Chebyshev's Bias

Michael Rubinstein and Peter Sarnak

CONTENTS

The title refers to the fact, noted by Chebyshev in 1853, that primes congruent to 3 modulo 4 seem to predominate over those congruent to 1. We study this phenomenon and its generalizations. Assuming the Generalized Riemann Hypothesis and the Grand Simplicity Hypothesis (about the zeros of the Dirichlet L-function), we can characterize exactly those moduli and residue classes for which the bias is present. We also give results of numerical investigations on the prevalence of the bias for several moduli. Finally, we briefly discuss generalizations of the bias to the distribution to primes in ideal classes in number fields, and to prime geodesics in homology classes on hyperbolic surfaces.

1. INTRODUCTION

Dirichlet [1837] proved that for any a and q with $(a, q) = 1$ there are infinitely many primes p with $p \equiv a \bmod q$, and that they are roughly equidistributed amongst these residue classes. We denote the set of such residue classes by A_q. It was later proved by Hadamard and de la Vallée Poussin that the number $\pi(x, q, a)$ of primes $p \leq x$ with $p \equiv a \bmod q$ has the behavior

$$\pi(x, q, a) \sim \frac{\mathrm{Li}(x)}{\varphi(q)} \sim \frac{1}{\varphi(q)} \frac{x}{\log x}$$

as $x \to \infty$, where $\varphi(q) = |A_q|$ is the Euler phi function and

$$\mathrm{Li}(x) = \int_2^x \frac{dt}{\log t}.$$

Chebyshev noted in 1853 that there are many more primes congruent to 3 than 1 modulo 4. Much has been written about this since then, but we have found the literature to be a little confused and inaccurate. We have, therefore, tried our best to cite below the original sources where appropriate. A

1058-6458/94 $0.50 per page

good survey appears in [Kaczorowski]. In this paper we take a somewhat different point of view in our attempt to analyze Chebyshev's phenomenon and its generalizations, which we call "Chebyshev's bias". Our purpose has been to examine these issues both theoretically and numerically and, in particular, to give numerical values to these biases.

Let $a_1, a_2, \dots, a_r \in A_q$ be distinct, and define $P_{q;a_1,\dots,a_r}$ as the set of real $x \geq 2$ such that

$$\pi(x,q,a_1) > \pi(x,q,a_2) > \cdots > \pi(x,q,a_r).$$

Are these sets nonempty? What are their densities? A definitive result of Littlewood [1914] asserts that both $P_{4;1,3}$ and $P_{4;3,1}$ extend to infinity and the same is true of $P_{3;1,2}$ and $P_{3;2,1}$. Leech [1957] found the first member of $P_{4;1,3}$. It is 26861 and indicates the bias of primes towards 3 mod 4. The first member of $P_{3;1,2}$ is 608981813029, and was found by Bays and Hudson [1978]. Evidently, strong biases occur in the initial segments and one asks whether they persist or perhaps even grow.

Define $\delta(P)$ to be the logarithmic density of P, that is, set

$$\overline{\delta}(P) := \limsup_{X\to\infty} \frac{1}{\log X} \int_{t\in P\cap[2,X]} \frac{dt}{t},$$

$$\underline{\delta}(P) := \liminf_{X\to\infty} \frac{1}{\log X} \int_{t\in P\cap[2,X]} \frac{dt}{t},$$

and set $\delta(P) = \overline{\delta}(P) = \underline{\delta}(P)$ if the latter two limits are equal. That the logarithmic density is the appropriate one to use here is well known [Wintner 1941] and will become clear in Section 2; suffice it to say that the usual densities of our $P_{q;a_1,\dots,a_r}$'s do not exist. We will see that, on certain natural hypotheses, $\delta(P_{4;3,1}) = 0.9959\dots$ and $\delta(P_{3;2;1}) = 0.9990\dots$, showing very strong biases.

In order to investigate these densities and biases we introduce the vector-valued function

$$E_{q;a_1,\dots,a_r}(x) = \frac{\log x}{\sqrt{x}} \times (\varphi(q)\pi(x,q,a_1) - \pi(x),\ \dots,\ \varphi(q)\pi(x,q,a_r) - \pi(x))$$

for $x \geq 2$. The normalization is such that, if we assume the Generalized Riemann Hypothesis (GRH), as we shall do throughout this paper, $E_{q;a_1,\dots,a_r}(x)$ varies roughly boundedly (see Section 2). With this normalization, $E_{q;a_1,\dots,a_r}$ has a limiting distribution.

Theorem 1.1 (see Section 2.1). *Assume GRH. Then $E_{q;a_1,\dots,a_r}$ has a limiting distribution $\mu_{q;a_1,\dots,a_r}$ on $\mathbb{R}^r$, that is,*

$$\lim_{X\to\infty} \frac{1}{\log X} \int_2^X f(E_{q;a_1,\dots,a_r}(x)) \frac{dx}{x} = \int_{\mathbb{R}^r} f(x)\, d\mu_{q;a_1,\dots,a_r}(x)$$

for all bounded continuous functions f on $\mathbb{R}^r$.

Special one-dimensional cases of this theorem (in somewhat different forms) have been known for some time. See [Wintner 1938], and more recently [Kueh 1988; Heath-Brown 1992].

Note that, if $r = \varphi(q)$ in Theorem 1.1, the sum of the components of $E_{q;a_1,\dots,a_r}$ is $O(\log x/\sqrt{x})$. It follows that in this case $\mu_{q;a_1,\dots,a_r}$ is supported on the hyperplane $\sum_{j=1}^r x_j = 0$.

The measures $\mu_{q;a_1,\dots,a_r}$ carry all the information concerning the densities and biases that we are interested in, and we seek to understand their shapes, means, and so on. For example, if $\mu_{q;a_1,\dots,a_r}$ is absolutely continuous, we have

$$\delta(P_{q;a_1,\dots,a_r}) = \mu_{q;a_1,\dots,a_r}(\{x \in \mathbb{R}^r \mid x_1 > x_2 > \cdots > x_r\}).$$

Note, however, that assuming only GRH we don't know that $\delta(P_{q;a_1,\dots,a_r})$ exists, since we have not been able to establish Theorem 1.1 with f a characteristic function of a nice set.

The proof of Theorem 1.1 also yields a method of approximating μ. In Section 2 we construct measures $\mu^T_{q;a_1,\dots,a_r}$ defined in terms of the zeros $\frac{1}{2}+i\gamma_\chi$ of the L-functions $L(s,\chi)$, where χ runs over the Dirichlet characters modulo q with $|\gamma_\chi| \leq T$. These measures satisfy

$$\left|\mu_{q;a_1,\dots,a_r}(f) - \mu^T_{q;a_1,\dots,a_r}(f)\right| \ll_q c_f \frac{\log T}{\sqrt{T}},$$

where f is Lipschitz with constant c_f, and the notation $x \ll_\tau y$ means $x \le ky$ with k depending on τ only.

Concerning the localization of the measures μ, we can show that they are very localized but not compactly supported. Set

$$B'_R = \{x \in \mathbb{R}^r \mid |x| \ge R\},$$
$$B^+_R = \{x \in B'_R \mid \varepsilon(a_j)x_j > 0\},$$
$$B^-_R = -B^+_R,$$

where $\varepsilon(a) = 1$ if $a \equiv 1 \bmod q$ and $\varepsilon(a) = -1$ otherwise.

Theorem 1.2 (see Section 2.2). *Assume GRH. There are positive constants* c_1, c_2, c_3 *and* c_4, *depending only on* q, *such that*

$$\mu_{q;a_1,\dots,a_r}(B'_R) \le c_1 \exp(-c_2\sqrt{R}),$$
$$\mu_{q;a_1,\dots,a_r}(B^\pm_R) \ge c_3 \exp(-\exp c_4 R).$$

This theorem asserts that the tails of the distributions are "exponentially" small. However, it is the double exponential lower bound that is presumably closer to the true size, as will be seen later. Note that it is only for special sets like $B^\pm_R$ that we can establish a nonzero lower bound. It seems quite difficult (without further hypotheses) to show that all orthants have positive mass (if, say, $r < \varphi(q)$); see Remark 2.5.

We mention two related cases where Theorems 1.1 and 1.2 apply:

First, the case "$q = 1$", concerning the density of

$$P_1 = \{x \ge 2 \mid \pi(x) > \mathrm{Li}(x)\}.$$

Again, Littlewood [1914] showed that P_1 extends to infinity. However, in this case the bias is so keen that no member of P_1 is known. It is known [Skewes 1933; te Riele 1986] that the first member of P_1 is at most 10^{370}. Denote by μ_1 the limiting distribution of

$$E_1(x) := (\pi(x) - \mathrm{Li}(x))\frac{\log x}{\sqrt{x}}.$$

Then, assuming the Riemann Hypothesis, we have, for $\lambda \gg 1$,

$$c_7 \exp(-\exp(c_8\lambda)) \le \mu_1[\lambda, \infty) \le c_5 \exp(-c_6\sqrt{\lambda}),$$
$$c_7 \exp(-\exp(c_8\lambda)) \le \mu_1(-\infty, -\lambda] \le c_5 \exp(-c_6\sqrt{\lambda}) \tag{1.1}$$

for absolute positive constants c_5, c_6, c_7 and c_8. In particular we have $\underline{\delta}(P_1) > 0$; this may also be deduced from [Wintner 1941]. We will see later that $\delta(P_1) = .00000026\dots$. So, although the initial segment in which $\pi(x)$ loses to $\mathrm{Li}(x)$ is extremely long, the probability that $\pi(x)$ beats $\mathrm{Li}(x)$, while very small, is still palpable.

The second case concerns the excess of primes that are quadratic residues over those that are nonresidues, a problem recently analyzed in [Davidoff 1994]. For $q = p^\alpha$, $q = 2p^\alpha$ or $q = 4$, where p is an odd prime and α is a positive integer, let

$$P_{q;N,R} = \{x \ge 2 \mid \pi_N(x,q) > \pi_R(x,q)\},$$
$$P_{q;R,N} = \{x \ge 2 \mid \pi_R(x,q) > \pi_N(x,q)\},$$

where $\pi_R(x,q)$ is the number of prime quadratic residues not exceeding x and $\pi_N(x,q)$ is the number of prime quadratic nonresidues not exceeding x. We will see that there is always a bias towards nonresidues. From now on, when we write $q; N, R$ it will be understood that q is of the form p^α, $2p^\alpha$, or 4.

As in Theorem 1.2 and in the estimates (1.1), one can give lower bounds for the tails of the limiting distribution $\mu_{q;N,R}$ of

$$E_{q;N,R} := (\pi_N(x,q) - \pi_R(x,q))\frac{\log x}{\sqrt{x}}$$

(see Section 2.2). Consequently we have

$$\underline{\delta}(P_{q;R,N})\underline{\delta}(P_{q;N,R}) > 0.$$

In particular (under GRH), we have $\underline{\delta}(P_{4;1,3}) > 0$, hence the usual density of $P_{4;3,1}$ relative to the usual measure dx cannot be zero (we note that $\underline{\delta}(P_{4;1,3}) > 0$ could also be deduced by the method in [Wintner 1941]). This is contrary to a conjecture in [Knapowsky and Turan 1962]. Kaczorowski has also recently disproved the aforementioned conjecture by a somewhat different approach based on his K-functions [Kaczorowski]. He also gives some numerical approximations to the unequal upper and lower (usual) densities of $P_{4;3,1}$.

To further analyze the measures $\mu_{q;a_1,\dots,a_r}$, we need to make some further hypotheses about the zeros of the L-functions $L(s,\chi)$. In practice, we can verify these hypotheses to the extent that we want to approximate the measures $\mu^T_{q;a_1,\dots,a_r}$ and hence $\mu_{q;a_1,\dots,a_r}$. So we view the following as a working hypothesis. It appears that the first person to propose (or realize the significance of) the hypothesis below, at least for $\zeta(s)$, was Wintner [1938, Ch. 13; 1941]. Later authors [Hooley 1977; Montgomery 1980] introduced it for similar purposes.

Grand Simplicity Hypothesis (GSH)**.** The set of $\gamma \geq 0$ such that $L(\frac{1}{2}+i\gamma,\chi)=0$, for χ running over primitive Dirichlet characters, is linearly independent over $\mathbb{Q}$.

Note that GSH implies that all the zeros are simple and that $L(\frac{1}{2},\chi)\neq 0$ for all such χ. A lot of evidence exists for the last two statements, while numerical evidence for GSH is more modest [Odlyzko; Rumely 1993]. In any event, there is no reason to suspect that different zeros satisfy any relations, and this is the main rationale for believing GSH. For more general L-functions GSH may fail, but in a predictable way: see Section 5.

Montgomery [1980], using GSH for the Riemann zeta function, investigated the sizes of the tails of μ_1. He showed that

$$\exp(-c_2\sqrt{R}\exp\sqrt{2\pi R}) \leq \mu_1(B_R^{\pm}) \leq \exp(-c_1\sqrt{R}\exp\sqrt{2\pi R}).$$

In particular, we see that the double exponential lower bound of Theorem 1.2, which depends only on GRH and not on GSH, is closer to the true size of the tails.

Under GRH and GSH we have the following explicit formula for the Fourier transform of $\mu_{q;a_1,\dots,a_r}$ (see Section 3.1):

$$\hat{\mu}_{q;a_1,\dots,a_r}(\xi_1,\dots,\xi_r) = \exp\Bigl(i\sum_{j=1}^{r} c(q,a_j)\xi_j\Bigr) \times \prod_{\substack{\chi\neq\chi_0\\ \chi \bmod q}} \prod_{\gamma_\chi>0} J_0\biggl(\frac{2\left|\sum_{j=1}^r \chi(a_j)\xi_j\right|}{\sqrt{\frac{1}{4}+\gamma_\chi^2}}\biggr), \tag{1.2}$$

where χ_0 is the principal character,

$$c(q,a) = -1 + \sum_{\substack{b^2\equiv a(q)\\ 0\leq b\leq q-1}} 1, \tag{1.3}$$

and $J_0(z)$ is the Bessel function

$$J_0(z) = \sum_{m=0}^{\infty} \frac{(-1)^m(\frac{1}{2}z)^{2m}}{(m!)^2}.$$

Again, special cases of (1.2) are known [Wintner 1941; Hooley 1977].

The infinite product in (1.2) converges absolutely for fixed $(\xi_1,\dots,\xi_r)$. This follows from the expansion $J_0(z)=1-\frac{1}{4}z^2+\cdots$ at $z=0$ and the fact that $\sum_{\gamma_\chi} 1/(\frac{1}{4}+\gamma_\chi^2)<\infty$. If one is equipped with many zeros of $L(s,\chi)$, one can use (1.2) and some variations thereof to compute $\mu_{q;a_1,\dots,a_r}$ and $\delta_{q;a_1,\dots,a_r}$. This is carried out in Section 4 and is the basis for our computations of the δ's.

Remark 1.3. If $r<\varphi(q)$, we can easily deduce from (1.2) that $\hat{\mu}_{q;a_1,\dots,a_r}(\xi)$ is rapidly decreasing as $|\xi|\to\infty$. It follows that $\mu_{q;a_1,\dots,a_r}=f(x)\,dx$ with $f(x)$ rapidly decreasing, and even that f is entire. Also $f(x)$ is doubly exponentially localized. If $r=\varphi(q)$ then $\hat{\mu}(\xi)=\hat{\mu}(\xi+\lambda(1,\dots,1))$ for any $\lambda\in\mathbb{R}$. This implies that $\mu_{q;a_1,\dots,a_r}$ is supported on $\sum_{j=1}^r x_j=0$, a fact we have already noted. In this case, $\hat{\mu}(\xi)$ is rapidly decreasing as $|\xi|\to\infty$ as long as $\xi\perp(1,\dots,1)$. So again $\mu_{q;a_1,\dots,a_r}=f(x)\,dV(x)$, where dV is the volume form on $\sum_{j=1}^r x_j=0$ and f is analytic and localized. In either case it follows, under GRH and GSH, that $\delta(P_{q;a_1,\dots,a_r})$ exists and is nonzero, answering in particular the question about whether $P_{q;a_1,\dots,a_r}$ is nonempty.

The first factor in (1.2) causes a shift in the mean of the distribution $\mu_{q;a_1,\dots,a_r}$, placing it at $-(c(q,a_1),$

$\ldots, c(q, a_r))$. This is the source of the Chebyshev bias. For $q = 3, 4$ it leads to the bias towards nonresidues (primes congruent to 2 mod 3 and 3 mod 4) discussed on page 1.

A closer investigation of the symmetries of the density function of $\mu_{q;a_1,\ldots,a_r}$ in $(x_1, \ldots, x_r)$ reveals the nature of the biases. We say that

$$(q; a_1, \ldots, a_r)$$

is unbiased, or that the Rényi–Shanks primes race [Shanks 1959] is unbiased, if the density function of $\mu_{q;a_1,\ldots,a_r}$ is invariant under permutations of $(x_1, \ldots, x_r)$. In this case we have $\delta(P_{q;a_1,\ldots,a_r}) = (r!)^{-1}$. (The converse seems very plausible as well.)

Theorem 1.4 (see Proposition 3.1). *Under GRH and GSH, $(q; a_1, \ldots, a_r)$ is unbiased if and only if either $r = 2$ and $c(q, a_1) = c(q, a_2)$, where c is defined by (1.3), or $r = 3$ and there exists $\rho \neq 1$ such that $\rho^3 \equiv 1 \bmod q$, $a_2 \equiv a_1\rho \bmod q$ and $a_3 \equiv a_1\rho^2 \bmod q$.*

So with the aid of GSH we can in essence completely resolve the issue of the existence of a bias. The symmetry analysis of $\mu_{q;N,R}$ shows that $\frac{1}{2} < \delta(P_{q;N,R}) < 1$, that is, there is always a bias towards nonresidues. At the beginning of Section 4 we list these densities for $q = 3, 4, 5, 7, 11, 13$.

From (1.2) we can also find out what happens as $q \to \infty$. Interestingly, all biases disappear and a central limit behavior emerges. From (1.2) we can determine the covariance matrix of the distribution $\mu_{q;a_1,\ldots,a_r}$. Its entries are

$$b_{a_i,a_j} = \sum_{\substack{\chi \neq \chi_0 \\ \chi \bmod q}} \chi\Big(\frac{a_i}{a_j}\Big) B(\chi^*),$$

where χ^* is the primitive character inducing χ and where

$$B(\chi^*) = \tfrac{1}{2}\log\Big(\frac{q_{\chi^*}}{\pi}\Big) + \frac{\Gamma'}{2\Gamma}\Big(\frac{3-\chi^*(-1)}{4}\Big) + \frac{L'}{L}(1, \chi^*). \tag{1.4}$$

In view of a theorem of Littlewood [1928] we have $L'/L(1, \chi^*) = O(\log\log q_{\chi^*})$, still assuming GRH, so $B(\chi^*)$ is dominated by the $\log q_{\chi^*}$ term. This growth in the variance is responsible for dissolving the bias.

Theorem 1.5 (see Section 3.2). *Assume GRH and GSH. Then, for r fixed,*

$$\max_{a_1,\ldots,a_r \in A_q} \left|\delta(P_{q;a_1,\ldots,a_r}) - \frac{1}{r!}\right| \to 0 \text{ as } q \to \infty.$$

Even in the extreme case of $P_{q;N,R}$, where all the residues and nonresidues are grouped separately, the bias dissolves: that is, $\delta(P_{q;N,R}) \to \frac{1}{2}$ as $q \to \infty$. In fact we have the following central limit theorem.

Theorem 1.6 (see Section 3.2). *Assume GRH and GSH. Let $\tilde{\mu}_{q;N,R}$ be the limiting distribution of*

$$\frac{E_{q;N,R}(x)}{\sqrt{\log q}}.$$

Then $\tilde{\mu}_{q;N,R}$ converges in measure to the Gaussian $(2\pi)^{-1/2}e^{-x^2/2}\,dx$ as $q \to \infty$.

A central limit theorem for a close relative of $\mu_{q;a_1}$ as $q \to \infty$ was derived in [Hooley 1977].

In Section 4 we give the results of our numerical investigations into the issues discussed above. For the computations of the measures μ we used thousands of zeros of $\zeta(s)$ and $L(s, \chi_1)$, where χ_1 is the nonprincipal real character mod q. The values of $\zeta(s)$ were provided to us by Odlyzko and te Riele, and those of $L(s, \chi_1)$ by Rumely. At the beginning of Section 4 we give the values of $\delta(P_1^{\text{comp}})$ and $\delta(P_{q;N,R})$, for $q = 3, 4, 5, 7, 11, 13$. As expected, the bias is most extreme for $q = 1$ (that is, $\pi(x)$ vs. $\mathrm{Li}(x)$), and decreases, albeit not steadily, as q increases. Figure 1 shows graphs of the distributions μ_1 and $\mu_{q;R,N}$, for $q = 3, 4, 5, 7, 11, 13$, comparing the curves predicted by the use of (1.2) and our table of zeros with numerical distributions involving primes up to 10^{10}. The fits are quite good and the dissolving bias and central limit behavior are already present.

In Section 5 we briefly discuss generalizations of the bias of distribution to primes in ideal classes in number fields and to prime geodesics in homology classes on hyperbolic surfaces.

2. APPLICATIONS OF THE GENERALIZED RIEMANN HYPOTHESIS

2.1. Existence of the Limiting Distribution

The main tool in establishing the existence of the limiting distribution μ is the explicit formula of Riemann relating $\pi(x,q,a)$ to zeros of Dirichlet L-functions $L(s,\chi)$. Fix q and let χ run over the Dirichlet characters modulo q, with χ_0 the principal character. Set

$$\psi(x,\chi) := \sum_{n\le x} \chi(n)\Lambda(n),$$

where $\Lambda(n) = \log p$ if $n = p^m$ for some $m \in \mathbb{Z}$ and $\Lambda(n) = 0$ otherwise. As is shown in [Davenport 1980, pp. 115–120], if $\chi \neq \chi_0$, $x \ge 2$ and $X \ge 1$ we have

$$\psi(x,\chi) = -\sum_{|\gamma_\chi|\le X} \frac{x^\rho}{\rho} + O\biggl(\frac{x\log^2(xX)}{X} + \log x\biggr),$$

where $\rho = \beta_\chi + i\gamma_\chi$ runs over the zeros of $L(s,\chi)$ in $0 < \mathrm{Re}(s) < 1$, and the implied constant in the O depends on q. Since we are assuming the Riemann Hypothesis for $L(s,\chi)$, we have $\beta_\chi = \frac{1}{2}$ and the preceding equation becomes

$$\psi(x,\chi) = -\sqrt{x}\sum_{|\gamma_\chi|\le X} \frac{x^{i\gamma_\chi}}{\frac{1}{2}+i\gamma_\chi} + O\biggl(\frac{x\log^2(xX)}{X} + \log x\biggr). \tag{2.1}$$

We recall some notations from Section 1. For a and q relatively prime, let $\pi(x,q,a)$ be the number of primes $p \le x$ with $p \equiv a \bmod q$, set

$$E(x,q,a) = (\varphi(q)\pi(x,q,a) - \pi(x))\,\frac{\log x}{\sqrt{x}}$$

(where φ is the Euler function), and let $c(q,a)$ be given by (1.3) (when $q = p^\alpha$, $q = 2p^\alpha$ or $q = 4$ this is the nonprincipal real character $\chi_1(a)$). Also,

$$\begin{aligned}\psi(x,q,a) &:= \sum_{\substack{n\le x\\ n\equiv a \bmod q}} \Lambda(n) = \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \bar\chi(a) \sum_{n\le x}\Lambda(n)\chi(n)\\ &= \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \bar\chi(a)\psi(x,\chi).\end{aligned} \tag{2.2}$$

Lemma 2.1. *As $x \to \infty$ we have*

$$E(x,q,a) = -c(q,a) + \sum_{\chi\neq\chi_0} \bar\chi(a)\frac{\psi(x,\chi)}{\sqrt{x}} + O\Bigl(\frac{1}{\log x}\Bigr).$$

We remark that the constant term $-c(q,a)$ is what accounts for the bias towards nonresidues.

Proof. Let $\theta(x,q,a) = \sum_{\substack{p\le x\\ p\equiv a(q)}} \log p$. Then

$$\pi(x,q,a) = \int_2^x \frac{d\theta(t,q,a)}{\log t},$$

and, from Dirichlet's theorem for progressions,

$$\psi(x,q,a) = \theta(x,q,a) + \Bigl(\sum_{b^2\equiv a(q)} 1\Bigr)\frac{\sqrt{x}}{\varphi(q)} + O\Bigl(\frac{\sqrt{x}}{\log x}\Bigr).$$

Solving for $\theta(x,q,a)$ and combining with (2.2), we get

$$\begin{aligned}\int_2^x \frac{d\theta(t,q,a)}{\log t} &= \frac{1}{\varphi(q)}\int_2^x \frac{d\psi(t)}{\log t} + \frac{1}{\varphi(q)}\sum_{\chi\neq\chi_0}\bar\chi(a)\int_2^x\frac{d\psi(t,\chi)}{\log t} - \frac{1}{\varphi(q)}\Bigl(\sum_{b^2\equiv a(q)}1\Bigr)\frac{\sqrt{x}}{\log x} + O\Bigl(\frac{\sqrt{x}}{\log^2 x}\Bigr)\\ &= \frac{1}{\varphi(q)}\Bigl(\pi(x) + \frac{\sqrt{x}}{\log x}\Bigr) + \frac{1}{\varphi(q)}\sum_{\chi\neq\chi_0}\bar\chi(a)\frac{\psi(x,\chi)}{\log x} - \frac{1}{\varphi(q)}\Bigl(\sum_{b^2\equiv a(q)}1\Bigr)\frac{\sqrt{x}}{\log x}\\ &\qquad + O\biggl(\sum_{\chi\neq\chi_0}\biggl|\int_2^x \frac{\psi(t,\chi)}{t\log^2 t}\,dt\biggr| + \frac{\sqrt{x}}{\log^2 x}\biggr)\\ &= \frac{\pi(x)}{\varphi(q)} + \frac{1}{\varphi(q)}\sum_{\chi\neq\chi_0}\bar\chi(a)\frac{\psi(x,\chi)}{\log x} - \frac{c(q,a)}{\varphi(q)}\frac{\sqrt{x}}{\log x} + O\biggl(\sum_{\chi\neq\chi_0}\biggl|\int_2^x \frac{\psi(t,\chi)}{t\log^2 t}\,dt\biggr| + \frac{\sqrt{x}}{\log^2 x}\biggr).\end{aligned} \tag{2.3}$$

Let $G(x,\chi) = \int_2^x \psi(t,\chi)\,dt$. Then, from (2.1), after integrating and letting $X \to \infty$, we have

$$G(x,\chi) = -\sum_{\gamma_\chi} \frac{x^{3/2+i\gamma_\chi}}{(\frac{1}{2}+i\gamma_\chi)(\frac{3}{2}+i\gamma_\chi)} + O(x\log x)$$

It is a crucial point throughout our analysis that this series over γ_χ converges absolutely, as is apparent from the asymptotic formula for the number of zeros [Davenport 1980, p. 101]:

$$\#\{|\gamma_\chi| \le T\} = \frac{T}{\pi}\log\frac{qT}{2\pi} - \frac{T}{\pi} + O(\log T + \log q), \tag{2.4}$$

and so it follows that $G(x,\chi) \ll_q x^{3/2}$, with the constant depending on q. Hence, after an integration by parts, the O term in (2.3) is $O(\sqrt{x}/\log^2 x)$, and (2.3) becomes

$$\pi(x,q,a) - \frac{\pi(x)}{\varphi(q)} = -\frac{c(q,a)}{\varphi(q)}\frac{\sqrt{x}}{\log x} + \frac{1}{\varphi(q)\log x}\sum_{\chi\ne\chi_0}\bar\chi(a)\psi(x,\chi) + O\Big(\frac{\sqrt{x}}{\log^2 x}\Big).$$

This completes the proof. □

Combining Lemma 2.1 with (2.1) we get, for $T \ge 1$ and $2 \le x \le X$:

$$E(x,q,a) = -c(q,a) - \sum_{\chi\ne\chi_0}\bar\chi(a)\sum_{|\gamma_\chi|\le T}\frac{x^{i\gamma_\chi}}{\frac{1}{2}+i\gamma_\chi} + \varepsilon_a(x,T,X), \tag{2.5}$$

where

$$\varepsilon_a(x,T,X) = -\sum_{\chi\ne\chi_0}\bar\chi(a)\sum_{T\le|\gamma_\chi|\le X}\frac{x^{i\gamma_\chi}}{\frac{1}{2}+i\gamma_\chi} + O_q\Big(\frac{\sqrt{x}\log^2 X}{X} + \frac{1}{\log x}\Big). \tag{2.6}$$

Now set $y = \log x$, so that $dy = dx/x$.

Lemma 2.2. *For $T \ge 1$ and $Y \ge \log 2$,*

$$\int_{\log 2}^{Y} |\varepsilon_a(e^y,T,e^Y)|^2\,dy \ll_q Y\frac{\log^2 T}{T} + \frac{\log^3 T}{T}.$$

Proof. We have

$$\begin{aligned}
&\int_{\log 2}^{Y}|\varepsilon_a(e^y,T,e^Y)|^2\,dy \\
&\ll \int_{\log 2}^{Y}\Big|\sum_{\chi\ne\chi_0}\bar\chi(a)\sum_{T\le|\gamma_\chi|\le e^Y}\frac{e^{iy\gamma_\chi}}{\frac{1}{2}+i\gamma_\chi}\Big|^2 dy + O(1) \\
&= \sum_{\substack{\chi\ne\chi_0\\ \lambda\ne\chi_0}}\sum_{\substack{T\le|\gamma_\chi|\le e^Y\\ T\le|\gamma_\lambda|\le e^Y}}\bar\chi(a)\lambda(a)\int_{\log 2}^{Y}\frac{e^{iy(\gamma_\chi-\gamma_\lambda)}}{(\frac{1}{2}+i\gamma_\chi)(\frac{1}{2}-i\gamma_\lambda)}\,dy \\
&\qquad + O(1) \\
&\ll_q \sum_{\substack{\chi\ne\chi_0\\ \lambda\ne\chi_0}}\sum_{\substack{T\le|\gamma_\chi|\le\infty\\ T\le|\gamma_\lambda|\le\infty}}\frac{1}{|\gamma_\chi||\gamma_\lambda|}\min\Big(Y,\frac{1}{|\gamma_\chi-\gamma_\lambda|}\Big).
\end{aligned}$$

Using (2.4) and comparing the sum with

$$\int_T^\infty\int_T^\infty \frac{\log x\log y}{xy}\min\Big(Y,\frac{1}{|y-x|}\Big)\,dx\,dy$$

we can get a bound of $O\left(Y\log^2 T/T + \log^3 T/T\right)$. □

Therefore, for each $T \ge 1$, (2.5) gives a finite quasiperiodic approximation of $E(e^y,q,a)$ (where $y = \log x$), with error ε whose mean square is uniformly small according to Lemma 2.2. This is the key to the proof, which we now turn to, of the existence of the limiting distribution.

Let $f : \mathbb{R}^r \to \mathbb{R}$ be a fixed continuous function satisfying a Lipschitz estimate

$$|f(x) - f(y)| \le c_f|x-y|. \tag{2.7}$$

Consider

$$\frac{1}{Y}\int_{\log 2}^{Y} f(E(y))\,dy,$$

where $E(y) = (E(e^y,q,a_1),\,\dots,\,E(e^y,q,a_r))$. Let

$$E^{(T)}(y) = (E_1^{(T)}(y),\,\dots,\,E_r^{(T)}(y)),$$

with

$$E_j^{(T)}(y) = -c(q,a_j) - \sum_{\chi\ne\chi_0}\bar\chi(a_j)\sum_{|\gamma_\chi|\le T}\frac{e^{iy\gamma_\chi}}{\frac{1}{2}+i\gamma_\chi}.$$

Lemma 2.3. *For each T there is a probability measure ν_T on $\mathbb{R}^r$ such that*

$$\nu_T(f) := \int_{\mathbb{R}^r} f(x)\, d\nu_T(x) = \lim_{Y\to\infty} \frac{1}{Y}\int_{\log 2}^{Y} f(E^{(T)}(y))\, dy$$

for all bounded continuous functions f on $\mathbb{R}^r$. In addition, there is a constant $c = c(q)$ such that the support of ν_T lies in the ball $B(0, c\log^2 T)$.

Proof. This is a general feature of quasiperiodic functions. For later calculations we give the proof. Let $\chi \neq \chi_0$ and list the zeros $\frac{1}{2} + i\gamma_\chi$ of $L(s,\chi)$ such that $0 \le \gamma_\chi \le T$ as $\gamma_1, \dots, \gamma_N$. (We need only focus on $\gamma_\chi \ge 0$ since, for χ real, we have $L(\frac{1}{2} + i\gamma, \chi) = 0$ if and only if $L(\frac{1}{2} - i\gamma, \chi) = 0$, while for χ complex we have $L(\frac{1}{2}+i\gamma, \chi) = 0$ if and only if $L(\frac{1}{2} - i\gamma, \bar{\chi}) = 0$.) We may write $E^{(T)}(y)$ in the form

$$E^{(T)}(y) = 2\,\mathrm{Re}\left(\sum_{l=1}^{N} b_l e^{iy\gamma_l}\right) + b_0, \tag{2.8}$$

where $b_0, \dots, b_N \in \mathbb{C}^r$ with

$$b_0 = -(c(q, a_1), \dots, c(q, a_r)),$$
$$b_l = -\left(\frac{\bar{\chi}_l(a_1)}{\frac{1}{2} + i\gamma_l}, \dots, \frac{\bar{\chi}_l(a_r)}{\frac{1}{2} + i\gamma_l}\right).$$

Define the function $g(y_1, \dots, y_N)$ on the N-torus $T^N = \mathbb{R}^N/\mathbb{Z}^N$ by

$$g(y_1, \dots, y_N) = f\left(2\,\mathrm{Re}\sum_{l=1}^{N} b_l e^{2\pi i y_l} + b_0\right).$$

Clearly, g is continuous on T^N and

$$f(E^{(T)}(y)) = g\left(\frac{\gamma_1 y}{2\pi}, \dots, \frac{\gamma_N y}{2\pi}\right).$$

Let A be the topological closure in T^N of the one-parameter subgroup

$$\Gamma(y) := \{(\gamma_1 y/2\pi, \dots, \gamma_N y/2\pi) \mid y \in \mathbb{R}\}.$$

A is a torus and the Kronecker–Weyl Theorem asserts that $\Gamma(y)$ is equidistributed in A. Since $g|_A$ is continuous on A, we have

$$\lim_{Y\to\infty} \frac{1}{Y}\int_{\log 2}^{Y} f(E^{(T)}(y))\, dy = \int_A g(a)\, da \tag{2.9}$$

where da is the normalized Haar measure on A. This proves the first part of the lemma, with

$$\nu_T(f) = \int_A g(a)\, da.$$

(We could actually compute ν_T—as we do later—if $A = T^N$.)

The second part of the lemma follows by noting, from the definition of $E^{(T)}(y)$, that

$$|E_j^{(T)}(y)| \ll \sum_{|\gamma_\chi|\le T} \frac{1}{|\gamma_\chi| + 1},$$

which, by (2.4), is $\ll_q \log^2 T$. □

Returning to (2.7), we have, from (2.5):

$$\begin{aligned}\frac{1}{Y}\int_{\log 2}^{Y} f(E(y))\, dy &= \frac{1}{Y}\int_{\log 2}^{Y} f(E^{(T)}(y) + \varepsilon^{(T)}(y))\, dy \\ &= \frac{1}{Y}\int_{\log 2}^{Y} f(E^{(T)}(y))\, dy + O\left(\frac{c_f}{Y}\int_{\log 2}^{Y} |\varepsilon^{(T)}(y)|\, dy\right),\end{aligned}$$

where $\varepsilon^{(T)}(y) := E(y) - E^{(T)}(y)$ and the implied constant depends on q only. By Lemma 2.2, this is further equal to

$$\begin{aligned}&\frac{1}{Y}\int_{\log 2}^{Y} f(E^{(T)}(y))\, dy + O\left(\frac{c_f}{\sqrt{Y}}\left(\int_{\log 2}^{Y} |\varepsilon^{(T)}(y)|^2\, dy\right)^{\frac{1}{2}}\right) \\ &= \frac{1}{Y}\int_{\log 2}^{Y} f(E^{(T)}(y))\, dy + O\left(c_f\left(\frac{\log T}{\sqrt{T}} + \frac{\log^2 T}{Y\sqrt{T}}\right)\right).\end{aligned}$$

Letting $Y \to \infty$ and using Lemma 2.3 we conclude that

$$\nu_T(f) - O\Big(\frac{c_f \log T}{\sqrt{T}}\Big) \le \liminf \frac{1}{Y}\int_{\log 2}^{Y} f(E(y))\,dy \le \limsup \frac{1}{Y}\int_{\log 2}^{Y} f(E(y))\,dy \le \nu_T(f) + O\Big(\frac{c_f \log T}{\sqrt{T}}\Big). \quad (2.10)$$

Since T can be as large as we please, we conclude that the lim sup and the lim inf coincide, i.e., that

$$\mu(f) := \lim_{Y\to\infty} \frac{1}{Y}\int_{\log 2}^{Y} f(E(y))\,dy \quad (2.11)$$

exists. Thus there exists a Borel measure μ on $\mathbb{R}^r$ such that (2.11) holds for all f satisfying (2.7). Moreover, for such f's,

$$|\mu(f) - \nu_T(f)| \le c_q c_f \frac{\log T}{\sqrt{T}}.$$

From (2.11) it is also clear that since the ν_T's are probability measures (total mass 1), so is μ. In fact, in view of the second part of Lemma 2.3 and of (2.11), we have

$$\mu(B'_\lambda) = \nu_T(B'_\lambda) + O\Big(\frac{\log T}{\sqrt{T}}\Big) = O\Big(\frac{\log T}{\sqrt{T}}\Big)$$

for $\lambda = c\log^2 T$ (recall that B'_λ is the complement of the open ball of radius λ). In other words,

$$\mu(B'_\lambda) = O(\sqrt{\lambda}e^{-c\sqrt{\lambda}}) = O(e^{-c_2\sqrt{\lambda}}),$$

where c_2 depends only on q.

2.2. Lower Bounds

We will now present a proof of the lower bound for $\mu_{q;N,R}[\pm\lambda, \pm\infty)$ for large λ. The basic principle of this analysis is the same as that used in [Littlewood 1914]; see also [Ingham 1932, Ch. 5]. The proof that $\mu_{q;a_1,\dots,a_r}(B_R^{\pm}) > 0$, as in Theorem 1.2, is similar; see Remark 2.5.

Fix q. All the constants c_j below depend on q only. Let χ_1 be the real nonprincipal character mod q, and $L(s,\chi_1)$ its L-function. The nontrivial zeros of $L(s,\chi_1)$ are denoted simply by $\frac{1}{2} + i\gamma$. Set

$$R(x) := \frac{\log x}{\sqrt{x}} \sum_{p\le x} \chi_1(p) = \frac{\log x}{\sqrt{x}}(\pi_R(x,q) - \pi_N(x,q)).$$

As in Section 2.1, we have, for $X \ge 1$ and $x \ge 2$:

$$R(x) = -1 - \sum_{|\gamma|\le X} \frac{x^{i\gamma}}{\frac{1}{2}+i\gamma} + O\Big(\frac{\sqrt{x}\log^2(xX)}{X} + \frac{1}{\log x}\Big). \quad (2.12)$$

For $\varepsilon > 0$ with $\xi - \frac{1}{2}\varepsilon \ge \log 2$, set

$$F_\varepsilon(\xi) := \frac{1}{\varepsilon}\int_{\xi-\frac{\varepsilon}{2}}^{\xi+\frac{\varepsilon}{2}} R(e^y)\,dy.$$

Because of (2.4), $\sum_\gamma 1/(\frac{1}{4}+\gamma^2) < \infty$. Together with a simple computation, this yields

$$R(e^y) = -2\sum_{0\le\gamma\le X} \frac{\sin\gamma y}{\gamma} + O\Big(1 + \frac{e^{y/2}(y+\log X)^2}{X}\Big). \quad (2.13)$$

Integrating this from $\xi - \frac{1}{2}\varepsilon$ to $\xi + \frac{1}{2}\varepsilon$ and letting $X \to \infty$ we get

$$F_\varepsilon(\xi) = \frac{4}{\varepsilon}\sum_{0\le\gamma} \frac{\sin\gamma\xi\,\sin\frac{1}{2}\gamma\varepsilon}{\gamma^2} + O(1).$$

Next we let ε be very small and introduce

$$\tilde{F}_\varepsilon(\xi) := \frac{4}{\varepsilon}\sum_{0\le\gamma\le\varepsilon^{-2}} \frac{\sin\gamma\xi\,\sin\frac{1}{2}\gamma\varepsilon}{\gamma^2}.$$

Thus, if $\xi - \frac{1}{2}\varepsilon \ge \log 2$,

$$F_\varepsilon(\xi) = \tilde{F}_\varepsilon(\xi) + O(1). \quad (2.14)$$

By studying $\tilde{F}_\varepsilon(\xi)$ as a function of the real variable ξ (in particular near $\xi = 0$), and by exploiting its almost-periodicity, we will be able to prove a lower bound for $\mu_{q;N,R}[\lambda,\infty)$. Now,

$$\tilde F_\varepsilon(\tfrac12\varepsilon)=4\varepsilon\sum_{0\le\gamma\le\varepsilon^{-2}}\Big(\frac{\sin\frac12\gamma\varepsilon}{\varepsilon\gamma}\Big)^2\ge 4\varepsilon\sum_{0\le\gamma\le 1/\varepsilon}\frac{1}{3^2}\ge c_0\log\varepsilon^{-1}$$

with $c_0>0$. That is,

$$\tilde F_\varepsilon(\varepsilon)\ge c_1\log\varepsilon^{-1}\quad\text{with } c_1>0. \tag{2.15}$$

Let $\gamma_1,\dots,\gamma_N$ denote the imaginary parts of the zeros of $L(s,\chi_1)$ with $0\le\gamma\le\varepsilon^{-2}$. We have $N\sim c_3\varepsilon^{-2}\log\varepsilon$ as $\varepsilon\to0$. Consider, for $M\to\infty$ (with ε fixed and very small), the integers m with $(\log 2)/\varepsilon\le m\le M/\varepsilon$ and the values $\tilde F_\varepsilon(\varepsilon+m\varepsilon)$. We have

$$\begin{aligned}|\tilde F_\varepsilon(\varepsilon+m\varepsilon)-\tilde F_\varepsilon(\varepsilon)|&\le 2\sum_{0\le\gamma\le\varepsilon^{-2}}\left|\frac{\sin(m+1)\gamma\varepsilon-\sin\gamma\varepsilon}{\gamma}\right|\\&\le 2\Big(\max_{0\le\gamma\le\varepsilon^{-2}}\|\gamma m\varepsilon\|\Big)\sum_{0\le\gamma\le\varepsilon^{-2}}\frac{1}{|\gamma|}\\&\le\max_{0\le\gamma\le\varepsilon^{-2}}\|\gamma m\varepsilon\|c_2\log^2\varepsilon^{-1},\end{aligned}\tag{2.16}$$

where $\|\;\|$ denotes the distance to the nearest integer multiple of 2π. We want the right-hand side of this inequality to be appropriately small. If

$$\max_{0\le\gamma\le\varepsilon^{-2}}\|\gamma m\varepsilon\|\le\frac{c_1}{2c_2\log\varepsilon^{-1}},\tag{2.17}$$

it follows from (2.15) and (2.16) that

$$\tilde F_\varepsilon(\varepsilon+m\varepsilon)\ge\tfrac12 c_1\log\varepsilon^{-1}.$$

From this and (2.14) we have, on adjusting c_1 appropriately so as to incorporate the $O(1)$, that

$$F_\varepsilon(\varepsilon+m\varepsilon)\ge\tfrac12 c_1\log\varepsilon^{-1}.$$

Let G_M be the set of m such that $(\log 2)/\varepsilon\le m\le M/\varepsilon$ and (2.17) holds. To get a lower bound on $|G_M|$ as $M\to\infty$ we use the box principle. In $\mathbb{R}^N/\mathbb{Z}^N$ consider the vector $(\varepsilon\gamma_1/2\pi,\,\dots,\,\varepsilon\gamma_N/2\pi)$. Divide $\mathbb{R}^N/\mathbb{Z}^N$ into disjoint boxes of side lengths (essentially) $c_1/(4c_2\log\varepsilon^{-1})$. There will be effectively $(4c_2c_1^{-1}\log\varepsilon^{-1})^N$ such boxes. Of the vectors $m(\varepsilon\gamma_1/2\pi,\,\dots,\,\varepsilon\gamma_N/2\pi)$, with

$$\frac{\log 2}{\varepsilon}\le m\le\frac{M}{\varepsilon},$$

at least

$$\nu=\left[\frac{M-\log 2}{\varepsilon(4c_2c_1^{-1}\log\varepsilon^{-1})^N}\right]$$

will be in one box. Corresponding to these integers $m_1<m_2<\dots<m_\nu$, we form $n_j=m_j-m_1$, with $0\le n_j\le M/\varepsilon$; these numbers satisfy (2.17). It follows that

$$|G_M|\ge\left[\frac{M-\log 2}{\varepsilon\left(4c_2c_1^{-1}\log\varepsilon^{-1}\right)^N}\right].\tag{2.18}$$

Let

$$\beta_m=\int_{m\varepsilon+\frac12\varepsilon}^{m\varepsilon+\frac32\varepsilon}R^2(e^y)\,dy.$$

As in Lemma 2.2 (with T fixed, say $T=2$) we have

$$\sum_{\log 2\le m\varepsilon\le M}\beta_m\le\int_{\log 2}^{M+\frac32\varepsilon}R^2(e^y)dy\le c_4M.\tag{2.19}$$

Lemma 2.4. *For $m\in G_M$, the measure of the set of $y\in[(m+\frac12)\varepsilon,\,(m+\frac32)\varepsilon]$ such that $R(e^y)\ge\frac14c_1\log\varepsilon^{-1}$ is at least $\varepsilon^2c_1^2\log^2\varepsilon^{-1}/(16\beta_m)$.*

Proof. Let

$$\nu(\lambda)=\varepsilon^{-1}\big|\{t\in[(m+\tfrac12)\varepsilon,\,(m+\tfrac32)\varepsilon]\mid R(e^t)>\lambda\}\big|$$

be the distribution function of R on this interval. We have $\int_{-\infty}^\infty d\nu(\lambda)=1$,

$$\int_{-\infty}^\infty\lambda\,d\nu(\lambda)=\frac1\varepsilon\int_{m\varepsilon+\frac12\varepsilon}^{m\varepsilon+\frac32\varepsilon}R(e^y)\,dy=F_\varepsilon(m\varepsilon+\varepsilon)\ge\tfrac12c_1\log\varepsilon^{-1}\tag{2.20}$$

for $m\in G_M$, and $\int_{-\infty}^\infty\lambda^2\,d\nu(\lambda)=\beta_m\varepsilon^{-1}$.

Since the total mass is 1 we have

$$\int_{-\infty}^{\frac14c_1\log\varepsilon^{-1}}\lambda\,d\nu(\lambda)\le\tfrac14c_1\log\varepsilon^{-1},$$

so (2.20) gives

$$\int_{\frac{1}{4}c_1\log\varepsilon^{-1}}^{\infty} \lambda\, d\nu(\lambda) \geq \tfrac{1}{4}c_1 \log \varepsilon^{-1}.$$

Thus, by Cauchy–Schwartz's inequality,

$$\begin{aligned}\tfrac{1}{4}c_1\log\varepsilon^{-1} &\leq \Big(\int_{\frac{1}{4}c_1\log\varepsilon^{-1}}^{\infty}\lambda^2\, d\nu(\lambda)\Big)^{\frac{1}{2}}\Big(\int_{\frac{1}{4}c_1\log\varepsilon^{-1}}^{\infty} d\nu(\lambda)\Big)^{\frac{1}{2}} \\ &\leq \Big(\frac{\beta_m}{\varepsilon}\Big)^{\frac{1}{2}}\big(\nu[\tfrac{1}{4}c_1\log\varepsilon^{-1},\,\infty)\big)^{1/2}.\end{aligned}$$

Hence

$$\nu[\tfrac{1}{4}c_1\log\varepsilon^{-1},\,\infty) \geq \frac{\varepsilon c_1^2 \log^2 \varepsilon^{-1}}{16\beta_m},$$

proving the lemma. □

Continuing with the estimation of the lower bound, we have

$$\begin{aligned}&\big|\{y\in[\log 2,\, M+\tfrac{3}{2}\varepsilon]\mid R(e^y)\geq \tfrac{1}{4}c_1\log\varepsilon^{-1}\}\big| \\ &\geq \sum_{m\in G_M}\big|\{y\in[(m+\tfrac{1}{2})\varepsilon,\,(m+\tfrac{3}{2})\varepsilon]\mid R(e^y)\geq\tfrac{1}{4}c_1\log\varepsilon^{-1}\}\big|,\end{aligned}$$

which by Lemma 2.4 is bounded below by

$$\sum_{m\in G_M}\frac{\varepsilon^2 c_1^2\log^2\varepsilon^{-1}}{16\beta_m}. \tag{2.21}$$

Also,

$$\begin{aligned}|G_M| &= \sum_{m\in G_M} 1 = \sum_{m\in G_M}\sqrt{\beta_m}\frac{1}{\sqrt{\beta_m}} \\ &\leq \Big(\sum_{m\in G_M}\beta_m\Big)^{\frac{1}{2}}\Big(\sum_{m\in G_M}\frac{1}{\beta_m}\Big)^{\frac{1}{2}} \\ &\leq \Big(\sum_{\log 2/\varepsilon\leq m\leq M/\varepsilon}\beta_m\Big)^{\frac{1}{2}}\Big(\sum_{m\in G_M}\frac{1}{\beta_m}\Big)^{\frac{1}{2}} \\ &\leq \sqrt{c_4 M}\Big(\sum_{m\in G_M}\frac{1}{\beta_m}\Big)^{\frac{1}{2}}\end{aligned}$$

by (2.19). On the other hand, we have the lower bound (2.18) for $|G_M|$. Hence

$$\sum_{m\in G_M}\frac{1}{\beta_m} \geq \frac{1}{c_4 M}\frac{(M-\log 2)^2}{\varepsilon^2(4c_2c_1^{-1}\log\varepsilon^{-1})^{2N}}.$$

Combining this with the bound (2.21) gives

$$\begin{aligned}&\frac{M}{(M-\log 2)^2} \\ &\quad\times\big|\{y\in[\log 2,\, M+\tfrac{3}{2}\varepsilon]\mid R(e^y)\geq\tfrac{1}{4}c_1\log\varepsilon^{-1}\}\big| \\ &\qquad\geq \frac{c_1^2\log^2\varepsilon^{-1}}{16c_4(4c_2c_1^{-1}\log\varepsilon^{-1})^{2N}}.\end{aligned}$$

As $M\to\infty$, the left-hand side gives

$$\mu_{q;R,N}[\tfrac{1}{4}c_1\log\varepsilon^{-1},\,\infty).$$

So, if we choose $\lambda = \frac{1}{4}c_1\log\varepsilon^{-1}$, we get $N \sim c\varepsilon^{-2}\log\varepsilon^{-2} \leq \exp(A\lambda)$ for some A. That is, for suitable constants $A_1, A_2 > 0$ depending on q we have

$$\mu_{q;R,N}[\lambda,\infty) \geq \frac{A_1}{\exp(\exp(A_2\lambda))}.$$

Returning to (2.15), we note that

$$\tilde{F}_\varepsilon(-\varepsilon) \leq -c_1\log\varepsilon^{-1},$$

so one can repeat the whole argument to show that

$$\mu_{q;R,N}(-\infty,-\lambda] \geq \frac{A_1}{\exp(\exp(A_2\lambda))}.$$

This concludes the proof.

Remark 2.5. The reason we were able to obtain the lower bounds in this section is in part that $R(e^y)$ in (2.13) has essentially the form

$$R(e^y) = -2\sum_{0\leq\gamma\leq X} a_\gamma\frac{\sin\gamma y}{\gamma}$$

with $a_\gamma \equiv 1$. If one tries to apply the same method to show that $\pi(x,q,a)-\pi(x,q,b)$, for general $a,b\in A_q$, changes sign infinitely often, one runs into the difficulty that the coefficients are not positive (γ_χ now running over all χ mod q, $\chi\neq\chi_0$). In general, this problem appears formidable.

There are special a's and b's for which this can be overcome, and Theorem 1.2 falls into this category. The vector sum $E(y)$ in question is essentially of the form

$$-\sum_{\chi\neq\chi_0}\sum_{\gamma_\chi}\frac{\sin\gamma_\chi y}{\gamma_\chi}\left(\bar\chi(a_1),\ldots,\bar\chi(a_r)\right).$$

By an analysis similar to the one in this section we can force this vector (for many y's) to be a large multiple of

$$\left(\tfrac12(\varepsilon(a_1)+1)\varphi(q)-1,\ \ldots,\ \tfrac12(\varepsilon(a_r)+1)\varphi(q)-1\right).$$

The proof of Theorem 1.2 then follows along the lines above.

3. APPLICATIONS OF THE GRAND SIMPLICITY HYPOTHESIS

3.1. The Product Formula for $\hat\mu$

We turn to some consequences of assuming GSH, which says that the $\gamma\geq 0$ are linearly independent over the rationals. Suppose that χ^*, of conductor q_{χ^*} (which divides q), induces a character χ. Then $L(s,\chi^*)$ and $L(s,\chi)$ have the same zeros on $\operatorname{Re}(s)=\frac12$. It follows that GSH implies that $\{\gamma_\chi \mid \chi \bmod q\}$ is linearly independent over $\mathbb{Q}$. Hence the set of $y(\gamma_1,\ldots,\gamma_N)$ for $y\in\mathbb{R}$ is uniformly distributed in T^N, and (2.8), (2.9) and (2.11) imply that

$$\hat\mu_{q;a_1,\ldots,a_r}(\xi)=\lim_{N\to\infty}\exp\biggl(i\sum_{m=1}^r c(q,a_m)\xi_m\biggr)\prod_{j=1}^N\hat\mu_{\gamma_j}(\xi), \tag{3.1}$$

where μ_{γ_j} is the distribution of a typical term

$$-\biggl(\frac{\bar\chi(a_1)e^{i\gamma y}}{\frac12+i\gamma}+\frac{\chi(a_1)e^{-i\gamma y}}{\frac12-i\gamma},\ldots,\frac{\bar\chi(a_r)e^{i\gamma y}}{\frac12+i\gamma}+\frac{\chi(a_r)e^{-i\gamma y}}{\frac12-i\gamma}\biggr)$$

in (2.8). Writing $\bar\chi(a_j)=u_j+iv_j$, we get

$$-\frac{2}{\sqrt{\frac14+\gamma^2}}\bigl(u_1\sin(\gamma y+w_\gamma)+v_1\cos(\gamma y+w_\gamma),\ \ldots,\ u_r\sin(\gamma y+w_\gamma)+v_r\cos(\gamma y+w_\gamma)\bigr),$$

where $\cos w_\gamma/\sin w_\gamma=2\gamma$. Noting that $\sin(\gamma y)$ has density

$$\begin{cases}\dfrac{1}{\pi\sqrt{1-t^2}} & \text{if } -1<t<1,\\ 0 & \text{otherwise,}\end{cases}$$

we find that a typical $\hat\mu_\gamma(\xi)$ in (3.1) equals

$$\frac12\int_{-1}^1\exp\biggl(iR_\gamma\sum_{m=1}^r\xi_m(u_mt+v_m\sqrt{1-t^2})\biggr)\frac{dt}{\pi\sqrt{1-t^2}}+\frac12\int_{-1}^1\exp\biggl(iR_\gamma\sum_{m=1}^r\xi_m(u_mt-v_m\sqrt{1-t^2})\biggr)\frac{dt}{\pi\sqrt{1-t^2}},$$

where $R_\gamma=2/\sqrt{\frac14+\gamma^2}$. If we set $U=\sum_{m=1}^r\xi_mu_m$ and $V=\sum_{m=1}^r\xi_mv_m$, this becomes

$$\begin{aligned}\hat\mu_\gamma(\xi)&=\frac1\pi\int_{-1}^1\tfrac12\bigl(\exp(iR_\gamma(Ut+V\sqrt{1-t^2}))+\exp(iR_\gamma(Ut-V\sqrt{1-t^2}))\bigr)\frac{dt}{\sqrt{1-t^2}}\\&=\frac1\pi\int_{-1}^1\exp(iR_\gamma Ut)\cos(R_\gamma V\sqrt{1-t^2})\frac{dt}{\sqrt{1-t^2}}\\&=J_0(R_\gamma\sqrt{U^2+V^2}),\end{aligned}$$

where

$$J_0(z)=\sum_0^\infty\frac{(-1)^m(\frac12z)^{2m}}{(m!)^2} \tag{3.2}$$

is the Bessel function of the first kind. Hence, (3.1) becomes

$$\hat\mu_{q;a_1,\ldots,a_r}(\xi)=\exp\biggl(i\sum_{j=1}^rc(q,a_j)\xi_j\biggr)\times\prod_{\substack{\chi\neq\chi_0\\ \chi \bmod q}}\prod_{\gamma_\chi>0}J_0\biggl(\frac{2\left|\sum_{j=1}^r\chi(a_j)\xi_j\right|}{\sqrt{\frac14+\gamma_\chi^2}}\biggr). \tag{3.3}$$

Note that the factor $\exp\bigl(i\sum_{j=1}^rc(q,a_j)\xi_j\bigr)$ arises from the constant term in (2.8) and it accounts for the Chebyshev bias. Similarly, using (2.12), we have

$$\hat\mu_{q;R,N}(\xi)=e^{i\xi}\prod_{\gamma_{\chi_1}>0}J_0\biggl(\frac{2\xi}{\sqrt{\frac14+\gamma_{\chi_1}^2}}\biggr). \tag{3.4}$$

Also, as in (2.12), we have, for $X \geq 1$ and $x \geq 2$:

$$(\pi(x) - \mathrm{Li}(x))\frac{\log x}{\sqrt{x}} = -1 - \sum_{|\gamma| \leq X} \frac{x^{i\gamma}}{\frac{1}{2} + i\gamma} + O\Big(\frac{\sqrt{x}\log^2(xX)}{X} + \frac{1}{\log x}\Big),$$

where γ runs over the imaginary parts of the nontrivial zeros of $\zeta(s)$ that lie in the upper half-plane. Thus the formula for $\hat{\mu}_1$ is the same as that for $\hat{\mu}_{q;R,N}$, the only difference being in the set of γ's.

3.2. An Investigation of the Symmetries

We focus first on (3.4) (so χ_1 is primitive) and investigate its symmetries. Because J_0 is an even function, so is

$$\prod_{\gamma > 0} J_0\bigg(\frac{2\xi}{\sqrt{\frac{1}{4} + \gamma^2}}\bigg),$$

so (3.4) implies that the density function of $\mu_{q;R,N}$ is symmetric about $t = -1$. Therefore

$$\delta(P_{q;R,N}) = \int_0^\infty d\mu_{q;R,N}(t) < \tfrac{1}{2};$$

the inequality is strict because the density function of $\mu_{q;R,N}$ is entire and hence cannot be identically zero on $(-1, 0)$. See Remark 1.3. However, as $q \to \infty$, this bias towards nonresidues disappears, as is indicated in Theorem 1.5, which we now turn to.

Consider $\log \hat{\mu}_{q;R,N}(\xi/\sqrt{\log q})$. From (3.4) and (3.2) we see that for $|\xi| \leq A$, where A is any large fixed constant,

$$\log \hat{\mu}_{q;R,N}\Big(\frac{\xi}{\sqrt{\log q}}\Big) = \frac{i\xi}{\sqrt{\log q}} - \frac{\xi^2}{\log q}\sum_{\gamma>0}\frac{1}{\frac{1}{4}+\gamma^2} + O\bigg(\frac{A^4}{\log^2 q}\sum_{\gamma>0}\frac{1}{(\frac{1}{4}+\gamma^2)^2}\bigg). \tag{3.5}$$

Expression (4.14) below gives $\sum_{\gamma>0}(\frac{1}{4} + \gamma^2)^{-1}$ in terms of $L'/L(1, \chi_1)$. Under GRH, a simple adaptation of the argument in [Littlewood 1928, p. 927] shows that $L'/L(1, \chi) = O(\log \log q)$. Combining these results we have

$$\sum_{\gamma>0} \frac{1}{\frac{1}{4} + \gamma^2} = \tfrac{1}{2}\log q + O(\log\log q).$$

Moreover,

$$\sum_{\gamma>0} \frac{1}{(\frac{1}{4} + \gamma^2)^2} \leq 4 \sum_{\gamma>0} \frac{1}{\frac{1}{4} + \gamma^2},$$

so that (3.5) becomes

$$\log \hat{\mu}_{q;R,N}\Big(\frac{\xi}{\sqrt{\log q}}\Big) = -\tfrac{1}{2}\xi^2 + O\Big(\frac{A}{\sqrt{\log q}} + \frac{A^2 \log\log q}{\log q} + \frac{A^4}{\log q}\Big).$$

In other words, we have shown that, for $|\xi| \leq A$,

$$\hat{\mu}_{q;R,N}\Big(\frac{\xi}{\sqrt{\log q}}\Big)$$

approaches $e^{-\xi^2/2}$ uniformly.

Hence by Levy's Theorem [Lévy 1922], the measures $\tilde{\mu}_{q;N,R}$ (as in Theorem 1.6) converge in measure to the standard Gaussian. As a corollary we deduce that $\delta(P_{q;N,R}) = \tilde{\mu}_{q;N,R}[0, \infty)$ satisfies

$$\delta(P_{q;N,R}) \to \tfrac{1}{2} \quad \text{as } q \to \infty. \tag{3.6}$$

We turn to the proof of Theorem 1.5, which runs along similar lines. Let q be large (where q now is any integer) and let $a_1, \ldots, a_r$, with r fixed, be distinct elements of A_q. Let $\tilde{\mu}_{q;a_1,\ldots,a_r}$ be the measure on $\mathbb{R}^r$ whose Fourier transform is

$$\hat{\mu}_{q;a_1,\ldots,a_r}\bigg(\frac{\xi}{\sqrt{\varphi(q)\log q}}\bigg).$$

The claim is that $\tilde{\mu}_{q;a_1,\ldots,a_r}$ converges in measure to the Gaussian

$$\frac{e^{-(x_1^2+\cdots+x_r^2)}}{(2\pi)^{r/2}}\, dx_1 \ldots dx_r$$

as $q \to \infty$, independently of the choice of $a_1, \ldots, a_r$. As before, this follows from Levy's criterion. Fix

A and consider $\xi \in \mathbb{R}^r$ with $|\xi| \le A$. Then, using (3.3) and (3.2), we have

$$\begin{aligned}
&\log \hat{\hat{\mu}}_{q;a_1,\dots,a_r}(\xi) \\
&\quad = \log \hat{\mu}_{q;a_1,\dots,a_r}\Big(\frac{\xi}{\sqrt{\varphi(q)\log q}}\Big) \\
&\quad = \frac{i}{\sqrt{\varphi(q)\log q}} \sum_{j=1}^{r} c(q,a_j)\xi_j \\
&\qquad\qquad + \sum_{\substack{\chi\neq\chi_0 \\ \chi \bmod q}} \sum_{\gamma_\chi>0} \log J_0\Big(\frac{2\big|\sum_{j=1}^r \chi(a_j)\xi_j\big|}{\sqrt{\varphi(q)\log q(\frac14+\gamma_\chi^2)}}\Big) \\
&\quad = -\sum_{\substack{\chi\neq\chi_0 \\ \chi \bmod q}} \sum_{\gamma_\chi>0} \frac{\big|\sum_{j=1}^r \chi(a_j)\xi_j\big|^2}{\varphi(q)\log q(\frac14+\gamma_\chi^2)} \\
&\qquad + O\Big(\frac{d(q)A}{\sqrt{\varphi(q)\log q}} + \frac{A^4}{(\varphi(q)\log q)^2} \sum_{\chi\neq\chi_0} \frac{1}{(\frac14+\gamma_\chi^2)^2}\Big),
\end{aligned}$$

where $d(q) = \sum_{d\mid q} 1$ and we have used $c(q,a) < d(q)$. As in (1.4), we let χ^* denote the primitive character inducing χ. Its conductor q_{χ^*} divides q.

Now $L(s,\chi)$ and $L(s,\chi^*)$ have the same zeros on $\mathrm{Re}(s) = \frac12$, so

$$\sum_{\gamma_\chi} \frac{1}{\frac14+\gamma_\chi^2} = \sum_{\gamma_{\chi^*}} \frac{1}{\frac14+\gamma_{\chi^*}^2} = -2\,\mathrm{Re}\,B(\chi),$$

where $B(\chi)$ is given in (1.4). For a proof see [Davenport 1980]. As before, Littlewood's bound implies that

$$B(\chi^*) = -\tfrac12(\log q_{\chi^*}) + O(\log\log q).$$

Hence

$$\begin{aligned}
\log \hat{\hat{\mu}}_{q;a_1,\dots,a_r}(\xi) &= -\frac{1}{2\varphi(q)\log q} \\
&\quad \times \sum_{\chi\neq\chi_0} \log q_{\chi^*} \Big|\sum_{j=1}^r \chi(a_j)\xi_j\Big|^2 + O\Big(A^4 \frac{\log\log q}{\log q}\Big),
\end{aligned}$$

where we have also used $d(q) = O_\varepsilon(q^\varepsilon)$ for all $\varepsilon > 0$. In order to analyze the first term on the right-hand side above, let $\beta(a)$ denote the number of primitive characters to a modulus a. For each a dividing q, every such character induces a unique character mod q, so $\sum_{a\mid q}\beta(a) = \varphi(q)$. Also note that $\beta(a) \le \varphi(a) < a$. Now

$$\begin{aligned}
\sum_{\chi\neq\chi_0} \log q_{\chi^*} \Big|\sum_{j=1}^r \chi(a_j)\xi_j\Big|^2
&= \sum_{j,k} \xi_j\xi_k \sum_{\chi\neq\chi_0} \chi\Big(\frac{a_j}{a_k}\Big) \log q_{\chi^*} \\
&= \sum_{j,k} \xi_j\xi_k \sum_{\chi\neq\chi_0} \chi\Big(\frac{a_j}{a_k}\Big) \log q \\
&\qquad - \sum_{j,k} \xi_j\xi_k \sum_{\chi\neq\chi_0} \chi\Big(\frac{a_j}{a_k}\Big) \log q/q_{\chi^*}.
\end{aligned}$$

Denote the first summand on the right-hand side by I, and the second (including the minus sign) by II. Clearly

$$\mathrm{I} = (\varphi(q)-1)\log q \sum_{j=1}^r \xi_j^2 + O((\log q)A^2).$$

On the other hand,

$$\mathrm{II} = -\sum_{j,k} \xi_j\xi_k \sum_{\substack{a\mid q \\ a\neq 1}} \sideset{}{^*}\sum_{\lambda \bmod a} \lambda\Big(\frac{a_j}{a_k}\Big) \log q/a,$$

where $\sum^*$ indicates the sum over primitive characters mod a. So

$$\mathrm{II} \ll A^2 \sum_{a\mid q} \beta(a) \log \frac{q}{a}.$$

For any $\alpha < 1$,

$$\begin{aligned}
\mathrm{II} &\ll A^2 \sum_{\substack{a\mid q \\ a\le q^\alpha}} \beta(a)\log q + A^2 \sum_{\substack{a\mid q \\ a> q^\alpha}} \beta(a)\log q^{1-\alpha} \\
&\le A^2(\log q) d(q) q^\alpha + A^2(1-\alpha)(\log q)\varphi(q).
\end{aligned}$$

Hence

$$\limsup_{q\to\infty} \frac{|\mathrm{II}|}{\varphi(q)\log q} \le A^2(1-\alpha).$$

Since $1-\alpha$ can be chosen as small as we please we get

$$\lim_{q\to\infty} \frac{|\mathrm{II}|}{\varphi(q)\log q} = 0,$$

and moreover convergence is uniform for $|\xi| \le A$. Thus

$$\mathrm{I} + \mathrm{II} \sim \varphi(q) \log q \sum_{j=1}^{r} \xi_j^2.$$

We conclude that, for $|\xi| \le A$,

$$\hat{\tilde{\mu}}_{q;a_1,\dots,a_r}(\xi) \to \exp\biggl(-\sum_{j=1}^{r} \tfrac{1}{2}\xi_j^2\biggr).$$

This proves the central limit theorem for $\tilde{\mu}_{q;a_1,\dots,a_r}$. It follows that, for any $D \subset \mathbb{R}^r$ and for any permutation σ of the r-coordinates,

$$|\tilde{\mu}_{q;a_1,\dots,a_r}(D) - \tilde{\mu}_{q;a_1,\dots,a_r}(D^\sigma)| \to 0$$

as $q \to \infty$. That is, $\tilde{\mu}$ becomes unbiased, and, in particular,

$$\delta(P_{q;a_1,\dots,a_r}) = \tilde{\mu}_{q;a_1,\dots,a_r}(\{x \mid x_1 > x_2 > \cdots > x_r\})$$

approaches $1/r!$ as $q \to \infty$.

Next we study the symmetries of $\mu_{q;a_1,\dots,a_r}$.

Proposition 3.1. *The density function of $\mu_{q;a_1,\dots,a_r}$ is symmetric in $(x_1, \dots, x_r)$ if and only if either*

(a) $r = 2$ *and* $c(q, a_1) = c(q, a_2)$, *or*
(b) $r = 3$ *and there exists* $\rho \ne 1$ *satisfying these congruences modulo* q:

$$\rho^3 \equiv 1, \quad a_2 \equiv a_1\rho, \quad \textit{and} \quad a_3 \equiv a_1\rho^2.$$

The factor $\exp\bigl(i\sum_{j=1}^{r} c(q, a_j)\xi_j\bigr)$ in (3.3) shifts the mean of μ to $-(c(q, a_1), \dots, c(q, a_r))$ (note that the product of Bessel functions in (3.3) is an even function). Hence, if μ is symmetric, $c(q, a_j) = c(q, a_l)$ for all $1 \le j, l \le r$. We assume that this is the case and thus the symmetry issue is thrown onto the infinite product of Bessel functions.

Lemma 3.2. $B_\chi(\xi_1, \dots, \xi_r) := \bigl|\sum_{j=1}^{r} \chi(a_j)\xi_j\bigr|$ *is symmetric in* $(\xi_1, \dots, \xi_r)$ *for all* χ *if and only if one of the two conditions in Proposition* 3.1 *obtains.*

Proof. If $r = 2$, $B_\chi(\xi_1, \xi_2) = |\chi(a_1)\xi_1 + \chi(a_2)\xi_2|$. Now, $|\chi(a_1)| = |\chi(a_2)| = 1$, so

$$\begin{aligned}|\chi(a_1)\xi_1 + \chi(a_2)\xi_2| &= \bigl|\overline{\chi(a_1)\chi(a_2)}(\chi(a_1)\xi_1 + \chi(a_2)\xi_2)\bigr| \\ &= \bigl|\overline{\chi(a_2)}\xi_1 + \overline{\chi(a_1)}\xi_2\bigr| \\ &= \bigl|\overline{\overline{\chi(a_2)}\xi_1 + \overline{\chi(a_1)}\xi_2}\bigr| = B_\chi(\xi_2, \xi_1).\end{aligned}$$

If $r = 3$ and there exists ρ as stated, we have $\chi(a_2) = \chi(a_1)\chi(\rho)$, $\chi(a_3) = \chi(a_1)\chi^2(\rho)$, $\chi^3(\rho) = 1$. Hence, $\bigl|\chi(a_1)\xi_1 + \chi(a_1)\chi(\rho)\xi_2 + \chi(a_1)\chi^2(\rho)\xi_3\bigr| = \bigl|\xi_1 + \tau\xi_2 + \tau^2\xi_3\bigr|$, where $\tau^3 = 1$. But

$$\begin{aligned}\bigl|\xi_1 + \tau\xi_2 + \tau^2\xi_3\bigr| &= \bigl|\tau(\xi_1 + \tau\xi_2 + \tau^2\xi_3)\bigr| \\ &= \bigl|\xi_3 + \tau\xi_1 + \tau^2\xi_2\bigr| \\ &= \bigl|\xi_2 + \tau\xi_3 + \tau^2\xi_1\bigr|.\end{aligned}$$

Furthermore,

$$\begin{aligned}\bigl|\xi_1 + \tau\xi_2 + \tau^2\xi_3\bigr| &= \bigl|\overline{\xi_1 + \tau\xi_2 + \tau^2\xi_3}\bigr| \\ &= \bigl|\xi_1 + \tau\xi_3 + \tau^2\xi_2\bigr|.\end{aligned}$$

These equalities imply that $B_\chi(\xi_1, \xi_2, \xi_3)$ is symmetric in (ξ_1, ξ_2, ξ_3).

Conversely, if $|\chi(a_1)\xi_1 + \chi(a_2)\xi_2 + \chi(a_3)\xi_3|$ is symmetric in (ξ_1, ξ_2, ξ_3) for all χ, then so is

$$|\xi_1 + \chi(a_2/a_1)\xi_2 + \chi(a_3/a_1)\xi_3|.$$

Hence $\operatorname{Re}\chi(a_2/a_1) = \operatorname{Re}\chi(a_3/a_1)$, and similarly $\operatorname{Re}\chi(a_1/a_2) = \operatorname{Re}\chi(a_3/a_2)$. From this we deduce $\chi(a_2/a_1) = w$ and $\chi(a_3/a_1) = w^2$, with $w^3 = 1$. This being so for all χ, there exists $\rho \ne 1$ such that $a_2 \equiv a_1\rho \bmod q$, $a_3 \equiv a_1\rho^2 \bmod q$, and $\rho^3 \equiv 1 \bmod q$.

The same argument shows that $B_\chi(\xi_1, \dots, \xi_r)$ cannot be symmetric if $r \ge 4$. For if it were, then any three of the a_i's would be related as above, leading to a contradiction of the fact that the a_i's are distinct. □

We can now prove Proposition 3.1. If $r = 2$ and $c(q, a_1) = c(q, a_2)$, then since $B_\chi(\xi_1, \xi_2)$ is symmetric so is $\hat{\mu}(\xi_1, \xi_2)$ and also μ. If $r = 3$ and $a_2 \equiv a_1\rho \bmod q$, $a_3 \equiv a_1\rho^2 \bmod q$, then $c(q, a_1) = c(q, a_2) = c(q, a_3)$, so the exponential factor in $\hat{\mu}$ is symmetric in (ξ_1, ξ_2, ξ_3) and by Lemma 3.2 so is

$B_\chi(\xi_1,\xi_2,\xi_3)$. This shows that $\hat{\mu}$ is symmetric, and therefore also μ.

Conversely, if $r \geq 4$ or if condition (b) of Proposition 3.1 fails, then

$$B_\chi(\xi_1,\xi_2,\xi_3,\ldots,\xi_r) \neq B_\chi^\sigma(\xi_1,\xi_2,\xi_3,\ldots,\xi_r)$$

for some permutation σ. Assume that

$$\exp\biggl(i\sum_{j=1}^{r} c(q,a_j)\xi_j\biggr) \prod_{\substack{\chi\neq\chi_0\\ \chi \bmod q}} \prod_{\gamma_\chi>0} J_0\biggl(\frac{2B_\chi(\xi)}{\sqrt{\frac14+\gamma_\chi^2}}\biggr)$$
$$\equiv \exp\biggl(i\sum_{j=1}^{r} c(q,a_j)\xi_{\sigma(j)}\biggr) \prod_{\substack{\chi\neq\chi_0\\ \chi \bmod q}} \prod_{\gamma_\chi>0} J_0\biggl(\frac{2B_\chi^\sigma(\xi)}{\sqrt{\frac14+\gamma_\chi^2}}\biggr).$$

First, any χ for which $B_\chi(\xi) \equiv B_\chi^\sigma(\xi)$ can be removed on both sides of this identity without altering the relation. So we may assume that the above product over χ contains only terms such that $B_\chi(\xi) \not\equiv B_\chi^\sigma(\xi)$. In view of our assumption, the product is nonempty. Now choose ξ generically so that:

(i) $B_\chi(\xi) \neq 0$ and $B_\chi^\sigma(\xi) \neq 0$, for all χ mod q;
(ii) if $B_\chi(\xi)/B_\lambda^\sigma(\xi) \neq 1$, then

$$\frac{B_\chi(\xi)}{B_\lambda^\sigma(\xi)} \neq \sqrt{\frac{\frac14+\gamma_\chi^2}{\frac14+\gamma_\lambda^2}}$$

for all χ, λ mod q.

This can be done because our set of γ_χ's is countable. From (3.6) we have that, for ξ fixed as above and all $t \in \mathbb{R}$,

$$\exp\biggl(it\sum_{j=1}^{r} c(q,a_j)\xi_j\biggr) \prod_{\chi} \prod_{\gamma_\chi>0} J_0\biggl(\frac{2tB_\chi(\xi)}{\sqrt{\frac14+\gamma_\chi^2}}\biggr)$$
$$\equiv \exp\biggl(it\sum_{j=1}^{r} c(q,a_j)\xi_{\sigma(j)}\biggr) \prod_{\chi} \prod_{\gamma_\chi>0} J_0\biggl(\frac{2tB_\chi^\sigma(\xi)}{\sqrt{\frac14+\gamma_\chi^2}}\biggr).$$

The smallest zero in t of the left-hand side occurs at a number of the form

$$\frac{w\sqrt{\frac14+\gamma_\chi^2}}{2B_\chi(\xi)},$$

where w is the smallest zero of $J_0(z)$. The smallest zero on the right-hand side is at some

$$\frac{w\sqrt{\frac14+\gamma_\lambda^2}}{2B_\lambda^\sigma(\xi)}.$$

So we must have

$$\frac{w\sqrt{\frac14+\gamma_\chi^2}}{2B_\chi(\xi)} = \frac{w\sqrt{\frac14+\gamma_\lambda^2}}{2B_\lambda^\sigma(\xi)}.$$

In view of (ii) above, this implies

$$\frac{B_\chi(\xi)}{B_\lambda^\sigma(\xi)} = 1 = \frac{\sqrt{\frac14+\gamma_\chi^2}}{\sqrt{\frac14+\gamma_\lambda^2}}.$$

But the γ's are distinct, since we are assuming GSH, so $\chi = \lambda$. We conclude that $B_\chi(\xi) = B_\chi^\sigma(\xi)$, which contradicts an earlier condition.

4. NUMERICAL INVESTIGATIONS

We now describe the computations that led to the following numbers and the graphs at the end of this section.

$$\begin{aligned}
\delta(P_1^{\text{comp}}) &= 0.99999973\ldots\\
\delta(P_{3;N;R}) &= 0.9990\ldots\\
\delta(P_{4;N;R}) &= 0.9959\ldots\\
\delta(P_{5;N;R}) &= 0.9954\ldots\\
\delta(P_{7;N;R}) &= 0.9782\ldots\\
\delta(P_{11;N;R}) &= 0.9167\ldots\\
\delta(P_{13;N;R}) &= 0.9443\ldots
\end{aligned}$$

Let $f_{q;N,R}(t)$ and $f_1(t)$ be the density functions of $\mu_{q;R,N}$ and μ_1 respectively. In what follows, it will be more convenient to work with the distribution ω whose density function g is

$$g(t) := f(t-1),$$

where f stands for either $f_{q;R,N}$ or f_1. Its Fourier transform is

$$\hat{\omega}(\xi) = \prod_{\gamma>0} J_0\biggl(\frac{2\xi}{\sqrt{\frac14+\gamma^2}}\biggr) \tag{4.1}$$

and is symmetric about $t=0$ rather than $t=-1$.

We are interested in evaluating

$$\delta(P_{q;N,R}) = \int_{-\infty}^{1} d\omega_{q;R,N}(t),$$
$$\delta(P_1^{\mathrm{comp}}) = \int_{-\infty}^{1} d\omega_1(t).$$

Now, because $g_{q;R,N}$ is symmetric about 0, we have

$$\begin{aligned}\delta(P_{q;N,R}) &= \frac{1}{2}\left(\int_{-\infty}^{1} + \int_{-1}^{\infty}\right) d\omega_{q;R,N}(t)\\ &= \frac{1}{2} + \frac{1}{2}\int_{-1}^{1} d\omega_{q;R,N}(t)\\ &= \frac{1}{2} + \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\sin u}{u}\hat{\omega}_{q;R,N}(u)\, du, \quad (4.2)\end{aligned}$$

where the last equality follows from the inversion formula of characteristic functions; it was this expression that was used to compute the δ's. Similar equations hold for $\delta(P_1^{\mathrm{comp}})$.

The evaluation of these integrals involves three approximations. First, the integral was replaced by a sum of appropriately small rectangles. Then the infinite domain of summation was replaced by a large but finite domain. Finally, in place of the infinite product for $\hat{\omega}$, a finite product and a compensating polynomial were used. We now detail these three steps and estimate their cost to (4.2).

4.1. Replacing the Integral with a Sum

Consider the Poisson summation formula

$$\varepsilon \sum_{n\in\mathbb{Z}} \varphi(\varepsilon n) = \sum_{n\in\mathbb{Z}} \hat{\varphi}\big(\tfrac{n}{\varepsilon}\big) = \hat{\varphi}(0) + \sum_{\substack{n\in\mathbb{Z}\\ n\neq 0}} \hat{\varphi}\big(\tfrac{n}{\varepsilon}\big), \quad (4.3)$$

applied to

$$\begin{aligned}\varphi(u) &= \frac{1}{2\pi}\frac{\sin u}{u}\hat{\omega}(u),\\ \hat{\varphi}(x) &= \tfrac{1}{2}(\chi_{[-1,1]} * g)(x)\\ &= \frac{1}{2}\int_{x-1}^{x+1} g(u)\, du = \frac{1}{2}\int_{x-1}^{x+1} d\omega(u). \quad (4.4)\end{aligned}$$

We can justify using Poisson summation here as follows. As is well known [Watson 1948, p. 207],

$$|J_0(x)| \leq \min\big(1, \sqrt{2/(\pi\,|x|)}\big), \quad (4.5)$$

from which we deduce that $\hat{\mu}'(\xi)$ is rapidly decreasing. Furthermore, $g(u)$ is also rapidly decreasing, as we see from (1.1). Therefore, φ and $\hat{\varphi}$ are rapidly decreasing. Finally, $\hat{\omega}$ is continuous everywhere since the product in (4.1) converges absolutely for all ξ. Hence φ is also continuous. These facts allow us to apply Poisson summation [Stein and Weiss 1971].

Returning to (4.3), we have

$$\begin{aligned}&\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\sin u}{u}\hat{\omega}(u)\, du\\ &\qquad = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} \varepsilon \frac{\sin \varepsilon n}{\varepsilon n}\hat{\omega}(\varepsilon n) - \sum_{\substack{n\in\mathbb{Z}\\ n\neq 0}} \hat{\varphi}\big(\tfrac{n}{\varepsilon}\big) \quad (4.6)\end{aligned}$$

Therefore, to estimate the error of replacing the integral in (4.2) with the sum in (4.6), we need to get a bound on $\hat{\varphi}(n/\varepsilon)$. This amounts to bounding ω.

Montgomery [1980] shows that

$$\omega\Bigg[2\sum_{0<\gamma\leq X} R_\gamma, \infty\Bigg) \leq \exp\Bigg(-\frac{3}{4}\frac{\big(\sum_{0<\gamma\leq X} R_\gamma\big)^2}{\sum_{\gamma>X} R_\gamma^2}\Bigg) \quad (4.7)$$

with $R_\gamma = 2/\sqrt{\frac{1}{4}+\gamma^2}$, provided that the sums in this equation are nonempty.

It is possible to use this bound, together with (2.4), to get a double exponential bound on ω. However, to obtain a bound with explicit constants requires using explicit constants in the error term in (2.4). But we can avoid this by using the fact that, for $\zeta(s)$—which, for convenience, we call the $q = 1$ case—and for $L(s,\chi_1)$ with $q = 3,4,5,7,11,13$, all the positive γ's are greater than 2; this we know by looking at our computer files of zeros. Hence, for any $\lambda \geq 0$, with $q = 1,3,4,5,7,11,13$, we may find an X such that

$$0 \leq \lambda - 2\sum_{0<\gamma\leq X} R_\gamma < 2.$$

Combining this with (4.7) yields for $q = 1, 3, 4, 5, 7, 11, 13$ and for $\lambda \geq 2$ (so that the sum is nonempty):

$$\omega[\lambda, \infty) < \exp\biggl(-\frac{3}{4}\frac{\left(\frac{1}{2}(\lambda-2)\right)^2}{\sum_{\gamma>X} R_\gamma^2}\biggr) \leq \exp\biggl(-\frac{3}{4}\frac{\left(\frac{1}{2}(\lambda-2)\right)^2}{\sum_{\gamma>0} R_\gamma^2}\biggr).$$

Looking ahead to Table 2 and (4.13)–(4.14), we see that

$$\biggl(\sum_{\gamma>0} R_\gamma^2\biggr)^{-1} > 0.98$$

in all instances, so that

$$\omega[\lambda, \infty) \leq \exp(-\tfrac{1}{6}(\lambda-2)^2)$$

for $q = 1, 3, 4, 5, 7, 11, 13$ and $\lambda \geq 2$, where we used the fact that $\frac{1}{6} < \frac{3}{16} - 0.98^{-1}$. Hence, for $n \geq 1$ with $\frac{n}{\varepsilon} - 1 \geq 2$, (4.4) gives

$$\hat{\varphi}\Bigl(\frac{n}{\varepsilon}\Bigr) = \frac{1}{2}\int_{\frac{n}{\varepsilon}-1}^{\frac{n}{\varepsilon}+1} g(u)\,du \leq \tfrac{1}{2}\omega\bigl[\tfrac{n}{\varepsilon}-1,\, \infty\bigr) \leq \tfrac{1}{2}\exp\bigl(-\tfrac{1}{6}\bigl(\tfrac{n}{\varepsilon}-3\bigr)^2\bigr).$$

Now, because $g(u)$ is symmetric about 0, so is $\hat{\varphi}$. Thus, choosing $\varepsilon = \frac{1}{20}$, we find

$$\sum_{\substack{n\in\mathbb{Z}\\ n\neq 0}} \hat{\varphi}\bigl(\tfrac{n}{\varepsilon}\bigr) = 2\sum_{1}^{\infty} \hat{\varphi}\bigl(\tfrac{n}{\varepsilon}\bigr) \leq \sum_{1}^{\infty} \exp(-\tfrac{1}{6}(20n-3)^2) < 2\exp\left(-\tfrac{1}{6}(17)^2\right) = 10^{-20.617\ldots}.$$

Combining this with (4.2) and (4.6) yields

$$\delta(P_{q;N,R}) = \tfrac{1}{2} + \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} \varepsilon\frac{\sin\varepsilon n}{\varepsilon n}\hat{\omega}_{q;R,N}(\varepsilon n) + \text{error}, \tag{4.8}$$

where $\varepsilon = \frac{1}{20}$ and $|\text{error}| < 10^{-20}$. The same holds for $\delta(P_1^{\text{comp}})$.

4.2. The Cutoff

Next, in (4.8), we replaced the sum over $-\infty < n\varepsilon < \infty$ with a sum over $-C \leq n\varepsilon \leq C$, where C was chosen sufficiently large so that the tail ends of the sum contributed a pleasingly small amount. More precisely,

$$\frac{1}{2\pi}\biggl(\sum_{-\infty<n\varepsilon<\infty} - \sum_{-C\leq n\varepsilon\leq C}\biggr)\varepsilon\frac{\sin n\varepsilon}{n\varepsilon}\hat{\omega}(n\varepsilon) = \frac{1}{\pi}\sum_{n\varepsilon>C}\varepsilon\frac{\sin n\varepsilon}{n\varepsilon}\prod_{j=1}^{\infty} J_0\Bigl(\frac{2n\varepsilon}{\sqrt{\frac{1}{4}+\gamma_j^2}}\Bigr) < \frac{1}{\pi}\sum_{n\varepsilon>C}\varepsilon\frac{1}{n\varepsilon}\biggl|\prod_{j=1}^{M} J_0\Bigl(\frac{2n\varepsilon}{\sqrt{\frac{1}{4}+\gamma_j^2}}\Bigr)\biggr|$$

for $M = 1, 2, 3, \ldots$. By (4.5), the right-hand side above is dominated by

$$\frac{\prod_{j=1}^{M}(\frac{1}{4}+\gamma_j^2)^{1/4}}{\pi^{M/2+1}}\sum_{n\varepsilon>C}\frac{\varepsilon}{(n\varepsilon)^{M/2+1}} < \frac{\prod_{j=1}^{M}(\frac{1}{4}+\gamma_j^2)^{1/4}}{\pi^{M/2+1}}\biggl(\int_C^\infty \frac{1}{x^{M/2+1}}dx + \frac{\varepsilon}{C^{M/2+1}}\biggr) = \frac{\prod_{j=1}^{M}(\frac{1}{4}+\gamma_j^2)^{1/4}}{\pi^{M/2+1}}\biggl(\frac{2}{MC^{M/2}} + \frac{1}{20C^{M/2+1}}\biggr). \tag{4.9}$$

Using this bound and our computer files of the γ_j's, we found that the C's and M's listed in Table 1 gave us a small enough error to achieve eight-digit accuracy for $\delta(P_1^{\text{comp}})$ and four-digit accuracy for the other δ's. And so we have

$$\delta(P_{q;N,R}) = \frac{1}{2\pi}\sum_{-25\leq n\varepsilon\leq 25}\varepsilon\frac{\sin\varepsilon n}{\varepsilon n}\prod_{\gamma_{\chi_1}>0} J_0\biggl(\frac{2n\varepsilon}{\sqrt{\frac{1}{4}+\gamma_{\chi_1}^2}}\biggr) + \tfrac{1}{2} + \text{error} \tag{4.10}$$

	C	M	bound
$\delta(P_1^{\text{comp}})$	50	59	2×10^{-10}
$\delta(P_{3;N;R})$	25	36	4×10^{-7}
$\delta(P_{4;N;R})$	25	39	6×10^{-8}
$\delta(P_{5;N;R})$	25	42	2×10^{-8}
$\delta(P_{7;N;R})$	25	46	2×10^{-9}
$\delta(P_{11;N;R})$	25	52	1×10^{-10}
$\delta(P_{13;N;R})$	25	53	5×10^{-11}

TABLE 1. Error bounds for the computed values of $\delta(P_1^{\text{comp}})$ and $\delta(P_{q;N;R})$. The bounds are provided by (4.9); C and M are chosen accordingly.

and

$$\delta(P_1^{\mathrm{comp}}) = \frac{1}{2\pi} \sum_{-50 \le n\varepsilon \le 50} \varepsilon \frac{\sin \varepsilon n}{\varepsilon n} \prod_{\gamma_\zeta > 0} J_0\Big(\frac{2n\varepsilon}{\sqrt{\frac{1}{4} + \gamma_\zeta^2}}\Big) + \tfrac{1}{2} + \text{error}, \quad (4.11)$$

where the error includes the one shown in Table 1 and the one from (4.8).

4.3. Replacing the Infinite Product

Finally, we replaced the infinite product in (4.10)–(4.11) with a finite product and a polynomial that compensated for the missing tail end of the product:

$$\hat{\omega}(u) \approx p(u) \prod_{0<\gamma\le X} J_0\Big(\frac{2u}{\sqrt{\frac{1}{4} + \gamma^2}}\Big) \quad (4.12)$$

for $-C \le u \le C$, where $p(u) = \sum_{m=0}^{A} b_m u^{2m}$ approximates the product

$$\prod_{\gamma > X} J_0\Big(\frac{2u}{\sqrt{\frac{1}{4} + \gamma^2}}\Big) = \sum_{m=0}^{\infty} b_m u^{2m}.$$

Using the formula (3.2) for $J_0(z)$ and the fact—a consequence of (2.4)—that

$$\sum_{\gamma > X} \frac{1}{\frac{1}{4} + \gamma^2}$$

converges, say to $T_1 = T_1(X)$, we see that such an expansion is justified. In fact, comparing the b_m's with the coefficients of

$$\prod_{\gamma > X} \exp\bigg(\frac{1}{4}\Big(\frac{2u}{\sqrt{\frac{1}{4} + \gamma^2}}\Big)^2\bigg) = \exp\bigg(u^2 \sum_{\gamma > X} \frac{1}{\frac{1}{4} + \gamma^2}\bigg),$$

we find that $|b_m| < T_1^m/m!$. Therefore

$$\bigg|\sum_{m=A+1}^{\infty} b_m u^{2m}\bigg| < \sum_{A+1}^{\infty} \frac{T_1^m}{m!} |u|^{2m} < \frac{(T_1 u^2)^{A+1}}{(A+1)!} (1 + T_1 u^2 + (T_1 u^2)^2 + \cdots).$$

This last quantity equals

$$\frac{(T_1 u^2)^{A+1}}{(A+1)!} \frac{1}{1 - T_1 u^2}$$

if $T_1 u^2 < 1$, and so is less than $2(T_1 u^2)^{A+1}/(A+1)!$ if $T_1 u^2 < \frac{1}{2}$. Thus, the error introduced by replacing the infinite product in (4.10)–(4.11) with (4.12) is bounded, in norm, by

$$\frac{1}{2\pi} \sum_{-C \le n\varepsilon \le C} \varepsilon \frac{|\sin n\varepsilon|}{n\varepsilon} \prod_{0<\gamma\le X} \bigg|J_0\Big(\frac{2n\varepsilon}{\sqrt{\frac{1}{4} + \gamma^2}}\Big)\bigg| \times 2\frac{(T_1 n^2 \varepsilon^2)^{A+1}}{(A+1)!}$$

if $T_1 n^2 \varepsilon^2 < \frac{1}{2}$. To carry out this sum we first needed to compute the T_1's. This is described shortly. For $\delta(P_1^{\mathrm{comp}})$, using $X = 88190$, $A = 2$, $C = 50$, and $\varepsilon = \frac{1}{20}$, this sum is less than 3×10^{-10}. For all the other δ's, using $X = 9999$, $A = 1$, $C = 25$, and $\varepsilon = \frac{1}{20}$, this sum is less than 2×10^{-6}.

So, for most of our computations, we only needed a compensating polynomial of the form $p(u) = 1 + b_1 u^2$, the exception being the computation of $\delta(P_1^{\mathrm{comp}})$, where we used $p(u) = 1 + b_1 u^2 + b_2 u^4$.

From the definition of the b_j's, we see that

$$b_1 = -T_1(X) = -\bigg(\sum_{\gamma>0} - \sum_{0<\gamma\le X}\bigg) \frac{1}{\frac{1}{4} + \gamma^2}.$$

Now, it is known [Davenport 1980, pp. 80–83] that, assuming GRH,

$$\sum_{\gamma_\zeta > 0} \frac{1}{\frac{1}{4} + \gamma_\zeta^2} = \tfrac{1}{2}\gamma + 1 - \tfrac{1}{2}\log(4\pi) = .0230957089661210338\ldots \quad (4.13)$$

and

$$\sum_{\gamma_{\chi_1} > 0} \frac{1}{\frac{1}{4} + \gamma_{\chi_1}^2} = \tfrac{1}{2}\log\Big(\frac{q}{\pi}\Big) - \tfrac{1}{2}\gamma - \tfrac{1}{2}(\chi_1(-1) + 1)\log 2 + \frac{L'}{L}(1, \chi_1), \quad (4.14)$$

where, overloading the notation,

$$\gamma = \lim_{N\to\infty} \sum_{1}^{N} \frac{1}{n} - \log N = 0.577215664901532\ldots$$

is Euler's constant.

To compute $L'/L(1, \chi_1)$ we evaluated $L'(1, \chi_1)$ and $L(1, \chi_1)$ separately and then divided. $L(1, \chi_1)$

was calculated using the formulas in [Davenport 1980, pp. 8–9], according to which it equals

$$\begin{aligned} -\frac{1}{\sqrt{q}}\sum_{1}^{q-1}\chi_1(m)\log\Bigl(2\sin\Bigl(\frac{\pi m}{q}\Bigr)\Bigr) & \quad \text{if } q\equiv 1 \bmod 4,\\ -\frac{\pi}{q^{3/2}}\sum_{1}^{q-1} m\chi_1(m) & \quad \text{if } q\equiv 3 \bmod 4,\\ \frac{\pi}{4} & \quad \text{if } q=4. \end{aligned}$$

The $L'(1,\chi_1)$'s were computed using Dirichlet's formula [Davenport 1980, p. 11]

$$\Gamma(s)L(s,\chi_1)=\int_0^\infty \frac{h(e^{-u})}{1-e^{-uq}}u^{s-1}e^{-u}\,du,$$

where

$$h(x)=\sum_{1}^{q-1}\chi_1(m)x^{m-1}.$$

Differentiating we get,

$$\Gamma'(s)L(s,\chi_1)+\Gamma(s)L'(s,\chi_1) = \int_0^\infty \frac{h(e^{-u})}{1-e^{-uq}}\log(u)u^{s-1}e^{-u}\,du,$$

which at $s=1$ becomes

$$L'(1,\chi_1)=\gamma L(1,\chi_1)+\int_0^\infty\frac{h(e^{-u})}{1-e^{-uq}}\log(u)e^{-u}\,du.$$

since $\Gamma'(1)=-\gamma$, $\Gamma(1)=1$. Maple [Char et al. 1991] was used to perform the integral numerically, and combining the results with our earlier computed values of $L(1,\chi_1)$'s we got our $L'(1,\chi_1)$'s. With these numbers in our hands, we were then able, using (4.13)–(4.14), to evaluate the $T_1(0)$'s; see Table 2.

Thus, our final formula for $\delta(P_{q;N,R})$ is

$$\delta(P_{q;N,R})=\frac{1}{2\pi}\sum_{-25\le n\varepsilon\le 25}\varepsilon\frac{\sin(n\varepsilon)}{n\varepsilon}(1+b_1(n\varepsilon)^2) \times\prod_{0<\gamma\le 9999}J_0\Bigl(\frac{2n\varepsilon}{\sqrt{\frac14+\gamma^2}}\Bigr)+\tfrac12+\text{error}$$

where

$$b_1=-T_1(0)+\sum_{0<\gamma\le 9999}\frac{1}{\frac14+\gamma^2}.$$

Recall that the error in this formula accounts for replacing the integral by a sum of rectangles of width $\varepsilon=\frac{1}{20}$, cutting off the infinite domain of summation at ± 25, and replacing the infinite product by a finite product and a compensating polynomial of the form $p(u)=1+b_1u^2$. In all instances, using the estimates made earlier, the error was less than 2.5×10^{-6} in norm, and did not have an effect on the first four decimal places of the δ's given at the beginning of Section 4 (for $q=3,4,5,7,11,13$).

To compute $\delta(P_1^{\text{comp}})$, as already mentioned, we replaced the infinite product in (4.10)–(4.11) by

$$(1+b_1u^2+b_2u^4)\prod_{0<\gamma\le 88190}J_0\Bigl(\frac{2u}{\sqrt{\frac14+\gamma^2}}\Bigr),$$

where

$$\begin{aligned} b_1&=-\sum_{\gamma>88190}\frac{1}{\frac14+\gamma^2}\\ &=-\tfrac12\gamma-1+\tfrac12\log(4\pi)+\sum_{0<\gamma\le 88190}\frac{1}{\frac14+\gamma^2}\\ &=-.0230957089661210338\ldots+\sum_{0<\gamma\le 88190}\frac{1}{\frac14+\gamma^2} \end{aligned}$$

and

$$\begin{aligned} b_2&=\sum_{\gamma>88190}\frac{1}{4(\frac14+\gamma^2)^2}+\sum_{\gamma_j>\gamma_k>88190}\frac{1}{\frac14+\gamma_j^2}\frac{1}{\frac14+\gamma_k^2}\\ &=\sum_{\gamma>88190}\frac{4}{(1+4\gamma^2)^2}\\ &\quad+8\biggl(\Bigl(\sum_{\gamma>88190}\frac{1}{1+4\gamma^2}\Bigr)^2-\sum_{\gamma>88190}\frac{1}{(1+4\gamma^2)^2}\biggr)\\ &=\frac{b_1^2}{2}-4\Bigl(\sum_{\gamma>0}-\sum_{0<\gamma\le 88190}\Bigr)\frac{1}{(1+4\gamma^2)^2}. \end{aligned} \tag{4.15}$$

Now

$$\begin{aligned} \sum_{\gamma>0}\frac{1}{(1+4\gamma^2)^2}&=-\frac14\sum_{\gamma>0}\Bigl(\frac{1}{2\gamma+i}-\frac{1}{2\gamma-i}\Bigr)^2\\ &=\frac12\sum_{\gamma>0}\frac{1}{1+4\gamma^2}-\frac14\sum_{\text{all }\gamma}\frac{1}{(2\gamma+i)^2}. \end{aligned}$$

q	$L(1,\chi_1)$	$L'(1,\chi_1)$	$T_1(0)$
3	0.6045997880780726168646	0.2226629869686015094866	0.05661498492873617
4	0.7853981633974483096156	0.1929013167969124293631	0.07778398996179296
5	0.4304089409640040388894	0.3562406470307614988646	0.07827847699714324
7	1.1874104117237259487 84	0.0185659810930280571715	0.12761798914591051
11	0.9472258250994829364296	−0.0797737527762439195432	0.25375655672667782
13	0.6627353910718455897136	0.3114667901362450908264	0.19832628962613668

TABLE 2. Values of $L(1,\chi_1)$, $L'(1,\chi_1)$ and $T_1(0)=\sum_{\gamma>0}1/(\frac14+\gamma^2)$ for $q=3,4,5,7,11,13$.

The first of these sums we already know to equal $\frac18(\frac12\gamma+1-\frac12\log(4\pi))$. The second sum we determine by differentiating the formula [Davenport 1980, p. 80]

$$\frac{\zeta'}{\zeta}(s)+\frac{1}{s-1}+\frac12\frac{\Gamma'}{\Gamma}\Bigl(\frac{s}{2}+1\Bigr)+K=\sum_\rho\Bigl(\frac{1}{s-\rho}+\frac{1}{\rho}\Bigr),$$

where K is a constant and ρ runs over all the nontrivial zeros of $\zeta(s)$; then substituting $s=1$ and dividing by -16. On the right we get, assuming the Riemann Hypothesis,

$$-\frac14\sum_{\text{all }\gamma}\frac{1}{(2\gamma+i)^2}.$$

On the left, we use

$$\zeta(s)=\frac{1}{s-1}+\sum_0^\infty a_m(s-1)^m$$

with $a_0=\gamma=0.577215664901532\ldots$ and

$$a_1=\lim_{N\to\infty}\tfrac12\log^2(N+1)-\sum_1^N\frac{\log(m+1)}{m+1},$$

and also use

$$\frac12\frac{\Gamma'}{\Gamma}\Bigl(\frac{s}{2}+1\Bigr)=-\frac{\gamma}{2}-\sum_1^\infty\Bigl(\frac{1}{s+2n}-\frac{1}{2n}\Bigr)$$

to obtain

$$\begin{aligned}-\frac14\sum_{\text{all }\gamma}\frac{1}{(2\gamma+i)^2}&=-\tfrac1{16}(2a_1-a_0^2+\tfrac34\zeta(2)-1)\\&=-\tfrac1{16}(2a_1-\gamma^2+\tfrac18\pi^2-1).\end{aligned}$$

Hence

$$\begin{aligned}\sum_{\gamma>0}\frac{1}{(1+4\gamma^2)^2}&=\tfrac18(\tfrac12\gamma+1-\tfrac12\log(4\pi))\\&\qquad-\tfrac1{16}(2a_1-\gamma^2+\tfrac18\pi^2-1)\\&=\tfrac18(\tfrac12\gamma+1-\tfrac12\log(4\pi))\\&=0.0000023187897773415544 69\ldots\end{aligned}$$

The value of a_1 was obtained from Maple, which knows how to calculate the a_m's to great precision. With this number we were able, using (4.15), to evaluate b_2 and thus find

$$\begin{aligned}\delta(P_1^{\text{comp}})&=\frac{1}{2\pi}\sum_{-50\le n\varepsilon\le 50}\varepsilon\frac{\sin n\varepsilon}{n\varepsilon}(1+b_1(n\varepsilon)^2+b_2(n\varepsilon)^4)\\&\qquad\times\prod_{0<\gamma\le 88190}J_0\Bigl(\frac{2n\varepsilon}{\sqrt{\frac14+\gamma^2}}\Bigr)+\tfrac12+\text{error}\\&=.99999973\ldots,\end{aligned}$$

where the error is less than 6×10^{-10} in norm.

Figure 1 shows graphs of the density functions of μ_1 and $\mu_{q;R,N}$, for $q=1,3,4,5,7,11,13$, obtained by evaluating the Fourier transform of (3.4). Also shown are the histograms representing (logarithmic) distributions numerically computed for

$$F(x)=-1+\frac{\psi(x)-x}{\sqrt{x}} \tag{4.16}$$

and

$$F(x)=-1+\frac{\psi(x,\chi_1)}{\sqrt{x}}, \tag{4.17}$$

for x in the range $10^5\le x\le 10^{10}$. One can show, using the method of Lemma 2.1, that (4.16) and

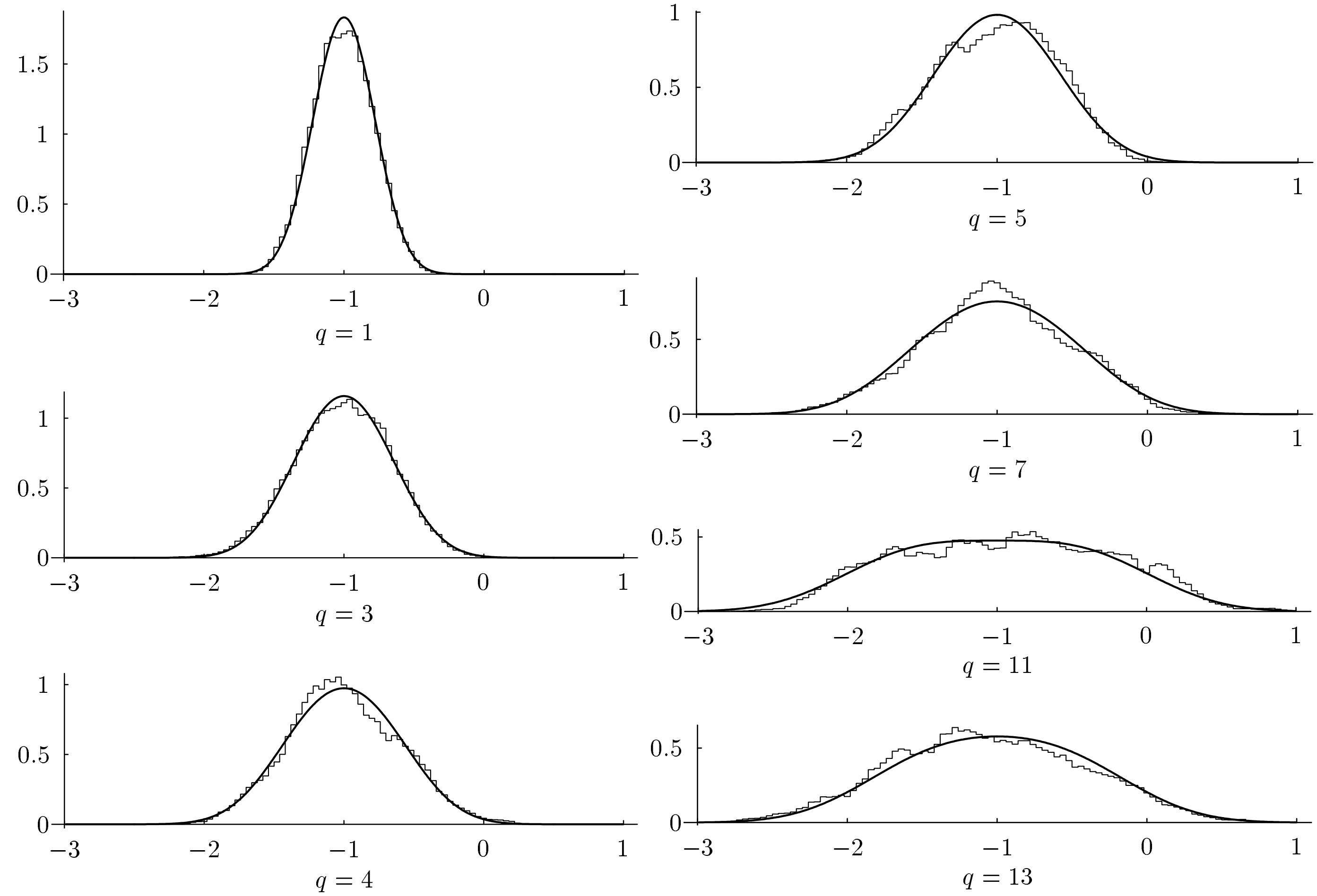

FIGURE 1. Predicted density functions (curves) of μ_1 and $\mu_{q;R,N}$, for $q = 3, 4, 5, 7, 11, 13$, compared with experimental data (histograms) of the logarithmic distribution of the function $F(x)$ of (4.16) (for $q = 1$) and (4.17) (for $q = 3, 4, 5, 7, 11, 13$). The real line was divided into intervals (buckets) of width $\frac{1}{25}$. Using a sieve, we then evaluated (4.16) and (4.17) at $x = n + \frac{1}{2}$, for all $10^5 \leq n \leq 10^{10}$, and for each x we added x^{-1} to the bucket containing $F(x)$. Finally, we scaled the histograms so as to have area one.

(4.17) have the same (logarithmic) distributions as $E_1(x)$ and $E_{q;R,N}(x)$, respectively. We chose to work with them because the term in $O(1/\log x)$ in Lemma 2.1 is significant enough to skew the distribution in the range that we examined.

5. GENERALIZATIONS

In this short section we discuss generalizations of the Chebyshev bias phenomenon. First, we examine the relative distribution of prime ideals in a number field. Given two ideal classes, one can examine whether there is a preference for primes to be in one class over the other. If we assume the Riemann Hypothesis for the corresponding ideal class L-functions, we obtain results similar to those in Sections 2 and 3. For example, if the class number is 2 there is a bias of primes to be nonprincipal. On the other hand, if the class number is odd there are no biases in pairwise comparisons.

Similarly, one can study the relative distribution of primes according to their splitting in Galois extensions (Chabotarev-type questions). Again, one can prove results analogous to those in Sections 2 and 3. In this case, one has to deal with general Artin L-functions [Lang 1970, Ch. 12]. Here a new feature emerges concerning GSH for such L-functions and some care must be exercised. First, such an L-function may factor into a product of primitive such L-functions and the factors may appear with exponent greater than 1. So, for example, the Dedekind zeta function of a nonabelian Galois extension $K/\mathbb{Q}$ will have multiple zeros and will not satisfy GSH. As far as GSH is concerned, we must restrict ourselves to principal primitive L-functions, as described in [Rudnick and Sarnak 1994], which discusses the statistical distribution of the zeros of such L-functions. In particular, distinct primitive principal L-functions have statistically independent zeros. However, the algebraic GSH for zeros of different primitive Artin L-functions is more subtle. The reason (or at least one reason) is that there is an example [Armitage 1972] of a primitive Artin L-function with a zero at $s = \frac{1}{2}$. This will naturally cause a bias in connection with the problem that we are discussing. This bias should still be considered as algebraic since the vanishing at $s = \frac{1}{2}$ is a consequence of an odd functional equation that emerges from computations of Serre [1971] on Artin conductors and root numbers. One might surmise that besides this relation there are no algebraic relations between the imaginary parts of the zeros of primitive Artin L-functions.

The other generalization that we discuss is an analogous problem in geometry. Let X be a compact hyperbolic surface (that is, of curvature -1). Denote by P the set of primitive closed geodesics (primes) on X and let

$$N(p) = \exp l(p),$$

where $l(p)$ denotes the length of p. Each p determines a homology class $C(p) \in H_1(X)$. Let $\pi_C(x)$ be the number of elements $p \in P$ such that $N(p) \le x$ and $C(p) = C$. In [Phillips and Sarnak 1987] it is shown that, for any C,

$$\pi_C(x) \sim \frac{(g-1)^g x}{\log^{g+1} x}$$

as $x \to \infty$, where $g \ge 2$ is the genus of X. Apparently, the situation is similar to the other examples that we have been considering: the primes are equidistributed amongst the homology classes. We can thus ask whether there are any biases towards one homology class as compared to another. One difference here is that the group $H_1(X) \cong \mathbb{Z}^{2g}$ into which the primes distribute themselves is infinite.

In fact, it turns out that in this case there are very strong biases. Not only can $\delta(P_{C_1,C_2})$ be zero, but there are always (for sufficiently large x) more primes in certain homology classes. The group $H_1(X)$ carries a natural norm coming from the conformal structure on X, defined as follows. Let $\operatorname{Har} X$ denote the space of harmonic one-forms on X. We have pairings $\langle\,,\,\rangle : H_1(X) \times \operatorname{Har} X \to \mathbb{R}$ and $(\,,\,) : \operatorname{Har} X \times \operatorname{Har} X \to \mathbb{R}$ given by

$$\langle C, w\rangle = \int_C w$$

and

$$(w_1, w_2) = \int_X w_1 \wedge *w_2.$$

Using duality we can therefore associate to each $c \in H_1(X)$ a unique dual harmonic one-form η_C that satisfies $\langle C, w\rangle = (\eta_C, w)$ for all $w \in \operatorname{Har} X$. Now, define a norm on $H_1(X)$ by setting $\|C\|^2 := (\eta_C, \eta_C)$. One can show by a careful analysis of the subleading term in the asymptotics developed in [Phillips and Sarnak 1987] that, if $\|C\| > \|D\|$, then $\pi_D(x) > \pi_C(x)$ for x sufficiently large. Thus there are "more" primes homologous to D than to C. In particular, there are more primes homologous to zero than to any other homology class. As we have seen in Sections 2 and 3, such a strong bias is never present in the arithmetic cases.

ACKNOWLEDGEMENTS

We would like to thank G. Davidoff, whose recent RUE project on $\pi_{q;N,R}(x)$ stimulated this work. Also, we thank J. Kaczorowski and A. Odlyzko for pointing us to the relevant literature. Odlyzko also supplied us, along with H. te Riele and R. Rumely,

with the zeros that were used in our computations. Part of this work constitutes the senior thesis of the first author [Rubinstein 1994].

REFERENCES

[Armitage 1972] J. V. Armitage, "Zeta functions with a zero at $s = \frac{1}{2}$", *Inv. Math.* **15** (1972), 199–205.

[Bays and Hudson 1978] C. Bays and R. H. Hudson, "Details of the first region of integers x with $\pi_{3,2}(x) < \pi_{3,1}(x)$", *Math. Comp.* **32** (1978), 571–76.

[Char et al. 1991] B. W. Char et al., *Maple V Language Reference Manual* and *Maple V Library Reference Manual*, Springer, New York, 1991.

[Davenport 1980] H. Davenport, *Multiplicative Number Theory* (2nd. ed), Graduate Texts in Mathematics **74**, Springer, Berlin, 1980.

[Davidoff 1994] G. Davidoff, "Some experimental results on a generalization of Littlewood's Theorem", preprint, Mount Holyoke College, 1994.

[Dirichlet 1837] L. Dirichlet, "Beweis des Satzes, daß jede unbegrenzte arithmetische Progression ... unendlich viele Primzahlen enthält", *Abh. König. Preuss. Akad.*, **34** (1837), 45–81. Reprinted on pp. 313–342 in *Dirichlets Werke*, vol. 1, Reimer, Berlin, 1889–97 and Chelsea, Bronx (NY), 1969.

[Heath-Brown 1992] R. Heath-Brown, "The distribution and moments of the error term in the Dirichlet divisor problem", *Acta Arith.* **60**(4) (1992), 389–415.

[Hooley 1977] C. Hooley. "On the Barban–Davenport–Halberstam Theorem: VII", *J. London Math. Soc.* (2) **16** (1977), 1–8.

[Ingham 1932] A. E. Ingham, *The Distribution of Prime Numbers*, Cambridge University Press, Cambridge, 1932.

[Kaczorowski] J. Kaczorowski, "The boundary values of generalized Dirichlet series and a problem of Chebyshev", preprint, Poznań University.

[Knapowsky and Turan 1962] S. Knapowsky and P. Turan, "Comparative prime number theory I", *Acta. Math. Acad. Sci. Hungar.* **13** (1962), 299–314. Reprinted on pp. 1329–1343 in *Collected papers of Paul Turan*, vol. 2 (edited by P. Erdős), Akadémiai Kiadó, Budapest, 1990.

[Kueh 1988] K. L. Kueh "The moments of infinite series", *J. reine angew. Math.* (1988) **385**, 1–9.

[Lang 1970] S. Lang, *Algebraic Number Theory*, Addison-Wesley, Reading (MA), 1970.

[Leech 1957] J. Leech, "Note on the distribution of prime numbers", *J. London Math. Soc.* **32** (1957), 56–58.

[Lévy 1922] P. Lévy, "Sur la détermination des lois de probabilité par leurs fonctions caractéristiques", *C. R. Acad. Sci. Paris* **175** (1922), 854–856.

[Littlewood 1914] J. E. Littlewood, "Distribution des Nombres Premiers", *C. R. Acad. Sci. Paris* **158** (1914), 1869–1872.

[Littlewood 1928] J. E. Littlewood, "On the Class-Number of the Corpus $P(\sqrt{-k})$", *Proc. London Math. Soc.* (2), **27** (1928), 358–372.

[Montgomery 1980] H. L. Montgomery "The Zeta Function and Prime Numbers", pp.14–24 in *Proceedings of the Queen's Number Theory Conference, 1979* (edited by P. Ribenboim), Queen's University, Kinston (Ont.), 1980.

[Odlyzko] A. Odlyzko, "The 10^{20}-th zero of the Riemann Zeta Function and 70 million of its neighbours", preprint.

[Phillips and Sarnak 1987] R. Phillips and P. Sarnak, "Geodesics in homology classes", *Duke Math. J.* **55** (1987), 287–297.

[te Riele 1986] H. J. J. te Riele, "On the sign change of the difference $\pi(x) - \mathrm{Li}(x)$", *Math. Comp.* **48** (1986), 667–81.

[Rubinstein 1994] M. Rubinstein, "Chebyshev's bias", senior thesis, Princeton University, 1994.

[Rudnick and Sarnak 1994] Z. Rudnick and P. Sarnak, "Zeros of principal L-functions and random matrix theory", preprint, Princeton University, 1994.

[Rumely 1993] R. Rumely, "Numerical computations concerning ERH", *Math Comp.*, **61** (1993), 415–440.

[Serre 1971] J.-P. Serre, "Conducteurs d'Artin des charactères réels", *Inv. Math.* **14** (1971), 173–183.

[Shanks 1959] D. Shanks, "Quadratic residues and the distribution of primes", *Math. Tables* (later *Math. Comp.*) **13** (1959), 272–284.

[Skewes 1933] S. Skewes, "On the difference $\pi(x) - \mathrm{Li}(x)$", *J. London Math. Soc.* **8** (1933), 277–83.

[Stein and Weiss 1971] E. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, Princeton, 1971.

[Watson 1948] G. N. Watson, *A Treatise on the Theory of Bessel Functions* (2nd ed.), Cambridge University Press, Cambridge, 1948.

[Wintner 1938] A. Wintner, "Asymptotic Distributions and Infinite Convolutions", Notes distributed by the Institute for Advanced Study (Princeton), 1938.

[Wintner 1941] A. Wintner, "On the distribution function of the remainder term of the prime number theorem", *Amer. J. Math.* **63** (1941), 233–48.

Michael Rubinstein, Department of Mathematics, Princeton University, Washington Road, Princeton, NJ, 08544 (miker@math.princeton.edu)

Peter Sarnak, Department of Mathematics, Princeton University, Washington Road, Princeton, NJ, 08544 (sarnak@math.princeton.edu)

Received July 7, 1994; accepted in revised form December 10