

for any Dirichlet character χ modulo q, so that

$$\begin{aligned}\frac{\log x}{\sqrt{x}}\big(\varphi(q)\pi(x;q,a) - \pi(x)\big) = {}& \sum_{\chi \,(\mathrm{mod}\, q)}^{*} \overline{\chi(a)}\, \theta_\chi(x) \\ & + \frac{\log x}{\sqrt{x}} \int_2^x \frac{\theta_\chi(t)}{t^{1/2}(\log t)^2} dt + \mathrm{O}\left(\frac{\log x}{\sqrt{x}}\right).\end{aligned} \tag{5.8}$$

We begin by handling the integral for a nontrivial character χ. We have $\theta_\chi(x) = \psi_\chi(x) + \mathrm{O}(1/\log x)$ if $\chi^2 \neq 1_q$, which implies

$$\int_2^x \frac{\theta_\chi(t)}{t^{1/2}(\log t)^2} dt = \int_2^x \frac{\psi_\chi(t)}{t^{1/2}(\log t)^2} dt + \mathrm{O}\left(\frac{x^{1/2}}{(\log x)^3}\right)$$

since

$$\int_2^x \frac{1}{t^{1/2}(\log t)^2} dt \ll \frac{x^{1/2}}{(\log x)^2}.$$

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

If χ^2 is trivial, we have an additional constant term $\theta_\chi(x) - \psi_\chi(x) = 1 + \mathrm{O}(1/\log x)$, and we get

$$\int_2^x \frac{\theta_\chi(t)}{t^{1/2}(\log t)^2}dt = \int_2^x \frac{\psi_\chi(t)}{t^{1/2}(\log t)^2}dt + \int_2^x \frac{1}{t^{1/2}(\log t)^2}dt + \mathrm{O}\left(\frac{x^{1/2}}{(\log x)^3}\right)$$
$$= \int_2^x \frac{\psi_\chi(t)}{t^{1/2}(\log t)^2}dt + \mathrm{O}\left(\frac{x^{1/2}}{(\log x)^2}\right).$$

Thus, in all cases, we get

$$\frac{\log x}{\sqrt{x}} \int_2^x \frac{\theta_\chi(t)}{t^{1/2}(\log t)^2}dt = \frac{\log x}{\sqrt{x}} \int_2^x \frac{\psi_\chi(t)}{t^{1/2}(\log t)^2}dt + \mathrm{O}\left(\frac{1}{\log x}\right).$$

Now comes the subtle point we previously mentioned. If we were to use the pointwise bound $\psi_\chi(t) \ll (\log t)^2$ (which is essentially the content of the Generalized Riemann Hypothesis) in the remaining integral, we would only get

$$\frac{\log x}{\sqrt{x}} \int_2^x \frac{\psi_\chi(t)}{t^{1/2}(\log t)^2}dt \ll \log x,$$

which is too big. So we need to use the integration process nontrivially. Precisely, by Corollary C.5.11, we have

$$\int_2^x \psi_\chi(t)dt \ll x$$

for all $x \geqslant 2$ (this reflects a “smoothing” effect due to the convergence of the series with terms $1/|\frac{1}{2} + i\gamma|^2$, where γ are the ordinates of zeros of $\mathrm{L}(s,\chi)$). Using integration by parts, we can then deduce that

$$\frac{\log x}{\sqrt{x}} \int_2^x \frac{\psi_\chi(t)}{t^{1/2}(\log t)^2}dt \ll \frac{\log x}{\sqrt{x}} \left(\frac{x}{x^{1/2}(\log x)^2} + \int_2^x \frac{t^{1/2}dt}{(\log t)^2}\right) \ll \frac{1}{\log x}.$$

Finally, we transform the first term of (5.8) to express it in terms of ψ_χ, again using (5.7). For any element $a \in (\mathbf{Z}/q\mathbf{Z})^\times$ and $x \in \Omega_X$, we have

$$\frac{\log x}{\sqrt{x}}\big(\varphi(q)\pi(x;q,a) - \pi(x)\big)$$
$$= -\sum_{\substack{\chi^2=1\\ \chi\neq 1}} \overline{\chi(a)} + \sideset{}{^*}\sum_{\chi \pmod q} \overline{\chi(a)}\psi_\chi(x) + \mathrm{O}\left(\frac{1}{\log x}\right)$$
$$= m_q(a) + \sideset{}{^*}\sum_{\chi \pmod q} \overline{\chi(a)}\psi_\chi(x) + \mathrm{O}\left(\frac{1}{\log x}\right),$$

where the implied constant depends on q. Since the error term is $\ll (\log x)^{-1}$ for $x \in \Omega_X$, it converges to zero in probability, and this concludes the proof. □

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

We keep further on the notation of the lemma, except that we also sometimes write E_{X} for $\mathsf{E}_{\mathrm{X},q}$. Since E_{X} tends to 0 in probability and m_q is a fixed function on $(\mathbf{Z}/q\mathbf{Z})^\times$, Theorem 5.2.2 will follow (by Corollary B.4.2) from the convergence in law of the random functions

$$\mathsf{M}_{\mathrm{X},q} = \sideset{}{^*}\sum_{\chi \ (\mathrm{mod}\, q)} \psi_\chi \, \overline{\chi}.$$

Now we express these functions in terms of zeros of L-functions. Here and later, a sum over zeros of a Dirichlet L-function always means implicitly that zeros are counted with their multiplicity.

We will denote by I_{X} the identity variable $x \mapsto x$ on Ω_{X}; thus, for a complex number s, the random variable $\mathsf{I}_{\mathrm{X}}^s$ is the function $x \mapsto x^s$ on Ω_{X}.

Below, when we have a random function X on $(\mathbf{Z}/q\mathbf{Z})^\times$, and a nonnegative random variable Y, the meaning of a statement of the form $\mathrm{X} = \mathrm{O}(\mathrm{Y})$ is that $\|\mathrm{X}\| = \mathrm{O}(\mathrm{Y})$, where the norm is the euclidean norm, that is, we have

$$\|\mathrm{X}\|^2 = \sum_{a \in (\mathbf{Z}/q\mathbf{Z})^\times} |\mathrm{X}(a)|^2.$$

Lemma 5.3.2 *We have*

$$\mathsf{M}_{\mathrm{X},q} = - \sideset{}{^*}\sum_{\chi \ (\mathrm{mod}\, q)} \Bigl(\sum_{|\gamma| \leqslant \mathrm{X}} \frac{\mathsf{I}_{\mathrm{X}}^{i\gamma}}{\frac{1}{2} + i\gamma} \Bigr) \overline{\chi} + \mathrm{O}\Bigl(\frac{(\log \mathrm{X})^2}{\mathrm{X}^{1/2}} \Bigr),$$

where γ ranges over ordinates of zeros of $\mathrm{L}(s, \chi)$*, counted with multiplicity, and the implied constant depends on q.*

Proof The key ingredient is the (approximate) *explicit formula* of Prime Number Theory, which can be stated in the form

$$\psi_\chi = - \sum_{\substack{\mathrm{L}(\beta + i\gamma) = 0 \\ |\gamma| \leqslant \mathrm{X}}} \frac{\mathsf{I}_{\mathrm{X}}^{\beta - \frac{1}{2} + i\gamma}}{\beta + i\gamma} + \mathrm{O}\left(\frac{\mathsf{I}_{\mathrm{X}}^{1/2} \log(\mathrm{X})^2}{\mathrm{X}} \right),$$

where the sum is over zeros of the Dirichlet L-functions with $0 \leqslant \beta \leqslant 1$, counted with multiplicity (see Theorem C.5.6). Under the assumption of the Generalized Riemann Hypothesis modulo q, we always have $\beta = \frac{1}{2}$, and this formula implies

$$\psi_\chi = - \sum_{|\gamma| \leqslant \mathrm{X}} \frac{\mathsf{I}_{\mathrm{X}}^{i\gamma}}{\frac{1}{2} + i\gamma} + \mathrm{O}\left(\frac{(\log \mathrm{X})^2}{\mathrm{X}^{1/2}} \right).$$

Summing over the characters (the number of which is $\varphi(q) - 1 \leqslant q$), the formula follows. ☐

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

Probabilistically, we have now a finite linear combination (of length depending on X) of the random variables $\mathsf{l}_{\mathrm{X}}^{i\gamma}$. The link with probability theory, and to the existence of the Rubinstein–Sarnak distribution, is then performed by the following theorem (quite similar to Proposition 3.2.5).

Proposition 5.3.3 *Let $k \geqslant 1$ be an integer. Let* F *be a finite set of real numbers, and let $(\alpha(t))_{t\in \mathrm{F}}$ be a family of elements in $\mathbf{C}^k$. The random vectors*

$$\sum_{t\in \mathrm{F}} \mathsf{l}_{\mathrm{X}}^{it}\,\alpha(t)$$

on Ω_{X} converge in law as $\mathrm{X} \to +\infty$.

Proof After a simple translation, this is a direct consequence of the Kronecker Equidistribution Theorem B.6.5. Indeed, consider the vector

$$z = \left(\frac{t}{2\pi}\right)_{t\in \mathrm{F}} \in \mathbf{R}^{\mathrm{F}}.$$

By Kronecker's Theorem, the probability measures μ_{Y} on $(\mathbf{R}/\mathbf{Z})^{\mathrm{F}}$ defined for $\mathrm{Y} > 0$ by

$$\mu_{\mathrm{Y}}(\mathrm{A}) = \frac{1}{\mathrm{Y}}|\{y \in [0,\mathrm{Y}] \mid yz \in \mathrm{A}\}|,$$

for any measurable set A, converge in law to the probability Haar measure μ on the subgroup T of $(\mathbf{R}/\mathbf{Z})^{\mathrm{F}}$ generated by the classes modulo $\mathbf{Z}^{\mathrm{F}}$ of the elements yz, where y ranges over $\mathbf{R}$.

We extend the isomorphism $\theta \mapsto e(\theta)$ from $\mathbf{R}/\mathbf{Z}$ to $\mathbf{S}^1$ componentwise to define an isomorphism of $(\mathbf{R}/\mathbf{Z})^{\mathrm{F}}$ to $(\mathbf{S}^1)^{\mathrm{F}}$. For any continuous function f on $(\mathbf{S}^1)^{\mathrm{F}}$, we observe that

$$\begin{aligned}\int_{(\mathbf{R}/\mathbf{Z})^{\mathrm{F}}} f(e(v))d\mu_{\mathrm{Y}}(v) &= \frac{1}{\mathrm{Y}}\int_0^{\mathrm{Y}} f(e(yz))dy \\ &= \frac{1}{\mathrm{Y}}\int_0^{\mathrm{Y}} f\big((e^{ity})_{t\in\mathrm{F}}\big)dy \\ &= \frac{1}{\mathrm{Y}}\int_1^{e^{\mathrm{Y}}} f((x^{it})_{t\in\mathrm{F}})\frac{dx}{x} = \mathbf{E}_{\mathrm{X}}\big(f((\mathsf{l}_{\mathrm{X}}^{it})_{t\in\mathrm{F}})\big)\end{aligned}$$

for $\mathrm{X} = e^{\mathrm{Y}}$, after the change of variable $x = e^y$. Hence the vector $(\mathsf{l}_{\mathrm{X}}^{it})_{t\in\mathrm{F}}$ converges in law as $\mathrm{X} \to +\infty$ to the image of μ by $v \mapsto e(v)$. Now we finish the proof of the proposition by composition with the continuous map from $(\mathbf{S}^1)^{\mathrm{F}}$ to $\mathbf{C}^k$ defined by

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

$$(z_t)_{t\in \mathrm{F}} \mapsto \sum_{t\in\mathrm{F}} z_t\alpha(t),$$

using Proposition B.3.2. □

From the proof, we see that we can make the result more precise:

Corollary 5.3.4 *With notation and assumptions as in Proposition 5.3.3, the random vectors*

$$\sum_{t\in\mathrm{F}} \mathsf{I}_{\mathrm{X}}^{it}\,\alpha(t)$$

on Ω_{X} converge in law as $\mathrm{X}\to+\infty$ to

$$\sum_{t\in\mathrm{F}} \mathrm{I}_t\,\alpha(t),$$

where $(\mathrm{I}_t)_{t\in\mathrm{F}}$ is a random variable with values in $(\mathbf{S}^1)^{\mathrm{F}}$ with law given by the probability Haar measure of the closure of the subgroup of $(\mathbf{S}^1)^{\mathrm{F}}$ generated by all elements $(x^{it})_{t\in\mathrm{F}}$ for $x\in\mathbf{R}$.

Remark 5.3.5 This proposition explains why the logarithmic weight in (5.1) is absolutely natural. It also hints that it is necessary. Indeed, the statement of the proposition becomes false if the probability measure $\mathbf{P}_{\mathrm{X}}$ on Ω_{X} is replaced by the uniform measure. This is already visible in the simplest case where $\mathrm{F}=\{t\}$ contains a single nonzero real number t; for instance, taking the test function f to be the identity, observe that with this other probability measure, the expectation of $x\mapsto x^{it}$ is

$$\frac{1}{\mathrm{X}-1}\int_1^{\mathrm{X}} x^{it}dx = \frac{1}{it+1}\frac{\mathrm{X}^{it+1}-1}{\mathrm{X}-1} \sim \frac{\mathrm{X}^{it}}{it+1},$$

which has no limit as $\mathrm{X}\to+\infty$.

Let $\mathrm{T}\geqslant 2$ be a parameter. It follows from Lemma 5.3.2 and Proposition 5.3.3 that for $\mathrm{X}\geqslant\mathrm{T}$, we have

$$\mathsf{M}_{\mathrm{X},q} = \mathsf{N}_{\mathrm{X,T},q} + \sideset{}{^*}\sum_{\chi\,(\mathrm{mod}\,q)} \Bigl(\sum_{\mathrm{T}\leqslant|\gamma|\leqslant\mathrm{X}} \frac{\mathsf{I}_{\mathrm{X}}^{i\gamma}}{\frac{1}{2}+i\gamma}\Bigr)\overline{\chi} + \mathrm{O}\Bigl(\frac{(\log \mathrm{X})^2}{\sqrt{\mathrm{X}}}\Bigr), \qquad (5.9)$$

where

$$\mathsf{N}_{\mathrm{X,T},q} = -\sideset{}{^*}\sum_{\chi\,(\mathrm{mod}\,q)} \Bigl(\sum_{|\gamma|\leqslant\mathrm{T}} \frac{\mathsf{I}_{\mathrm{X}}^{i\gamma}}{\frac{1}{2}+i\gamma}\Bigr)\,\overline{\chi}$$

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

are random functions that converge in law as $X \to +\infty$ for any fixed $T \geqslant 2$. The next lemma will allow us to check that the remainder term in this approximation is small.

Lemma 5.3.6 *Let $k \geqslant 1$ be an integer. Let F be a countable set of real numbers, and let $(\alpha(t))_{t \in F}$ be a family of elements in $\mathbf{C}^k$. Assume that the following conditions hold for all $T \geqslant 2$ and all $t_0 \in \mathbf{R}$:*

$$\sum_{t \in F} \|\alpha(t)\|^2 \, |t|^{1/2} \, \log(1 + |t|) < +\infty, \tag{5.10}$$

$$\sum_{\substack{t \in F \\ |t| \geqslant T}} \frac{\|\alpha(t)\|}{|t|^{1/4}} \ll \frac{(\log T)^2}{T^{1/4}}, \tag{5.11}$$

$$|\{t \in F \mid |t - t_0| \leqslant 1\}| \ll \log(1 + |t_0|). \tag{5.12}$$

Then we have

$$\lim_{\substack{T \leqslant X \\ T \to +\infty}} \Big\| \sum_{\substack{t \in F \\ |t| \geqslant T}} \mathsf{l}_X^{it} \alpha(t) \Big\|_{L^2} = 0,$$

where the limit is over pairs (T, X) with $T \leqslant X$ and T tends to infinity.

In this statement, we use the Hilbert space $L^2(\Omega_X; \mathbf{R}^k)$ of $\mathbf{R}^k$-valued L^2-functions on Ω_X, with norm defined by

$$\|f\|_{L^2}^2 = \mathbf{E}_X(\|f\|^2)$$

for $f \in L^2(\Omega_X; \mathbf{R}^k)$.

Proof Note first that an explicit computation of the integral gives

$$\mathbf{E}_X(\mathsf{l}_X^{i(t_1 - t_2)}) = \frac{1}{\log X} \frac{X^{i(t_1 - t_2)} - 1}{t_1 - t_2}$$

for $t_1 \neq t_2$, hence the general bound

$$|\mathbf{E}_X(\mathsf{l}_X^{i(t_1 - t_2)})| \leqslant \min\left(1, \frac{1}{\log X} \frac{2}{|t_1 - t_2|}\right). \tag{5.13}$$

We will use this bound slightly wastefully (using the first estimate even when it is not the best of the two) to gain some flexibility.

All sums below involving t, t_1, t_2 are restricted to $t \in \mathrm{F}$. Assume $2^5 \leqslant \mathrm{T} \leqslant \mathrm{X}$. We have

$$\Big\| \sum_{\substack{t\in\mathrm{F}\\|t|\geqslant\mathrm{T}}} \mathsf{l}_{\mathrm{X}}^{it}\,\alpha(t)\Big\|_{\mathrm{L}^2} = \mathbf{E}_{\mathrm{X}}\left(\Big\|\sum_{\mathrm{T}\leqslant|t|\leqslant\mathrm{X}} \mathsf{l}_{\mathrm{X}}^{it}\,\alpha(t)\Big\|^2\right)$$
$$= \sum_{\mathrm{T}\leqslant|t_1|,|t_2|\leqslant\mathrm{X}} \alpha(t_1)\cdot\alpha(t_2)\,\mathbf{E}_{\mathrm{X}}\left(\mathsf{l}_{\mathrm{X}}^{i(t_1-t_2)}\right).$$

We write this double sum as $\mathrm{S}_1 + \mathrm{S}_2$, where S_1 is the contribution of the terms where $|t_1 - t_2| \leqslant |t_1t_2|^{1/4}$, and S_2 is the remainder.

In the sum S_1, we first claim that if $\mathrm{T} \geqslant \sqrt{2}$, then the condition $|t_1 - t_2| \leqslant |t_1t_2|^{1/4}$ implies $|t_2| \leqslant 2|t_1|$. Indeed, suppose that $|t_2| > 2|t_1|$. We have

$$|t_2| \leqslant |t_1 - t_2| + |t_1| \leqslant |t_1t_2|^{1/4} + \tfrac{1}{2}|t_2|,$$

hence $|t_2| \leqslant 2|t_1t_2|^{1/4}$, which implies $|t_2| \leqslant 2^{4/3}|t_1|^{1/3}$, and further

$$2|t_1| < |t_2| \leqslant 2^{4/3}|t_1|^{1/3},$$

which implies that $\mathrm{T} \leqslant |t_1| < \sqrt{2}$, reaching a contradiction.

Exchanging the roles of t_1 and t_2, we see also that $|t_1| \leqslant 2|t_2|$. In particular, it now follows that we also have

$$|t_2 - t_1| \leqslant |t_1t_2|^{1/4} \leqslant 2|t_1|^{1/2} \quad \text{and} \quad |t_2 - t_1| \leqslant |t_1t_2|^{1/4} \leqslant 2|t_2|^{1/2}.$$

Still for $\mathrm{T} \geqslant \sqrt{2}$, we get

$$\begin{aligned}
|\mathrm{S}_1| &\leqslant \sum_{\substack{\mathrm{T}\leqslant|t_1|,|t_2|\leqslant\mathrm{X}\\|t_2-t_1|\leqslant|t_1t_2|^{1/4}}} |\alpha(t_1)\cdot\alpha(t_2)| \\
&\leqslant \frac{1}{2}\sum_{\substack{\mathrm{T}\leqslant|t_1|,|t_2|\leqslant\mathrm{X}\\|t_2-t_1|\leqslant|t_1t_2|^{1/4}}} \left(\|\alpha(t_1)\|^2 + \|\alpha(t_2)\|^2\right) \\
&\leqslant \sum_{\mathrm{T}\leqslant|t_1|\leqslant\mathrm{X}} \|\alpha(t_1)\|^2 \sum_{\substack{\mathrm{T}\leqslant|t_2|\leqslant\mathrm{X}\\|t_2-t_1|\leqslant2|t_1|^{1/2}}} 1 + \sum_{\mathrm{T}\leqslant|t_2|\leqslant\mathrm{X}} \|\alpha(t_2)\|^2 \sum_{\substack{\mathrm{T}\leqslant|t_1|\leqslant\mathrm{X}\\|t_2-t_1|\leqslant2|t_2|^{1/2}}} 1 \\
&\ll \sum_{\mathrm{T}\leqslant|t|\leqslant\mathrm{X}} \|\alpha(t)\|^2|t|^{1/2}\log(1+|t|)
\end{aligned}$$

by (5.12). This quantity tends to 0 as $\mathrm{T} \to +\infty$ since the series over all t converges by assumption (5.10).

For the sum S_2, we have

$$|S_2| \leqslant \frac{2}{\log X} \sum_{\substack{T\leqslant |t_1|,|t_2|\leqslant X\\ |t_2-t_1|>|t_1t_2|^{1/4}}} \frac{\|\alpha(t_1)\|\,\|\alpha(t_2)\|}{|t_1-t_2|}$$
$$\leqslant \frac{2}{\log X} \sum_{\substack{T\leqslant |t_1|,|t_2|\leqslant X\\ |t_2-t_1|>|t_1t_2|^{1/4}}} \frac{\|\alpha(t_1)\|\,\|\alpha(t_2)\|}{|t_1t_2|^{1/4}},$$

and therefore

$$|S_2| \leqslant \frac{2}{\log X} \sum_{T\leqslant |t_1|\leqslant X} \frac{\|\alpha(t_1)\|}{|t_1|^{1/4}} \sum_{T\leqslant |t_2|\leqslant X} \frac{\|\alpha(t_2)\|}{|t_2|^{1/4}} \ll \frac{1}{\log X}\frac{(\log T)^4}{T^{1/2}},$$

by (5.11). The lemma now follows. □

Remark 5.3.7 Although we have stated this lemma in some generality, it is far from the best that can be achieved along such lines.

The assumptions might look complicated, but note that (5.12) means that the density of F is roughly logarithmic; then (5.10) and (5.11) are certainly satisfied if the series with terms $\|\alpha(t)\|$ is convergent, and more generally when $\|\alpha(t)\|$ is comparable with $(1+|t|)^{-\alpha}$ with $\alpha > 3/4$.

We will now finish the proof of Theorem 5.2.2. We apply Lemma 5.3.6 to the set F of ordinates γ of zeros of some $L(s,\chi)$, for χ a nontrivial character modulo q, and to

$$\alpha(\gamma) = \sideset{}{^*}\sum_{\substack{\chi \ (\mathrm{mod}\, q)\\ L(\frac{1}{2}+i\gamma)=0}} \frac{1}{\frac{1}{2}+i\gamma}\,\overline{\chi}$$

for $\gamma \in F$, viewing $\alpha(\gamma)$ as a vector in $\mathbf{C}^{(\mathbf{Z}/q\mathbf{Z})^\times}$, and taking into account the multiplicity of the zero $\frac{1}{2}+i\gamma$ for any character χ such that $L(\frac{1}{2}+i\gamma,\chi)=0$. We need to check the three assumptions of the lemma.

From the asymptotic von Mangoldt formula (C.10), we first know that (5.12) holds for the zeros of a fixed L-function modulo q, with an implied constant depending on q, and hence it holds also for F.

We next have

$$\|\alpha(\gamma)\| \leqslant \sideset{}{^*}\sum_{\substack{\chi \ (\mathrm{mod}\, q)\\ L(\frac{1}{2}+i\gamma)=0}} \frac{1}{|\frac{1}{2}+i\gamma|}\,\|\overline{\chi}\| = \varphi(q)^{1/2} \sideset{}{^*}\sum_{\substack{\chi \ (\mathrm{mod}\, q)\\ L(\frac{1}{2}+i\gamma)=0}} \frac{1}{|\frac{1}{2}+i\gamma|} \leqslant \frac{\varphi(q)^{3/2}}{|\frac{1}{2}+i\gamma|} \tag{5.14}$$

by a trivial estimate of the number of characters of which $\frac{1}{2}+i\gamma$ can be a zero.

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

Condition (5.10) follows from (5.14), since we even have

$$\sum_{L(\frac{1}{2}+i\gamma,\chi)=0} \frac{1}{|\frac{1}{2}+i\gamma|^{1+\varepsilon}} < +\infty \tag{5.15}$$

for any fixed $\varepsilon > 0$ and any $\chi \pmod q$, and condition (5.11) is again an easy consequence of (5.15) and (5.14).

From (5.9), we conclude that for $X \geqslant T \geqslant 2$, we have

$$\mathsf{M}_X = \mathsf{N}_{X,T} + \mathsf{E}'_{X,T},$$

where

$$\mathsf{E}'_{X,T} = \sideset{}{^*}\sum_{\chi \pmod q} \Big(\sum_{T\leqslant|\gamma|\leqslant X} \frac{\mathsf{I}_X^{i\gamma}}{\frac{1}{2}+i\gamma} \Big) \overline{\chi} + O\Big(\frac{(\log X)^2}{\sqrt{X}} \Big).$$

These random functions converge to 0 in L^2, hence in L^1, by Lemma 5.3.6 as applied before. By Proposition B.4.4 (and Remark B.4.6), we conclude that the random functions M_X converge in law, and that their limit is the same as the limit as $T \to +\infty$ of the law of the limit of

$$- \sideset{}{^*}\sum_{\chi \pmod q} \Big(\sum_{|\gamma|\leqslant T} \frac{\mathsf{I}_X^{i\gamma}}{\frac{1}{2}+i\gamma} \Big) \overline{\chi}.$$

In the next section, we compute these limits, and hence the law of N_q, assuming that the zeros of the Dirichlet L-functions are "as independent as possible," so that Proposition 5.3.3 becomes explicit in the special case of interest.

To finish the proof of Theorem 5.2.2, we need to check the last assertion, namely, that the support of N_q is contained in the hyperplane (5.3). But note that

$$\begin{aligned} \sum_{a\in(\mathbf{Z}/q\mathbf{Z})^\times} \mathsf{N}_X(x)(a) &= \frac{\log x}{\sqrt{x}} \sum_{a\in(\mathbf{Z}/q\mathbf{Z})^\times} (\varphi(q)\pi(x;q,a) - \pi(x)) \\ &= \frac{\log x}{\sqrt{x}} \sum_{\substack{p\leqslant x \\ p \pmod q \notin (\mathbf{Z}/q\mathbf{Z})^\times}} 1 \ll \frac{\log x}{\sqrt{x}} \end{aligned}$$

for all $x \in \Omega_X$, since at most finitely many primes are not congruent to some $a \in (\mathbf{Z}/q\mathbf{Z})^\times$. Hence the random variables

$$\sum_{a\in(\mathbf{Z}/q\mathbf{Z})^\times} \mathsf{N}_X(a)$$

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

converge in probability to 0 as $\mathrm{X} \to +\infty$, and by Corollary B.3.4, it follows that the support of N_q is contained in the zero set of the linear form

$$f \mapsto \sum_{a \in (\mathbf{Z}/q\mathbf{Z})^\times} f(a),$$

that is, in H_q.

5.4 The Generalized Simplicity Hypothesis

The proof of Theorem 5.2.2 now allows us to understand what is needed for the next step, which we take to be the explicit determination of the random variable N_q. Indeed, the proof tells us that N_q is the limit, as $\mathrm{T} \to +\infty$, of the random variables that are themselves the limits in law as $\mathrm{X} \to +\infty$ of the random function given by the finite sum

$$m_q - \sideset{}{^*}\sum_{\chi \pmod q} \sum_{\substack{\mathrm{L}(\frac{1}{2}+i\gamma,\chi)=0 \\ |\gamma| \leqslant \mathrm{T}}} \frac{\mathsf{I}_{\mathrm{X}}^{i\gamma}}{\frac{1}{2}+i\gamma} \overline{\chi(a)},$$

which converge by Proposition 5.3.3. The *proof* of that proposition shows how this limit $\mathrm{N}_{q,\mathrm{T}}$ can be computed in principle. Precisely, let X_T be the set of pairs (χ,γ), where χ runs over nontrivial Dirichlet characters modulo q and γ runs over the ordinates of the nontrivial zeros of $\mathrm{L}(s,\chi)$ with $|\gamma| \leqslant \mathrm{T}$. Then, by Corollary 5.3.4, we have

$$\mathrm{N}_{q,\mathrm{T}} = m_q - \sideset{}{^*}\sum_{\chi \pmod q} \sum_{\substack{\mathrm{L}(\frac{1}{2}+i\gamma,\chi)=0 \\ |\gamma| \leqslant \mathrm{T}}} \frac{\mathrm{I}_{\chi,\gamma}}{\frac{1}{2}+i\gamma} \overline{\chi(a)}, \tag{5.16}$$

where $(\mathrm{I}_{\chi,\gamma})$ is distributed on $(\mathbf{S}^1)^{\mathrm{X}_\mathrm{T}}$ according to the probability Haar measure of the closure S_T of the subgroup generated by the elements $(x^{i\gamma})_{(\chi,\gamma)\in\mathrm{X}_\mathrm{T}}$ for $x \in \mathbf{R}$.

Thus, *to compute* N_q *explicitly*, we "simply" need to know what the subgroup $\mathrm{S}_{q,\mathrm{T}}$ is. If (hypothetically) this subgroup was *equal* to $(\mathbf{S}^1)^{\mathrm{X}_{q,\mathrm{T}}}$, then the $(\mathrm{I}_{\chi,\gamma})$ would simply be independent and uniformly distributed on $\mathbf{S}^1$, and we would immediately obtain a formula for N_q from (5.16) as a sum of a series of independent terms.

This hypothesis is however too optimistic. Indeed, there is an "obvious" type of dependency among the ordinates γ, which amount to restrictions on the subgroup S_T in $(\mathbf{S}^1)^{\mathrm{X}_\mathrm{T}}$. Beyond these relations, there are none that are

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

immediately apparent. The *Generalized Simplicity Hypothesis* modulo q is then the statement that, in fact, these obvious relations should exhaust all possible constraints satisfied by S_T.[3]

These systematic relations between the elements of X_T are simply the following: a complex number $\frac{1}{2} + i\gamma$ is a zero of $L(s,\chi)$ if and only if the conjugate $\frac{1}{2} - i\gamma$ is a zero of $L(s,\overline{\chi})$, simply because $\overline{L(\bar{s},\chi)} = L(s,\overline{\chi})$ as holomorphic functions; hence (χ,γ) belongs to X_T if and only if $(\overline{\chi},-\gamma)$ does.

We are therefore led to the so-called Generalized Simplicity Hypothesis modulo q.

Definition 5.4.1 Let $q \geqslant 1$ be an integer. The *Generalized Simplicity Hypothesis* holds modulo q if the family of *nonnegative* ordinates γ of the nontrivial zeros of all nontrivial Dirichlet L-functions modulo q, with multiplicity taken into account, is linearly independent over $\mathbf{Q}$.

We emphasize that we are looking at the family of the ordinates, not just the set of values. In particular, the Generalized Simplicity Hypothesis modulo q implies that

- for a given $\gamma \geqslant 0$, there is at most one primitive Dirichlet character χ modulo q such that $L(\frac{1}{2} + i\gamma, \chi) = 0$;
- all nontrivial zeros are of multiplicity 1;
- we have $L(\frac{1}{2}, \chi) \neq 0$ for any nontrivial character χ.

All these statements are highly nontrivial conjectures!

Lemma 5.4.2 *Under the assumption of the Generalized Simplicity Hypothesis modulo q, the subgroup S_T is given by*

$$S_T = \{(z_{\chi,\gamma}) \in (\mathbf{S}^1)^{X_T} \mid z_{\overline{\chi},-\gamma} = \overline{z_{\chi,\gamma}} \text{ for all } (\chi,\gamma) \in X_T\}, \tag{5.17}$$

for all $T \geqslant 2$. In particular, denoting by X_T^+ the set of pairs (χ,γ) in X_T with $\gamma \geqslant 0$, the projection

$$(z_{\chi,\gamma}) \mapsto (z_{\chi,\gamma})_{(\chi,\gamma)\in X_T^+} \tag{5.18}$$

from S_T to $(\mathbf{S}^1)^{X_T^+}$ is surjective.

Proof Indeed, S_T is contained in the subgroup $\widetilde{S}_T$ in the right-hand side of (5.17), because each vector $(x^{i\gamma})_{(\chi,\gamma)\in X_T}$ has this property for $x \in \mathbf{R}$, by the relation between zeros of the L-functions of χ and $\overline{\chi}$.

[3] In other words, it is an application of Occam's Razor.

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

To show that $\mathrm{S_T}$ is not a proper subgroup of $\widetilde{\mathrm{S}}_\mathrm{T}$, it is enough to prove the last assertion, since an element of $\widetilde{\mathrm{S}}_\mathrm{T}$ is uniquely determined by the value of the projection (5.18). But if that projection is not surjective, then there exists a nonzero family of integers $(m_{\chi,\gamma})_{(\chi,\gamma)\in \mathrm{X}_\mathrm{T}^+}$ such that

$$\prod_{(\chi,\gamma)\in \mathrm{X}_\mathrm{T}^+} x^{im_{\chi,\gamma}\gamma} = 1$$

for all $x \in \mathbf{R}$, and this implies

$$\sum_{\chi \pmod q}^{*} \sum_{\gamma \geqslant 0} m_{\chi,\gamma}\gamma = 0,$$

which contradicts the Generalized Simplicity Hypothesis modulo q. □

Remark 5.4.3 If we were also considering problems involving the comparison of the number of primes in arithmetic progressions with different moduli, say, modulo q_1 and q_2, then there would be another systematic source of relations between the zeros of the L-functions modulo q_1 and q_2. Precisely, if d is a common divisor of q_1 and q_2, and χ_0 a Dirichlet character modulo d, corresponding to a character χ_0 of $(\mathbf{Z}/d\mathbf{Z})^\times$, then there is a Dirichlet character χ_i modulo q_i, for $i = 1, 2$, corresponding to the composition

$$(\mathbf{Z}/q_i\mathbf{Z})^\times \to (\mathbf{Z}/d\mathbf{Z})^\times \xrightarrow{\chi_0} \mathbf{C}^\times,$$

and we have

$$\mathrm{L}(s,\chi_i) = \prod_{p \mid q_i/d} (1 - \chi_0(p)p^{-s})\mathrm{L}(s,\chi_0),$$

which shows that the ordinates of the nontrivial zeros of $\mathrm{L}(s,\chi_1)$ and $\mathrm{L}(s,\chi_2)$ are the same.

Because of this, the correct formulation of the Generalized Simplicity Hypothesis, without reference to a single modulus q, is that *the nonnegative ordinates of zeros of the* L*-functions of all primitive Dirichlet characters are* **Q***-linearly independent*; this is the statement as formulated in [105].

We can now state precisely the computation of the law of the random function N_q under the assumption of the Generalized Simplicity Hypothesis modulo q.

To do this, let X^+ be the set of all pairs (χ,γ) where χ is a nontrivial Dirichlet character modulo q and $\gamma \geqslant 0$ is a *nonnegative* ordinate of a nontrivial zero of $\mathrm{L}(s,\chi)$, that is, we have $\mathrm{L}(\frac{1}{2}+i\gamma,\chi) = 0$. Let $(\mathrm{I}_{\chi,\gamma})_{(\chi,\gamma)\in \mathrm{X}^+}$ be a family of independent random variables all uniformly distributed over the circle $\mathbf{S}^1$.

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

Define further

$$\mathrm{I}_{\overline{\chi},-\gamma} = \bar{\mathrm{I}}_{\chi,\gamma} \tag{5.19}$$

for all ordinates $\gamma \geqslant 0$ of a zero of $\mathrm{L}(s,\chi)$. We have then defined random variables $\mathrm{I}_{\chi,\gamma}$ for all ordinates of a zero of $\mathrm{L}(s,\chi)$.

Theorem 5.4.4 (Rubinstein–Sarnak) *Let $q \geqslant 1$. In addition to the Generalized Riemann Hypothesis, assume the Generalized Simplicity Hypothesis modulo q. Then the law of N_q is the law of the series*

$$m_q - \sum_{\chi \,(\mathrm{mod}\, q)}^{*} \Bigg(\sum_{\substack{\gamma \\ \mathrm{L}(\frac{1}{2}+i\gamma,\chi)=0}} \frac{\mathrm{I}_{\chi,\gamma}}{\frac{1}{2}+i\gamma} \Bigg) \overline{\chi}, \tag{5.20}$$

where the series converges almost surely and in L^2 as the limit of partial sums

$$\lim_{\mathrm{T}\to+\infty} \sum_{\chi \,(\mathrm{mod}\, q)}^{*} \sum_{\substack{\mathrm{L}(\frac{1}{2}+i\gamma,\chi)=0 \\ |\gamma|\leqslant \mathrm{T}}} \frac{\mathrm{I}_{\chi,\gamma}}{\frac{1}{2}+i\gamma}. \tag{5.21}$$

In these formulas, for each Dirichlet character χ modulo q, the sum runs over the ordinates of zeros of $\mathrm{L}(s,\chi)$.

Remark 5.4.5 (1) Since the Generalized Simplicity Hypothesis modulo q implies that each zero has multiplicity one (even as we vary χ modulo q), there is no need to worry about this issue when defining the series over the zeros.

(2) This result shows that the random function N_q is probabilistically quite subtle. It is somewhat analogue to Bagchi's measure, or to one of its Bohr–Jessen specializations (see Theorem 3.2.1), with a sum (or a product) of rather simple individual independent random variables, but it retains important arithmetic features because the sum and the coefficients involve the zeros of Dirichlet L-functions (instead of the primes that occur in Bagchi's random Euler product).

One important contrasting feature, in comparison with either Theorem 3.2.1 (or Selberg's Theorem) is that the series defining N_q is not far from being absolutely convergent, which is not the case at all of the series

$$\sum_p \frac{\mathrm{X}_p}{p^s}$$

that occurs in Bagchi's Theorem when $\frac{1}{2} < \mathrm{Re}(s) < 1$.

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

Before giving the proof, we can draw some simple conclusions from Theorem 5.4.4, in the direction of confirming the existence of a bias for certain residue classes.

Under the assumptions of Theorem 5.4.4, we have $\mathbf{E}(\mathrm{N}_q) = m_q$, since the convergence also holds in L^2, and $\mathbf{E}(\mathrm{I}_{\chi,\gamma}) = 0$ for all (χ,γ). Using either (5.4) or (5.5), we know that

$$\frac{1}{\varphi(q)} \sum_{a\in(\mathbf{Z}/q\mathbf{Z})^\times} m_q(a) = 0 \quad \text{and} \quad \frac{1}{\varphi(q)} \sum_{a\in(\mathbf{Z}/q\mathbf{Z})^\times} m_q(a)^2 = \sigma_q = \sum_{\chi^2=1}{}^{*} 1.$$

It is natural to say that "not all residue classes modulo q are equal," as far as representing primes is concerned, if the average function m_q of N_q is not constant (assuming that Theorem 5.4.4 is applicable). This is equivalent (by (5.5)) to the existence of at least one $b \neq 1$ such that $b^2 = 1$, and therefore holds whenever $q \neq 2$, since one can always take $b = -1$.

This statement can be considered to be the simplest general confirmation of the Chebychev bias; note that $q = 2$ is of course an exception, since all primes (with one exception) are odd.

Remark 5.4.6 (1) The mean-square σ_q of m_q is also the size of the quotient group

$$(\mathbf{Z}/q\mathbf{Z})^\times/((\mathbf{Z}/q\mathbf{Z})^\times)^2$$

of invertible residues modulo quadratic residues, minus 1. Using the Chinese Remainder Theorem, this expression can be computed in terms of the factorization of q, namely, if we write

$$q = \prod_p p^{n_p},$$

then we obtain

$$\sigma_q = 2^{\min(n_2-1,2)} \prod_{\substack{p|q \\ p\geqslant 3}} 2 - 1$$

(because for p odd, the group of squares is of index 2 in $(\mathbf{Z}/p^{n_p}\mathbf{Z})^\times$ if $n_p \geqslant 1$, whereas for $p = 2$, it is trivial if $n_p = 1$ or $n_p = 2$, and of index 4 if $n_2 \geqslant 3$).

(2) Consider once more the case $q = 4$. Then $m_4(1) = -1$ and $m_4(3) = 1$, and in particular we certainly expect to have, in general, more primes congruent to 3 modulo 4 than there are congruent to 1 modulo 4.

In fact, using Theorem 5.4.4 and numerical tables of zeros of the Dirichlet L-functions modulo 4 up to some bound T, one can get approximations to the distribution of N_4 (e.g., through the characteristic function of N_4, and

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

approximate Fourier inversion). Rubinstein and Sarnak [105, §4] established in this manner that

$$\mathbf{P}(\mathrm{N}_4 \in \mathrm{H}_4 \cap \mathrm{C}) = 0.9959\ldots$$

(under the assumptions of Theorem 5.4.4 modulo 4). This confirms a very strong bias for primes to be $\equiv 3$ modulo 4, but also shows that one has sometimes $\pi(x;4,1) > \pi(x;4,3)$ (in fact, in the sense of the probability measure $\mathbf{P}_{\mathrm{X}}$, this happens with probability about $1/250$, and we have already mentioned that the first occurrence of this reverse inequality is for $\mathrm{X} = 26861$).

We now give the proof of Theorem 5.4.4. We first check that the series (5.20) converges almost surely and in L^2 in the sense of the limit (5.21).[4]

It suffices to prove that each value $\mathrm{N}_q(a)$ of the random function N_q converges almost surely and in L^2. To check this, we first observe that for any $\mathrm{T} \geqslant 2$, we have

$$\begin{aligned}
&\sideset{}{^*}\sum_{\chi \ (\mathrm{mod}\, q)} \sum_{\substack{\mathrm{L}(\frac{1}{2}+i\gamma,\chi)=0 \\ |\gamma|\leqslant \mathrm{T}}} \frac{\mathrm{I}_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)} \\
&= \sideset{}{^*}\sum_{\chi \ (\mathrm{mod}\, q)} \sum_{\substack{\mathrm{L}(\frac{1}{2}+i\gamma,\chi)=0 \\ 0<\gamma\leqslant \mathrm{T}}} \left(\frac{\mathrm{I}_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)} + \frac{\mathrm{I}_{\overline{\chi},-\gamma}}{\frac{1}{2}-i\gamma}\chi(a)\right) \\
&= 2\sideset{}{^*}\sum_{\chi \ (\mathrm{mod}\, q)} \sum_{\substack{\mathrm{L}(\frac{1}{2}+i\gamma,\chi)=0 \\ 0<\gamma\leqslant \mathrm{T}}} \mathrm{Re}\left(\frac{\mathrm{I}_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)}\right)
\end{aligned} \tag{5.22}$$

according to the definition (5.19) of $\mathrm{I}_{\chi,\gamma}$ for negative γ (we use here the fact that, under the Generalized Simplicity Hypothesis, no zero has ordinate $\gamma = 0$).

The right-hand side of (5.22) is the partial sum of a series of independent random variables, and we can apply Kolmogorov's Theorem, B.10.1. Indeed, we have

$$\mathbf{E}\left(\mathrm{Re}\left(\frac{\mathrm{I}_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)}\right)\right) = 0$$

[4] This convergence could be proved without any condition, not even the Generalized Riemann Hypothesis, but the series has no arithmetic meaning without such assumptions.

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

for any pair (χ, γ), and

$$\sum_{\chi \,(\mathrm{mod}\, q)}^{*} \sum_{\gamma > 0} \mathbf{V}\left(\mathrm{Re}\left(\frac{\mathrm{I}_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)}\right)\right) \leqslant \sum_{\chi \,(\mathrm{mod}\, q)}^{*} \sum_{\gamma > 0} \mathbf{E}\left(\left|\frac{\mathrm{I}_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\right|^2\right)$$
$$= \sum_{\chi \,(\mathrm{mod}\, q)}^{*} \sum_{\gamma > 0} \frac{1}{\frac{1}{4}+\gamma^2} < +\infty$$

by Proposition C.5.3 (2), so that the series converges almost surely and in L^2, by Kolmogorov's Theorem, as claimed.

Now we need only go through the steps described above when motivating Definition 5.4.1. The random function N_q is the limit as $\mathrm{T} \to +\infty$ of

$$\mathrm{N}_{q,\mathrm{T}} = m_q - \lim_{\mathrm{X}\to+\infty}\left(\sum_{\chi \,(\mathrm{mod}\, q)}^{*} \sum_{\substack{\mathrm{L}(\frac{1}{2}+i\gamma,\chi)=0 \\ |\gamma|\leqslant \mathrm{T}}} \frac{\mathsf{I}_{\mathrm{X}}^{i\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)}\right).$$

We write once more

$$m_q - \sum_{\chi \,(\mathrm{mod}\, q)}^{*} \sum_{\substack{\mathrm{L}(\frac{1}{2}+i\gamma,\chi)=0 \\ |\gamma|\leqslant \mathrm{T}}} \frac{\mathsf{I}_{\mathrm{X}}^{i\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)}$$
$$= m_q - 2\sum_{\chi \,(\mathrm{mod}\, q)}^{*} \sum_{\substack{\mathrm{L}(\frac{1}{2}+i\gamma,\chi)=0 \\ 0<\gamma\leqslant \mathrm{T}}} \mathrm{Re}\left(\frac{\mathsf{I}_{\mathrm{X}}^{i\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)}\right).$$

By Proposition 5.3.3, or Corollary 5.3.4, as explained above, and the Generalized Simplicity Hypothesis modulo q (precisely through Lemma 5.4.2), the limit as $\mathrm{X} \to +\infty$ of these random functions is simply

$$m_q - 2\sum_{\chi \,(\mathrm{mod}\, q)}^{*} \sum_{\substack{\mathrm{L}(\frac{1}{2}+i\gamma,\chi)=0 \\ 0<\gamma\leqslant \mathrm{T}}} \mathrm{Re}\left(\frac{\mathrm{I}_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\overline{\chi(a)}\right),$$

which in turn converge to the random function N_q as $\mathrm{T} \to +\infty$ by definition. This concludes the proof of Theorem 5.4.4.

Theorem 5.4.4 is equivalent to the computation of the characteristic function of N_q, viewed as a random vector, that is, of the function

$$t \mapsto \mathbf{E}(e^{it\cdot \mathrm{N}_q})$$

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

for $t \in \mathbf{R}^{(\mathbf{Z}/q\mathbf{Z})^\times}$, where

$$t \cdot f = \sum_{a \in (\mathbf{Z}/q\mathbf{Z})^\times} t_a f(a)$$

for $t = (t_a) \in \mathbf{R}^{(\mathbf{Z}/q\mathbf{Z})^\times}$ and $f \colon (\mathbf{Z}/q\mathbf{Z})^\times \to \mathbf{R}$. (Indeed, this is how the result is presented in [105, §3.1].)

To state the formula for the characteristic function, define the Bessel function J_0 on $\mathbf{R}$ by

$$\mathrm{J}_0(x) = \frac{1}{2\pi} \int_0^{2\pi} e^{ix\cos(t)} dt.$$

It is elementary that J_0 is a real-valued and even function of x.

Corollary 5.4.7 *Let $q \geqslant 2$ be an integer. Assume the Generalized Riemann Hypothesis and the Generalized Simplicity Hypothesis modulo q. The characteristic function of the law of the Rubinstein–Sarnak distribution N_q modulo q is given by*

$$\mathbf{E}(e^{it\cdot \mathrm{N}_q}) = \exp(it \cdot m_q) \prod_{\chi \,(\mathrm{mod}\, q)}^{*} \prod_{\substack{\gamma > 0 \\ \mathrm{L}(\frac{1}{2}+i\gamma,\chi)=0}} \mathrm{J}_0\left(\frac{2\,|t \cdot \overline{\chi}|}{(\frac{1}{4}+\gamma^2)^{1/2}}\right)$$

for $t \in \mathbf{R}^{(\mathbf{Z}/q\mathbf{Z})^\times}$, where, for each Dirichlet character χ modulo q, the product runs over the positive ordinates of zeros of $\mathrm{L}(s,\chi)$.

Proof Using the previous argument, we write the series defining N_q in the form

$$\begin{aligned} m_q - \sum_{\chi \,(\mathrm{mod}\, q)}^{*} \sum_{\gamma > 0} \left(\frac{\mathrm{I}_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\,\overline{\chi} + \frac{\mathrm{I}_{\overline{\chi},-\gamma}}{\frac{1}{2}-i\gamma}\,\chi \right) \\ = m_q - 2\,\mathrm{Re}\left(\sum_{\chi \,(\mathrm{mod}\, q)}^{*} \sum_{\gamma>0} \frac{\mathrm{I}_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\,\overline{\chi} \right). \end{aligned}$$

Since the characteristic function of a limit in law is the pointwise limit of the characteristic functions of the sequence involved, we obtain using the independence of the random variables $(\mathrm{I}_{\chi,\gamma})$ the convergent product formula

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

$$\begin{aligned}\mathbf{E}(e^{it\cdot \mathrm{N}_q}) &= e^{it\cdot m_q} \prod\nolimits^*_{\chi \,(\mathrm{mod}\, q)} \prod_{\gamma>0} \mathbf{E}\left(e^{-2it\cdot \mathrm{Re}\left(\frac{\mathrm{I}_{\chi,\gamma}}{1/2+i\gamma}\overline{\chi}\right)}\right)\\ &= e^{it\cdot m_q} \prod\nolimits^*_{\chi \,(\mathrm{mod}\, q)} \prod_{\gamma>0} \varphi\left(\frac{t\cdot\overline{\chi}}{\frac{1}{2}+i\gamma}\right),\end{aligned}$$

where, for $z \in \mathbf{C}$, we defined

$$\varphi(z) = \mathbf{E}\left(e^{-2i\,\mathrm{Re}(z\mathrm{I})}\right)$$

for a random variable I uniformly distributed over the unit circle. By invariance of the law of I under rotation (i.e., the law of $ze^{i\theta}\mathrm{I}$ is the same as that of $z\mathrm{I}$ for any $\theta \in \mathbf{R}$), applied to the angle θ such that $ze^{i\theta} = |z|$, we have

$$\varphi(z) = \mathbf{E}(e^{-2i\,\mathrm{Re}(|z|\mathrm{I})}) = \mathbf{E}(e^{-2i|z|\,\mathrm{Re}(\mathrm{I})}) = \frac{1}{2\pi}\int_0^{2\pi} e^{-2i|z|\cos(t)}dt = \mathrm{J}_0(2|z|).$$

Hence we obtain

$$\mathbf{E}(e^{it\cdot \mathrm{N}_q}) = e^{it\cdot m_q} \prod\nolimits^*_{\chi \,(\mathrm{mod}\, q)} \prod_{\gamma>0} \mathrm{J}_0\left(2\frac{|t\cdot\overline{\chi}|}{|\frac{1}{2}+i\gamma|}\right),$$

as claimed. □

Another consequence of Theorem 5.4.4 is an estimate for the probability that N_q takes large values.

Corollary 5.4.8 *There exists a constant $c_q > 0$ such that, for $\mathrm{A} > 0$, we have*

$$\begin{aligned}c_q^{-1}\exp(-\exp(c_q\mathrm{A}^{1/2})) &\leqslant \liminf_{\mathrm{X}\to+\infty} \mathbf{P}_{\mathrm{X}}(\|\mathsf{N}_{\mathrm{X},q}\| \geqslant \mathrm{A})\\ &\leqslant \limsup_{\mathrm{X}\to+\infty} \mathbf{P}_{\mathrm{X}}(\|\mathsf{N}_{\mathrm{X},q}\| > \mathrm{A}) \leqslant c_q\exp(-\exp(c_q^{-1}\mathrm{A}^{1/2})).\end{aligned}$$

Proof We view N_q as a random variable with values in the complex finite-dimensional Banach space of complex-valued functions on $(\mathbf{Z}/q\mathbf{Z})^\times$. We have the series representation

$$\mathrm{N}_q = m_q - 2\sum\nolimits^*_{\chi \,(\mathrm{mod}\, q)} \sum_{\substack{\gamma>0\\ \mathrm{L}(\frac{1}{2}+i\gamma,\chi)=0}} \mathrm{Re}\left(\frac{\mathrm{I}_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\,\overline{\chi}\right).$$

This series converges almost surely, the terms are independent and the random variables $\mathrm{I}_{\chi,\gamma}$ are bounded by 1 in modulus. Moreover,

$$\mathbf{P}(\|\mathrm{N}_q\| > \mathrm{A}) \leqslant \mathbf{P}(\|\widetilde{\mathrm{N}}_q\| > \mathrm{A}),$$

https://doi.org/10.1017/9781108888226.006 Published online by Cambridge University Press

where

$$\widetilde{\mathrm{N}}_q = m_q - 2 \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q)} \sum_{\substack{\gamma>0 \\ \mathrm{L}(\frac{1}{2}+i\gamma,\chi)=0}} \frac{\mathrm{I}_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\,\overline{\chi},$$

since $\|\mathrm{N}_q\| \leqslant \|\widetilde{\mathrm{N}}_q\|$. By Corollary C.5.5, the functions

$$-\frac{2}{\frac{1}{2}+i\gamma}\overline{\chi}$$

satisfy the bounds described in Remark B.11.14 (2), namely,

$$\sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q)} \sum_{\substack{0<\gamma<\mathrm{T} \\ \mathrm{L}(\frac{1}{2}+i\gamma,\chi)=0}} \left\| \frac{1}{\frac{1}{2}+i\gamma}\,\overline{\chi} \right\| \gg (\log \mathrm{T})^2$$

and

$$\sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q)} \sum_{\substack{\gamma>\mathrm{T} \\ \mathrm{L}(\frac{1}{2}+i\gamma,\chi)=0}} \left\| \frac{1}{\frac{1}{2}+i\gamma}\,\overline{\chi} \right\|^2 \ll \frac{\log \mathrm{T}}{\mathrm{T}}$$

for $\mathrm{T} \geqslant 1$. Thus by Remark B.11.14 (2), and the convergence in law of $\mathsf{N}_{\mathrm{X},q}$ to N_q, we deduce the upper bound

$$\limsup_{\mathrm{X}\to+\infty} \mathbf{P}_{\mathrm{X}}(\|\mathsf{N}_{\mathrm{X},q}\| > \mathrm{A}) \leqslant \mathbf{P}(\|\mathrm{N}_q\| > \mathrm{A}) \leqslant c\exp(-\exp(c^{-1}\mathrm{A}^{1/2}))$$

for some real number $c > 0$.

In the case of the lower bound, it suffices to prove it for $\mathrm{N}_q(a)$, where a is any fixed element of $(\mathbf{Z}/q\mathbf{Z})^\times$. Since the series expressing $\mathrm{N}_q(a)$ is not exactly of the form required for the lower bound in Remark B.11.14 (2) (and in Proposition B.11.13), we first transform it a bit. We have

$$\mathrm{Re}\left(\frac{\mathrm{I}_{\chi,\gamma}}{\frac{1}{2}+i\gamma}\,\overline{\chi(a)}\right) = \frac{1}{2(\frac{1}{4}+\gamma^2)}\,\mathrm{Re}(\mathrm{I}_{\chi,\gamma}\overline{\chi(a)}) + \frac{\gamma}{\frac{1}{4}+\gamma^2}\,\mathrm{Im}(\mathrm{I}_{\chi,\gamma}\overline{\chi(a)})$$

for any pair (χ,γ), which implies that

$$\mathrm{N}_q(a) = m_q(a) + e_q(a) - 2 \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q)} \sum_{\substack{\gamma>0 \\ \mathrm{L}(\frac{1}{2}+i\gamma,\chi)=0}} \frac{\gamma}{\frac{1}{4}+\gamma^2}\,\mathrm{Im}(\mathrm{I}_{\chi,\gamma}\overline{\chi(a)}),$$

where the random variable $e_q(a)$ (arising from the sum of the first terms in the previous expression) is uniformly bounded (by Proposition C.5.3 (2)). Now we can apply the lower bound in Remark B.11.14 (2) to the last series: the random variables $\mathrm{Im}(\mathrm{I}_{\chi,\gamma}\overline{\chi(a)})$ are independent, symmetric and bounded by 1, and the assumptions on the size of the coefficients are provided by Corollary C.5.5 again. □

5.5 Further Results

In recent years, the Chebychev bias has been a popular topic in analytic number theory; besides further studies of the original setting that we have discussed, it has also been generalized in many ways. We only indicate a few examples here, without any attempt to completeness.

In the first direction, there have been many studies of the properties of the Rubinstein–Sarnak measures, and of the consequences concerning various "races" between primes (see, for instance, the papers of Granville and Martin [50] and Harper and Lamzouri [56]). In parallel, attempts have been made to weaken the assumptions used by Rubinstein and Sarnak to establish properties of their measures (recall that the *existence* of the measure does not require the Generalized Simplicity Hypothesis). Among these, we refer in particular to the work of Devin [26], who found a much weaker condition that ensures that the Rubinstein–Sarnak measure is absolutely continuous.

Among generalizations, it seems worth mentioning the discussion by Sarnak [107] of a bias related to elliptic curves, as well as the recent extensive work of Fiorilli and Jouve [38] concerning Artin L-functions. In another direction, Kowalski [68] and later Cha–Fiorilli–Jouve [23] have considered analogue questions over finite fields, where the main difference is that relations between zeros of the analogues of the Dirichlet L-functions may well exist (although they are rare), leading to new phenomena.