BIAS IN THE DISTRIBUTION OF PRIMES MODULO 4

SHUBHRAJIT BHATTACHARYA AND ZHENGHENG BAO

ABSTRACT. In this brief note, we discuss the bias in the distribution of primes modulo 4. First observed by Chebyshev in 1853, this phenomenon attracted many mathematicians. Rubinstein and Sarnak first proved [RS94] unconditional results where they attached numerical values to such biases in general for any modulus q. This article is an overview of the case where we determine that there are *typically more primes* $3 \pmod 4$ *than* $1 \pmod 4$. It also discusses why we need to assume the Generalized Riemann Hypothesis.

1. INTRODUCTION

The distribution of primes is one of the most mysterious fields of study for number theorists. Throughout the century, many mathematicians have attempted to study different aspects of it. Dirichlet [DdS37] proved that for any positive integer $q > 1$ and $(q, a) = 1$, there are infinitely many primes $p \equiv a \pmod q$. This indicates roughly an equidistribution of primes in every congruence class modulo any positive integer $q > 1$. This became mathematically more precise when Hadamard and de la Vallée Poussin proved, in 1890s, the asymptotic formula

$$\pi(x; q, a) \sim \frac{1}{\varphi(q)}\pi(x) \sim \frac{1}{\varphi(q)}\frac{x}{\log x} \tag{1.1}$$

where $\pi(x)$ and $\pi(x; q, a)$ denote the number of primes $\leq x$ and number of primes $p \leq x$ with $p \equiv a \pmod q$ for $q > 1$ and $(a, q) = 1$ respectively. Since the number of residue classes $a \pmod q$ with $(a, q) = 1$ is given by the Euler totient function $\varphi(q)$, equation (1.1) tells us that for large x, primes are uniformly distributed across the reduced residue classes modulo any integer $q > 1$.

In 1853, Chebyshev noticed that there are more primes in the congruence class $3 \pmod 4$ than in $1 \pmod 4$. The table on the next page (data taken from [GM04, page 1]) presents a comparison between the quantities $\pi(x; 4, 1)$ and $\pi(x; 4, 3)$ for some $x \leq 50,000$. This table may lure us to conjecture that $\pi(x; 4, 3) > \pi(x; 4, 1)$ for all $x \geq 1$; however, Leech [Lee57] found that $\pi(x; 4, 3) < \pi(x; 4, 1)$ for $x = 26861$ and this is the first such x! Nevertheless, the large size of this first counter-example still indicates the bias of primes towards the congruence class $3 \pmod 4$.

But how do we mathematically establish the statement *there are more primes* $p \equiv 3 \pmod 4$ *than primes* $p \equiv 1 \pmod 4$? According to the Prime Number Theorem of Hadamard/de la Vallée Poussin,

$$\lim_{x\to\infty} \frac{\pi(x; 4, 3)}{\pi(x; 4, 1)} = 1 \tag{1.2}$$

that says there *are asymptotically an equal number of primes* in both residue classes as $x \to \infty$. Therefore, we have to devise stronger measurement techniques. Rubinstein & Sarnak [RS94] consider the sets $P_{q;a,b} = \{x \geq 2 \colon \pi(x; q, a) > \pi(x; q, b)\}$ for two reduced residue classes a, b modulo q. With an appropriate notion of size, they prove that $P_{4;3,1}$ is larger than $P_{4;1,3}$. In [RS94], they computed the logarithmic density, which will be defined in the following section, of such sets and found that the logarithmic density of $P_{4;3,1}$ equals $0.9959\ldots$, which clearly indicates the bias in the distribution of prime modulo 4!

x	Number of primes $4n+3$ up to x	Number of primes $4n+1$ up to x
100	13	11
200	24	21
300	32	29
400	40	37
500	50	44
600	57	51
700	65	59
800	71	67
900	79	74
1000	87	80
2000	155	147
3000	218	211
4000	280	269
5000	339	329
6000	399	383
7000	457	442
8000	507	499
9000	562	554
10,000	619	609
20,000	1136	1125
50,000	2583	2549

Table 1. The number of primes of the form $4n+1$ and $4n+3$ up to x.

However, results in [RS94] are under the assumption of the Generalized Riemann Hypothesis (GRH) and the Linear Independence hypothesis (LI), which will be defined later. The unconditional results known so far was proved, before Rubinstein & Sarnak's results, by Littlewood [Lit14] that both sets $P_{4;3,1}$ and $P_{4;1,3}$ are unbounded.

Even if there are huge biases for small q, they disappear as $q \to \infty$. Rubinstein & Sarnak [RS94] prove that $\lim_{q\to\infty} \max_{(a,q)=(b,q)=1} |\delta(P_{q;a,b}) - 1/2| = 0$. In this note, we confine ourselves to the modulo 4 race, first studied by Chebyshev, that gave birth to this fascinating subject in number theory, now known as *The Chebyshev Bias / The Prime Number Race*!

2. Preliminaries

A Dirichlet character modulo q is a group homomorphism $\chi\colon (\mathbb{Z}/q\mathbb{Z})^\times \longrightarrow \mathbb{C}^\times$, which is extended to $\chi\colon \mathbb{Z} \to \mathbb{C}^\times$ by assigning $\chi(n) = 0$ for $(n,q) > 1$. To any such Dirichlet character χ, one attaches an L-function, known as the Dirichlet L-function,

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}, \tag{2.1}$$

which is absolutely convergent for $\Re(s) > 1$. It can be shown that it has an analytic continuation to $\mathbb{C}$, with at most one pole at $s = 1$. These L-functions are the most important objects while counting primes in a specific congruence class modulo an integer $q > 1$. For the trivial character χ_0 modulo 1, we get the famous Riemann zeta function $L(\chi_0, s) = \zeta(s)$. The Riemann Zeta $\zeta(s)$ has some trivial zeros: $\zeta(-2) = \zeta(-4) = \zeta(-6) = \cdots = 0$. Other zeros lie in the critical strip $0 < \Re(s) < 1$, and contribute to the counting of primes. After playing around with his creations, Riemann conjectured that the nontrivial zeros of ζ all lie on the line $\Re(s) = 1/2$. This is known as the Riemann Hypothesis, and is still wide open! After studying these L-functions, mathematicians expect this to happen for all Dirichlet L-functions.

Conjecture 2.1 (The Generalized Riemann Hypothesis (GRH))**.** Let $L(s, \chi)$ be a Dirichlet L-function. If $\sigma + it$ is a complex number with $\sigma \in (0, 1]$ and $L(\sigma + it, \chi) = 0$, then $\sigma = 1/2$.

The key step in Rubinstein & Sarnak's paper is to consider the logarithmic density (definition (2.2)), which will be used to measure the sets $P_{4;3,1}$. Wintner [Win41] did some investigations with the remainder term in the prime number theory and from his work, it is evident that logarithmic density is a good way to measure $P_{q;a,b}$ in general, which is done in [RS94].

Definition 2.2 ([RS94])**.** Let P be a set of real numbers. Define the upper and lower logarithmic densities $\overline{\delta}(P)$ and $\underline{\delta}(P)$ respectively as

$$\overline{\delta}(P) = \limsup_{X\to\infty} \frac{1}{\log X} \int_{t\in P\cap[2,X]} \frac{dt}{t}, \quad \underline{\delta}(P) = \liminf_{X\to\infty} \frac{1}{\log X} \int_{t\in P\cap[2,X]} \frac{dt}{t} \tag{2.2}$$

and the *logarithmic density* of P exists and equals $\delta(P) = \overline{\delta}(P) = \underline{\delta}(P)$ if the latter two are equal.

Rubinstein & Sarnak's work was stimulated by Davidoff's 1994 paper [Dav94] which considered the race between quadratic residue and nonresidue primes modulo q. An integer a is said to be a quadratic residue if $x^2 \equiv a \pmod{q}$ has a solution, and a nonresidue otherwise. Both Davidoff and Rubinstein & Sarnak considered the case when $q = p^\alpha, 2p^\alpha, 4$ for some odd prime p. This is because of the existence of primitive roots of unity modulo these integers. In this note, we also assume that $q \in \{p^\alpha, 2p^\alpha, 4\}$ for odd primes p.

Following [RS94], define $\pi_R(x; q)$ to be the number of primes $p \le x$ such that p is a quadratic residue modulo q. Similarly, $\pi_N(x; q)$ will denote the number of primes $p \le x$ which are quadratic nonresidues modulo q. Also, define

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

According to the computations in Rubinstein & Sarnak [RS94], we see that there is always a bias toward nonresidues. For $q = 4$, the integer 3 is the only quadratic nonresidue and 1 is the only quadratic residue. Therefore, $P_{4;N,R} = P_{4;3,1}$ and we will see that $\delta(P_{4;3,1}) = 0.9959...$, which numerically establishes Chebyshev's observations.

Here we take a moment to introduce the arithmetic functions we will need for the rest of this note.

Definition 2.3. We define

$$\Lambda(n) = \begin{cases} \log p & n = p^a \\ 0 & \text{otherwise} \end{cases} \qquad \psi(x,\chi) = \sum_{n\le x} \chi(n)\Lambda(n) \tag{2.3}$$

$$\psi(x;q,a) = \sum_{\substack{n\le x \\ n\equiv a \pmod q}} \Lambda(n) \qquad \theta(x;q,a) = \sum_{\substack{p\le x \\ p\equiv a \pmod q}} \log p \tag{2.4}$$

Here, $\Lambda(n)$ is known as Von Manglodt's function, and θ and ψ are variants of Chebyshev's θ and ψ functions.

In order to compute the logarithmic densities of $P_{q;N,R}$ (respectively $P_{q;R,N}$), [RS94] proved the existence of limiting logarithmic distributions of the normalized functions $E_{q;N,R}$ (respectively $E_{q;R,N}$) defined by

$$E_{q;N,R}(x) = \frac{\log x}{\sqrt{x}} \left(\pi_N(x,q) - \pi_R(x,q)\right). \tag{2.5}$$

By a limiting logarithmic distribution, we mean a measure $\mu_{q;N,R}$ such that

$$\lim_{X\to\infty} \frac{1}{\log X} \int_2^X f(E_{q;N,R}(x))\frac{dx}{x} = \int_{\mathbb{R}} f(x)d\mu_{q;N,R}(x) \tag{2.6}$$

for any bounded continuous function f on $\mathbb{R}$. Note that if f is absolutely continuous then $\delta(P_{q;N,R}) = \mu_{q;N,R}(\{x \in \mathbb{R} \colon x > 0\})$. But even GRH does not guarantee the existence of the densities $\delta(P_{q;N,R})$. The proof of [RS94, Theorem 1.1] may not be true for characteristic functions of nice sets like $\{x \in \mathbb{R} \colon x > 0\}$. Instead, they constructed similar measures, defined in terms of the nontrivial zeros of the Dirichlet L-functions $L(s,\chi)$, that estimate the measures $\mu_{q;N,R}$. One more assumption is made about the nontrivial zeros of Dirichlet L-functions, known as the Linear Independence(LI) hypothesis/Grand Simplicity Hypothesis (GSH) [Dav13].

Conjecture 2.4. The set $\gamma \geq 0$ such that $L(1/2 + i\gamma, \chi) = 0$, as χ runs through the set of primitive Dirichlet characters, is linearly independent over $\mathbb{Q}$.

Two of the immediate consequences of LI are that 0 can't be the imaginary part of a nontrivial zero (in other words, $L(1/2,\chi) \neq 0$) and that all the zeros are simple. GRH and LI together give the following explicit formula for the Fourier transform of $\mu_{q;N,R}$ ((3.4) in [RS94]):

$$\hat{\mu}_{q;R,N}(\xi) = e^{i\xi} \prod_{\gamma_{\chi_1}>0} J_0\left(\frac{2\xi}{\sqrt{1/4+\gamma_{\chi_1}^2}}\right) \tag{2.7}$$

where χ_1 is the real nontrivial character modulo q and J_0 is the Bessel function defined as

$$J_0(z) = \sum_{m=0}^{\infty} \frac{(-1)^m(\frac{1}{2}z)^{2m}}{(m!)^2}. \tag{2.8}$$

Observe that $\chi_1(a) = 1$ (respectively -1) when a is a quadratic residue (respectively nonresidue) modulo q. Therefore,

$$\frac{\log x}{\sqrt{x}} \sum_{p \leq x} \chi_1(p) = \frac{\log x}{\sqrt{x}} (\pi_R(x;q) - \pi_N(x;q)) = -E_{q;N,R}(x). \tag{2.9}$$

Under GRH and LI, the limiting distribution $\tilde{\mu}_{q;N,R}$ of the normalized function $E_{q;N,R}/\sqrt{\log q}$ converges in measure to the Gaussian $(2\pi)^{-1/2}e^{-x^2/2}\,dx$ as $q \to \infty$ [RS94, Theorem 1.6]

Before sketching the method of computing the densities $\delta(P_{q;N,R})$, we briefly describe the role of GRH in Rubinstein & Sarnak's work [RS94].

3. Contributions of the Generalized Riemann Hypothesis

As we know, the Generalized Riemann Hypothesis proved to be shockingly important in number theory and many other branches of mathematics. Assuming GRH, mathematicians have proved many interesting results. Although the Generalized Riemann Hypothesis is widely believed to be true, no proof is known to us yet, even for the Riemann Hypothesis, which is a special case of the Generalized Riemann Hypothesis.

> *Right now, when we tackle problems without knowing the truth of the Riemann hypothesis, it's as if we have a screwdriver. But when we have it, it'll be more like a bulldozer.* [Kla00]
>
> **Peter Sarnak**

The main tool for establishing the existence of the densities $\mu_{q;N,R}$, as done in [RS94], is the explicit formula for the prime counting function $\pi(x)$, sometimes called *Riemann's Revolutionary Formula* ([GM04]), that states

$$\frac{\int_2^x \frac{dt}{\log t} - \pi(x)}{\sqrt{x}/\log x} \approx 1 + 2 \times \left(\sum_{\substack{\gamma > 0 \\ \zeta(1/2+i\gamma)=0}} \frac{\sin(\gamma \log x)}{\gamma} \right) \tag{3.1}$$

As an analog of this for $\pi(x;q,a)$, [RS94] considered $\psi(x,\chi)$ for a Dirichlet character χ modulo q, as defined in the previous section. As is shown in [Dav13, pp. 115–120], if $\chi \neq \chi_0$, $x \geq 2$ and $X \geq 1$ we have

$$\psi(x,\chi) = -\sum_{|\gamma_\chi| \leq X} \frac{x^\rho}{\rho} + O\left(\frac{x \log^2(xX)}{X} + \log x \right), \tag{3.2}$$

where $\rho = \beta_\chi + i\gamma_\chi$ runs over the zeros of $L(s,\chi)$ in $0 < \mathrm{Re}(s) < 1$, and the implied O-constant depends on q. Therefore, under GRH, we have $\beta_\chi = \frac{1}{2}$ and equation (3.2) becomes

$$\psi(x,\chi) = -\sqrt{x} \sum_{|\gamma_\chi| \leq X} \frac{x^{i\gamma_\chi}}{\frac{1}{2} + i\gamma_\chi} + O\left(\frac{x \log^2(xX)}{X} + \log x \right). \tag{3.3}$$

Define

$$E(x;q,a) = (\varphi(q)\pi(x;q,a) - \pi(x))\log x/\sqrt{x} \tag{3.4}$$

$$c(q,a) = -1 + \sum_{\substack{b^2\equiv a \pmod q\\ 1\le b\le q-1}} 1. \tag{3.5}$$

Then [RS94] proves the following formula:

$$E(x,q,a) = -c(q,a) + \sum_{\chi\neq\chi_0} \bar{\chi}(a)\frac{\psi(x,\chi)}{\sqrt{x}} + O\left(\frac{1}{\log x}\right). \tag{3.6}$$

When $q \in \{p^\alpha, 2p^\alpha, 4\}$ for odd primes p (in other words, there are primitive roots modulo q) and a is a quadratic residue then $c(q,a) = 1$, and $c(q,a) = -1$ when a is a quadratic nonresidue modulo q. This shows that the constant term $-c(q,a)$ in equation (3.6) is responsible for the bias toward nonresidues. For the case $q = 4$, note that $\pi_N(x;4) = \pi(x;4,3)$ and $\pi_R(x;4) = \pi(x;4,1)$ since 1 and 3 are the only quadratic residue and nonresidue modulo 4 respectively. Therefore, $E(x;4,3) - E(x;4,1) = 2E_{4;3,1}(x)$.

We have skipped most of the technical details/proof in this note but we would like to sketch a proof of expression (3.6) since this formula along with GRH indicates the reason behind the bias. As defined in equation (2.4), we may write $\pi(x;q,a)$ as the Riemann-Stieltjes integral

$$\pi(x,q,a) = \int_2^x \frac{d\theta(t,q,a)}{\log t}; \tag{3.7}$$

from the prime theorem in arithmetic progressions,

$$\psi(x,q,a) = \theta(x,q,a) + \left(\sum_{b^2\equiv a(q)} 1\right)\frac{\sqrt{x}}{\varphi(q)} + O\left(\frac{\sqrt{x}}{\log x}\right). \tag{3.8}$$

Solving for $\theta(x,q,a)$ and combining with the fact that

$$\psi(x,q,a) = \frac{1}{\varphi(q)}\sum_{\chi \bmod q}\bar{\chi}(a)\sum_{n\le x}\Lambda(n)\chi(n) = \frac{1}{\varphi(q)}\sum_{\chi \bmod q}\bar{\chi}(a)\psi(x,\chi), \tag{3.9}$$

we get

$$\begin{aligned}\int_2^x \frac{d\theta(t,q,a)}{\log t} &= \frac{1}{\varphi(q)}\int_2^x \frac{d\psi(t)}{\log t} + \frac{1}{\varphi(q)}\sum_{\chi\neq\chi_0}\bar\chi(a)\int_2^x \frac{d\psi(t,\chi)}{\log t} - \frac{1}{\varphi(q)}\left(\sum_{b^2\equiv a(q)} 1\right)\frac{\sqrt{x}}{\log x} \\ &\quad + O\left(\frac{\sqrt{x}}{\log^2 x}\right) \\ &= \frac{1}{\varphi(q)}\left(\pi(x) + \frac{\sqrt{x}}{\log x}\right) + \frac{1}{\varphi(q)}\sum_{\chi\neq\chi_0}\bar\chi(a)\frac{\psi(x,\chi)}{\log x} - \frac{1}{\varphi(q)}\left(\sum_{b^2\equiv a(q)} 1\right)\frac{\sqrt{x}}{\log x} \\ &\quad + O\left(\sum_{\chi\neq\chi_0}\left|\int_2^x \frac{\psi(t,\chi)}{t\log^2 t}dt\right| + \frac{\sqrt{x}}{\log^2 x}\right) \\ &= \frac{\pi(x)}{\varphi(q)} + \frac{1}{\varphi(q)}\sum_{\chi\neq\chi_0}\bar\chi(a)\frac{\psi(x,\chi)}{\log x} - \frac{c(q,a)}{\varphi(q)}\frac{\sqrt{x}}{\log x} \\ &\quad + O\left(\sum_{\chi\neq\chi_0}\left|\int_2^x \frac{\psi(t,\chi)}{t\log^2 t}dt\right| + \frac{\sqrt{x}}{\log^2 x}\right). \end{aligned} \tag{3.10}$$

As shown in the proof of in [RS94, Lemma 2.1], the error term in equation (3.10) becomes $O(\sqrt{x}/\log^2 x)$ and hence it gives us the following

$$\pi(x,q,a) - \frac{\pi(x)}{\varphi(q)} = -\frac{c(q,a)}{\varphi(q)}\frac{\sqrt{x}}{\log x} + \frac{1}{\varphi(q)\log x}\sum_{\chi\neq\chi_0}\bar\chi(a)\psi(x,\chi) + O\left(\frac{\sqrt{x}}{\log^2 x}\right). \tag{3.11}$$

Combining equation (3.6) with equation (3.3) we get, for $T \geq 1$ and $2 \leq x \leq X$,

$$E(x,q,a) = -c(q,a) - \sum_{\chi\neq\chi_0}\bar\chi(a)\sum_{|\gamma_\chi|\leq T}\frac{x^{i\gamma_\chi}}{\frac{1}{2}+i\gamma_\chi} + \varepsilon_a(x,T,X). \tag{3.12}$$

After estimating the term $\varepsilon_a(x,T,x)$, the existence of the limiting distributions is established. From this discussion, we get a sense of how [RS94] uses GRH in a crucial way. As we mentioned earlier, LI is equally important but we skip that discussion here. One may consult Rubinstein & Sarnak's original paper [RS94, Section 3] for details.

4. Computing $\delta(P_{4;3,1})$

In [RS94, Section 4] the following densities are computed. They show that the accuracy is at least 20 decimal digits, though we are only showing the first 4 digits.

$$\delta(P_{3;N;R}) = 0.9990\ldots$$
$$\delta(P_{4;N;R}) = 0.9959\ldots$$
$$\delta(P_{5;N;R}) = 0.9954\ldots$$
$$\delta(P_{7;N;R}) = 0.9782\ldots$$
$$\delta(P_{11;N;R}) = 0.9167\ldots$$
$$\delta(P_{13;N;R}) = 0.9443\ldots$$

Table 2. Densities of $P_{q;N,R}$ for $q = 3, 4, 5, 7, 11, 13$

We follow [RS94] to give an overview of the computations.

Recall that

$$\hat{\mu}_{q;R,N}(\xi) = e^{i\xi} \prod_{\gamma_{\chi_1}>0} J_0\left(\frac{2\xi}{\sqrt{1/4+\gamma_{\chi_1}^2}}\right) \tag{4.1}$$

where χ_1 is the real nontrivial character modulo q and J_0 is the Bessel function defined in equation (2.8). Let $f_{q;N,R}(t)$ denote the density function of $\mu_{q;R,N}$ and $g(t) = f_{q;N,R}(t-1)$. Since J_0 is an even function, so is the above product for $\hat{\mu}$. So $\mu_{q;R,N}$ is symmetric about $t = -1$. So g is symmetric about $t = 0$ (which is easier than f to work), and its Fourier transform, which we denote by $\hat{\omega}$, is

$$\hat{\omega}(\xi) = \prod_{\gamma>0} J_0\left(\frac{2\xi}{\sqrt{1/4+\gamma^2}}\right). \tag{4.2}$$

Notice that $\delta(P_{q;N,R}) = \int_{-\infty}^{1} d\omega_{q;R,N}(t)$. By symmetry of $g_{q;R,N}$ about $t = 0$, we know that

$$\delta(P_{q;N,R}) = \frac{1}{2}\left(\int_{-\infty}^{1} + \int_{1}^{\infty}\right) d\omega_{q;R,N}(t) \tag{4.3}$$

$$= \frac{1}{2} + \frac{1}{2}\int_{-1}^{1} d\omega_{q;R,N}(t) \tag{4.4}$$

$$= \frac{1}{2} + \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\sin u}{u}\hat{\omega}_{q;R,N}(u)du \tag{4.5}$$

Here, the middle equality is because the integral from $-\infty$ to ∞ of the $\omega_{q;R,N}$, which is a probability measure, is 1; and the last equality is by the Fourier inversion formula of the characteristic function $\chi_{[-1,1]}$. Now we have expressed $\delta(P_{q;N,R})$ in terms of $\hat{\omega}_{q;R,N}$, and we have the formula (4.2) for $\hat{\omega}_{q;R,N}$. The integral (4.3) is estimated in [RS94] in three steps. We only sketch the first step here, which is to replace the integral with a sum, according to [RS94, p. 189].

The way to replace the integral with a sum is to use the Poisson summation formula. The Poisson formula tells us that

$$\varepsilon \sum_{n\in\mathbb{Z}} \phi(\varepsilon n) = \sum_{n\in Z} \hat{\phi}\left(\frac{n}{\varepsilon}\right) = \hat{\phi}(0) + \sum_{n\in\mathbb{Z},n\neq 0} \hat{\phi}\left(\frac{n}{\varepsilon}\right) \tag{4.6}$$

for any continuous and rapidly decreasing function ϕ. We apply this formula to

$$\phi(u) = \frac{1}{2\pi}\frac{\sin u}{u}\hat{\omega}(u),$$

obtaining

$$\hat{\phi}(x) = \frac{1}{2}\int_{x-1}^{x+1} g(u)du = \frac{1}{2}\int_{x-1}^{x+1} d\omega(u). \tag{4.7}$$

The fact that ϕ satisfies the conditions for applying the Poisson summation formula follows from [Wat48] and [SW71]. We skip the details here. Hence

$$\begin{aligned}\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\sin u}{u}\hat{\omega}(u)du &= \hat{\phi}(0)\\ &= \varepsilon\sum_{n\in\mathbb{Z}}\phi(\varepsilon n) - \sum_{n\in\mathbb{Z},n\neq 0}\hat{\phi}\left(\frac{n}{\varepsilon}\right)\\ &= \frac{1}{2\pi}\sum_{n\in\mathbb{Z}}\varepsilon\frac{\sin\varepsilon n}{\varepsilon n}\hat{\omega}(\varepsilon n) - \sum_{n\in\mathbb{Z},n\neq 0}\hat{\phi}\left(\frac{n}{\varepsilon}\right).\end{aligned} \tag{4.8}$$

Define $R_\gamma = \frac{2}{\sqrt{1/4+\gamma^2}}$. Now we go to our special case $q = 4$ for this paper. By calculations with computer, we know that when $q = 4$, all positive γ's are greater than 2. Then $\gamma > 2$ implies that $R_\gamma < 1$, so for any $\lambda \geq 0$, there is an X such that

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

[Mon80] shows that

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

Combining these two formulas, we get

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

Note that $\sum_{\gamma>0} R_\gamma^2$ can be computed in terms of $L(1,\chi_1)$ and $L'(1,\chi_1)$. When $q = 3$, we have $\left(\sum_{\gamma>0} R_\gamma^2\right)^{-1} > 0.98$ by some computer calculations, see for example, [RS94, Table 2 on p. 193]. So $\omega[\lambda,\infty) \leq \exp(-1/6(\lambda-2)^2)$ by some computation. Hence for $n \geq 1$ with $\frac{n}{\varepsilon} - 1 \geq 2$, equation (4.7) gives

$$\hat{\phi}\left(\frac{n}{\varepsilon}\right) = \frac{1}{2}\int_{\frac{n}{\varepsilon}-1}^{\frac{n}{\varepsilon}+1} g(u)du \leq \frac{1}{2}\omega\left[\frac{n}{\varepsilon}-1,\infty\right) \leq \frac{1}{2}\exp\left(-\frac{1}{6}\left(\frac{n}{\varepsilon}-3\right)^2\right).$$

By choosing $\varepsilon = 1/20$, we get

$$\sum_{n\in\mathbb{Z},n\neq 0}\hat{\phi}\left(\frac{n}{\varepsilon}\right) < 10^{-20.617\ldots}.$$

This gives

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

where $\varepsilon = 1/20$ and $|\text{error}| < 10^{-20}$. This is an infinite sum. Further investigations are made in [RS94] to truncate it to a finite sum $-C < n\varepsilon < C$, where Rubinstein and Sarnak showed that C can be taken to be 25 for the case $q = 4$ so that the error term becomes negligible. The modified formula with a finite sum for $\delta(P_{4;3,1})$ [RS94, p. 192] gives us $\delta(P_{4;3,1}) = 0.9959\ldots$.

REFERENCES

[Dav94] G Davido. Some experimental results on a generalization of littlewood's theorem. *preprint, Mount Holyoke College*, 1994.

[Dav13] Harold Davenport. *Multiplicative number theory*, volume 74. Springer Science & Business Media, 2013.

[DdS37] L Dirichlet and Beweis des Satzes. da jede unbegrenzte arithmetische progression::: unendlich viele primzahlen enth alt", abh. k onig. *Preuss. Akad*, 34:45–81, 1837.

[GM04] Andrew Granville and Greg Martin. Prime number races, 2004.

[Kla00] Erica Klarreich. Prime time: Will we ever solve the riemann hypothesis?, Nov 2000.

[Lee57] John Leech. Note on the distribution of prime numbers. *Journal of The London Mathematical Society-second Series*, pages 56–58, 1957.

[Lit14] J. E. Littlewood. Distribution des nombres premiers, 1914.

[Mon80] H. L. Montgomery. The zeta function and prime numbers. *Proceedings of the Queen's Number Theory Conference*, pages 13–24, 1980.

[RS94] Michael Rubinstein and Peter Sarnak. Chebyshev's bias. *Experimental Mathematics*, 3(3):173 – 197, 1994.

[SW71] E. Stein and G. Weiss. *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton University Press, 1971.

[Wat48] G. N. Watson. *A Treatise on the Theory of Bessel Functions*, volume (2nd ed.). Cambridge University Press, 1948.

[Win41] Aurel Wintner. On the distribution function of the remainder term of the prime number theorem. *American Journal of Mathematics*, 63(2):233–248, 1941.