RUBINSTEIN AND SARNAK: A TURNING POINT IN COMPARATIVE PRIME NUMBER THEORY

ARNAB BOSE AND REGINALD SIMPSON

ABSTRACT. In this overview, the impact of Rubinstein and Sarnak's 1994 paper on the study of comparative prime number theory is examined. The results of the paper itself are stated and examined, and placed in the historical context of both earlier and later results. The methods used to acheive these results are also examined, from both a mathematical and historical perspective. The impact of the paper on the study of prime number races since its publication is discussed, and it is found to be a key foundation for much contemporary research.

1. INTRODUCTION AND OVERVIEW

This is an overview of the influential and significant 1994 paper "Chebyshev's Bias" by Rubinstein and Sarnak [9]. In their paper, the two authors proved several results concerning prime number races underneath two assumptions about the zeroes of the the Dirichlet L-functions (including the Riemann zeta function): the Generalized Riemann Hypothesis (abbreviated "GRH") and Linear Independence (abbreviated "LI"). GRH asserts that all the nontrivial zeroes[1] of *every* Dirichlet L-function have real part $1/2$. Likewise, LI holds that the imaginary components of the nontrivial zeroes of *every* Dirichlet L-function are linearly independent over the rationals[2]. The systematic manner in which the authors chose to frame and discuss the problem of prime number races in terms of a limiting multivariate distribution for the logarithmic density of certain sets provided a framework for understanding the problem that is still used in contemporary research.

The study of prime number races began with Chebyshev in 1853, who made the observation that it seemed that there were more primes 3 (mod 4) than 1 (mod 4) (see the discussion in [3, p. 227]). This phenomenon was called Chebyshev's Bias. For any x easily calculable before the arrival of digital computers it appeared that $\pi(x;4,1) \leq \pi(x;4,3)$. Likewise, it appeared that $\pi(x;3,1) \leq \pi(x;3,2)$. However, it was shown by Littlewood [6] that there are infinitely many x for which the previous inequalities (individually) hold but also infinitely many x for which the previous inequalities (individually) do not hold. The first x for which $\pi(x;4,1) > \pi(x;4,3)$ was computed by Leech to be $26,861$ [5] and the first x for which $\pi(x;3,1) > \pi(x;3,2)$ was computed by Bays and Hudson to be $608,981,813,029$ [1].

Over the course of the 20th century, a field of study generalizing this problem took form, called comparative prime number theory. Chebyshev's bias was generalized into the concept of biased prime number races. In a prime number race, one considers a modulus $q \geq 3$ and the set of reduced residue classes A_q of q. In the general form chosen by Rubinstein and Sarnak, let $a_1, \ldots, a_r \in A_q$ be distinct, and let $P_{q;a_1,\ldots,a_r}$ be the set of real $x \geq 2$ such that $\pi(x;q,a_1) > \cdots > \pi(x;q,a_r)$. In

[1] A nontrivial zero is a $\rho \in \mathbb{C}$ where $L(\rho,\chi) = 0$ and $0 < \Re\rho < 1$. The zeroes ρ of $L(s,\chi)$ with χ nonprimitive and $\Re\rho = 0$ are not nontrivial zeroes.

[2] In their paper, Rubinstein and Sarnak call this assumption the Grand Simplicity Hypothesis due to the implication that each zero must be simple and $1/2$ must not be a zero. In this overview LI will be used due to becoming in the years since the more standard term.

this terminology, the results of Littlewood [6] determine that the sets $P_{4;3,1}$, $P_{4;1,3}$, $P_{3;2,1}$ and $P_{3;1,2}$ are unbounded.

In the limit, the logarithmic density of primes belonging to a specific reduced residue class modulo q approaches $1/\varphi(q)$ of the total number of primes in accordance with the prime number theorem for arithmetic progressions [8], but the phenomenon where the sets where some reduced residue classes modulo q produce more primes than others over large ranges of x is what gives rise to the notion of a prime number race among the residue classes $a_1, \dots, a_r$. Using the general form chosen here one can extend questions about a bias in a pair of residue classes to questions about biases of entire orderings of the prime number race among residue classes.

To formally study the existence of such biases, Rubinstein and Sarnak proved statements about the density of sets $P_{q;a_1,\dots,a_r}$. Earlier research by Wintner [11] determined some sets of this form have logarithmic densities under certain assumptions. The logarithmic density of a set $P \subseteq [2,\infty)$ is defined to be $\delta(P) = \lim_{X\to\infty} \frac{1}{\log X} \int_{t\in(P\cap[2,X])} \frac{dt}{t}$ with $\overline{\delta}(P)$ and $\underline{\delta}(P)$ being the $\limsup$ and $\liminf$ respectively.

For a residue q where all sets of the form $P_{q;a_1,\dots,a_r}$ have a logarithmic density, one could say a prime number race among the reduced residue classes $a_1, \dots, a_r \in A_q$ for some modulus q is unbiased if and only if for any $b_1, \dots, b_r$ which is a reordering of $1, \dots, r$, with $\delta(P_{q;a_1,\dots,a_r}) = \delta(P_{q;a_{b_1},\dots,a_{b_r}}) = \frac{1}{r!}$. To state what Rubinstein and Sarnak prove for such sets precisely one last concept is needed. Let $\vec{E}_{q;a_1,\dots,a_r} \colon [2,\infty) \to \mathbb{R}^r$ be the vector-valued function

$$\vec{E}_{q;a_1,\dots,a_r}(x) = \frac{\log x}{\sqrt{x}} \langle \varphi(q)\pi(x;q,a_1) - \pi(x), \dots, \varphi(q)\pi(x;q,a_r) - \pi(x) \rangle.$$

With this function defined, it is now possible to state Rubinstein and Sarnak's first theorem.

Theorem 1.1. [9, Theorem 1.1] *Assume GRH. Then $\vec{E}_{q;a_1,\dots,a_r}$ 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 $\mathbb{R}^r$.

Earlier results for special cases do exist (see [10]) but this is the first statement that gives the existence of a distribution to all prime number races. Note that this statement is not quite sufficient to prove the logarithmic densities exist (since characteristic functions are not continuous). However, if the measure $\mu_{q;a_1,\dots,a_r}(x)$ is absolutely continuous, then

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

Conveniently, LI implies that $\mu_{q;a_1,\dots,a_r}(x)$ is absolutely continuous, and thus implies the logarithmic densities of all sets of the form $P_{q;a_1,\dots,a_r}$ exist so our earlier definition of what it means for a prime number race to be unbiased extends to all possible prime number races. The question of which races are biased or unbiased then comes down to an examination of the limiting distributions of $\mu_{q;a_1,\dots,a_r}$ assuming both GRH and LI. Define

$$c(q,a) = -1 + \#\{b \in \mathbb{Z} \colon 0 \le b \le q, b^2 \equiv a \pmod{q}\}, \tag{1.1}$$

which is -1 if a is a quadratic nonresidue modulo q and is positive when a is a quadratic residue modulo q and $q > 3$[3].

[3]Trivially, $c(1,1) = c(2,1) = 0$.

The next result requires making a careful distinction in order to place it into the correct historical context. If $a_1, \dots, a_r \in A_q$ is such that for any reordering $b_1, \dots, b_r \in A_q$ the the limiting distribution $\mu_{q;a_1,\dots,a_r} = \mu_{q;a_{b_1},\dots,a_{b_r}}$, the corresponding prime number race among $a_1, \dots, a_r$ is said to be unbiased in distribution. A prime number race being unbiased in distribution implies being unbiased, but Rubinstein and Sarnak did not prove the converse.

Theorem 1.2. [9, Theorem 1.4] *Assume GRH and LI. The prime number race $(q; a_1, \dots, a_r)$ is unbiased in distribution if and only if $r = 2$ with $c(q, a_1) = c(q, a_2)$ or $r = 3$ where there is some $\rho \neq 1$ such that $\rho^3 \equiv 1 \pmod q$, $a_2 \equiv \rho a_1 \pmod q$, $a_3 \equiv \rho^2 a_1 \pmod q$.*

While Rubinstein and Sarnak suggest that unbiased in distribution is equivalent to the race itself being unbiased, the absence of a definitive proof has served to inspire later researchers. Lamzouri [4] provides an overview of this area of research, and additionally shows that there exist biased r-way races if q is sufficiently large compared to r.

Next, Rubinstein and Sarnak show that despite all other races being biased, there are bounds on how biased they can be.

Theorem 1.3. [9, Theorem 1.5]. *Assume GRH and LI. Then for any fixed r,*

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

In particular, this means that as q grows in size, any prime number races among a fixed number of participants becomes less biased.

For the last theorem, the situation concerning the race of all quadratic nonresidues against all quadratic residues is considered for q being an odd prime power, 2 times an odd prime power, or 4. Define $\pi_N(x; q)$ as the number of primes $\leq x$ which are nonresidues modulo q and $\pi_R(x; q)$ as the number of primes which are residues $\leq x$. Let

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

and let $P_{q;N,R}$ be the set of $x \geq 2$ where $E_{q;N,R}(x) > 0$.

Theorem 1.4. [9, Theorem 1.6] *Assume GRH and LI. Let $\hat{\mu}_{q;N,R}$ be the limiting distribution of*

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

Then $\hat{\mu}_{q;N,R}$ converges in measure to the standard normal distribution as q goes to infinity.

The above results are proven in the next two sections of their work. Afterwards, they conduct numerical investigations and prove the existence of some densities for a a few specific prime number races.

2. Applications of the Generalized Riemann Hypothesis

In this section, a summarized proof of Theorem 1.1 by Rubinstein and Sarnak will be provided. The idea is to show that each entry of $E_{q;a_1,\dots,a_r}$ can be estimated by an approximation using a finite number of zeros ρ of the Dirichlet L-function $L(s, \chi)$. Lastly, by application of Lemma 2.2 the existence of a limiting distribution of the approximation of an entry of $E_{q;a_1,\dots,a_r}$ is guaranteed.

Recalling some notations from Section 1, set

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

Also, set

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

for which

$$\psi(x,\chi)=-\sum_{|\gamma|\le X}\frac{x^{\rho}}{\rho}+O\left(\frac{x\log^2(xX)}{X}+\log x\right) \tag{2.1}$$

by [2]. The sum is over all nontrivial zeros $\rho=\beta+i\gamma$ of the associated $L(s,\chi)$ with $|\gamma|\le X$.

Since GRH is assumed, taking $\beta=\frac{1}{2}$ for all zeros ρ, it follows that

$$\psi(x,\chi)=-\sqrt{x}\sum_{|\gamma|\le X}\frac{x^{i\gamma}}{\frac{1}{2}+i\gamma}+O\Big(\frac{x\log^2(xX)}{X}+\log x\Big). \tag{2.2}$$

Using equation (2.2), the following lemma will be proven.

Lemma 2.1. *As $x\to\infty$,*

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

where the constant $c(a,q)$ is as defined in equation (1.1).

Proof. Start by letting

$$\theta(x;q,a)=\sum_{\substack{p\le x\\ p\equiv a\ (\mathrm{mod}\ q)}}\log p$$

and

$$\psi(x;q,a)=\sum_{\substack{n\le x\\ n\equiv a\ (\mathrm{mod}\ q)}}\Lambda(n).$$

Rewrite $\psi(x;q,a)$ as

$$\psi(x;q,a)=\sum_{\substack{p\le x\\ p\equiv a\ (\mathrm{mod}\ q)}}\log p+\sum_{\substack{p^2\le x\\ p^2\equiv a\ (\mathrm{mod}\ q)}}\log p+\sum_{k\ge 3}\sum_{\substack{p^k\le x\\ p^k\equiv a\ (\mathrm{mod}\ q)}}\log p. \tag{2.3}$$

The first sum in equation (2.3) is $\theta(x;q,a)$. Also, that $\sum_{p\le\sqrt{x}}\log p=\theta(\sqrt{x})$. By the prime number theorem, $\theta(\sqrt{x})\sim\sqrt{x}$. Hence, it follows that

$$\sum_{\substack{p^2\le x\\ p^2\equiv a\ (\mathrm{mod}\ q)}}\log p=\left(\sum_{b^2\equiv a\ (\mathrm{mod}\ q)}1\right)\frac{\sqrt{x}}{\varphi(q)}+O\left(\frac{\sqrt{x}}{\log x}\right),$$

by Dirichlet's density theorem and counting each residue class b which is a solution to $b^2\equiv a\ (\mathrm{mod}\ q)$. The third sum in equation (2.3) can be incorporated in the error term $O(\frac{\sqrt{x}}{\log x})$. Therefore,

$$\psi(x;q,a)=\theta(x;q,a)+\left(\sum_{b^2\equiv a\ (\mathrm{mod}\ q)}1\right)\frac{\sqrt{x}}{\varphi(q)}+O\left(\frac{\sqrt{x}}{\log x}\right). \tag{2.4}$$

Furthermore,

$$\psi(x;q,a) = \frac{1}{\varphi(q)} \sum_{\chi} \bar{\chi}(a)\psi(x,\chi) \tag{2.5}$$

and

$$\pi(x;q,a) = \int_2^x \frac{d\theta(t;q,a)}{\log t}. \tag{2.6}$$

Using equations (2.5) and (2.6) and solving equation (2.4) for $\theta(x;q,a)$, it follows that

$$\pi(x;q,a) = \frac{1}{\varphi(q)} \int_2^x \frac{d\psi(t)}{\log t} + \frac{1}{\varphi(q)} \sum_{\chi \neq \chi_0} \bar{\chi}(a) \int_2^x \frac{d\psi(t,\chi)}{\log t} - \frac{1}{\varphi(q)} \left(\sum_{b^2 \equiv a \pmod q} 1 \right) \frac{\sqrt{x}}{\log x} + O\left(\frac{\sqrt{x}}{\log^2 x} \right). \tag{2.7}$$

Equation (2.7) can be simplified using integration by parts. It is worth pointing out the main part of the calculation. Let

$$G(x,\chi) = \int_2^x \psi(t,\chi)dt.$$

Using equation (2.2), integrating and letting $X \to \infty$, it follows that

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

Using the formula for $N(X,\chi)$ it follows that the series above is absolutely convergent. Therefore, $G(x,\chi) \ll x^{3/2}$, and in conclusion,

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

This finishes the proof. □

Now, computing $E(x,q,a)$ using the estimate (2.1) for $\psi(x,\chi)$ and Lemma 2.1 yields

$$E(x,q,a) = -c(q,a) - \sum_{|\gamma| \leq X} \frac{x^{i\gamma}}{1/2+i\gamma} + \epsilon_a(x,X),$$

where the error term $\epsilon_a(x,X)$ can be made arbitrarily small by making X large enough. Therefore, for $a_1, \ldots, a_r \in A_q$, the quantity $E(x,q,a_j)$ can be estimated by

$$E_j^{(X)}(y) = -c(q,a_j) - \sum_{\substack{\chi \pmod q \\ \chi \neq \chi_0}} \bar{\chi}(a_j) \sum_{|\gamma_\chi| \leq X} \frac{x^{i\gamma_\chi y}}{1/2+i\gamma_\chi},$$

where γ_χ refers to the imaginary components of the nontrivial zeros ρ of $L(s,\chi)$.

Hence, $E_{q;a_1,\ldots,a_r}$ can be estimated by

$$E^{(X)}(y) = \langle E_1^{(X)}(y), E_2^{(X)}(y), \cdots, E_r^{(X)}(y) \rangle$$

Lastly, this section is concluded by stating a lemma (without proof) which guarantees a probability measure for $E^{(X)}(y)$, and hence, a probability measure μ for $E_{q;a_1,\ldots,a_r}$.

Lemma 2.2. *For each X, there is a probability measure ν_X on $\mathbb{R}^r$ such that*

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

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

3. Implications of Linear Independence

In the third section of their paper, Rubinstein and Sarnak prove Theorem 1.4, Theorem 1.3, and Theorem 1.2 in that order. Theorems 1.4 and 1.3 are proven directly. Theorem 1.2 is proven from a proposition concerning the symmetry of the limiting distribution $\mu_{q;a_1,\ldots,a_r}$. However, before starting the proofs of the three theorems, they first obtain a formula for the characteristic function $\hat{\mu}_{q;a_1,\ldots,a_r}$ of the limiting distribution after assuming both GRH and LI.

Rubinstein and Sarnak apply the Kronecker–Weyl Theorem and the same logic as previously found in Lemma 2.3 to show that for $\xi = \langle \xi_1, \ldots, \xi_r\rangle \in \mathbb{R}^r$ (with $\xi \perp \langle 1, \ldots, 1\rangle$ if $r = \varphi(q)$),

$$\hat{\mu}_{q;a_1,\ldots,a_r}(\xi) = \lim_{N\to\infty} \exp\left(i\sum_{m=1}^{r} c(q,a_m)\xi_m\right)\prod_{j=1}^{N}\hat{\mu}_{\gamma_j}(\xi),$$

where $\gamma_1, \ldots, \gamma_N$ are the imaginary parts of the first N nontrivial zeroes of any $L(s,\chi)$ with non-principal χ (mod q) in the upper half-plane, and $\mu_\gamma(\xi)$ is the distribution function of the vector

$$E_\chi(y) = -\left\langle \frac{\bar{\chi}(a_1)e^{i\gamma y}}{\frac{1}{2}+i\gamma} + \frac{\chi(a_1)e^{-i\gamma y}}{\frac{1}{2}-i\gamma}, \ldots, \frac{\bar{\chi}(a_r)e^{i\gamma y}}{\frac{1}{2}+i\gamma} + \frac{\chi(a_r)e^{-i\gamma y}}{\frac{1}{2}-i\gamma}\right\rangle,$$

which is a typical term in the sum $E^{(T)}(y)$ as expressed in Lemma 2.3.

Rubinstein and Sarnak split χ and $\bar{\chi}$ into their real and imaginary parts, expressing the vector as a sum of sines and cosines. They then use the derivative of arcsine to find the density function of $E_\chi(y)$ and after some symbolic manipulation find that

$$\hat{\mu}_\gamma(\xi) = J_0\left(\frac{2\sum_{j=1}^{r}\chi(a_j)\xi_j}{\sqrt{\frac{1}{4}+\gamma^2}}\right),$$

where J_0 is the Bessel function of the first kind. Consequently,

$$\hat{\mu}_{q;a_1,\ldots,a_r}(\xi) = \exp\left(i\sum_{m=1}^{r} c(q,a_m)\xi_m\right)\prod_{\substack{\chi \pmod q\\ \chi\neq\chi_0}}\prod_{\gamma_\chi>0} J_0\left(\frac{2\sum_{j=1}^{r}\chi(a_j)\xi_j}{\sqrt{\frac{1}{4}+\gamma^2}}\right). \tag{3.1}$$

This is the equation of the characteristic function of the limiting multivariate distribution $\mu_{q;a_1,\ldots,a_r}$, which fully determines all properties of the limiting distribution. However, as can be seen from the discussion above, it is only known to be true if one assumes GRH and LI. They also derive two related equations

$$\hat{\mu}_{q;R,N}(\xi) = e^{i\xi} \prod_{\gamma_{\chi_1}>0} J_0\left(\frac{2\xi}{\sqrt{\frac{1}{4}+\gamma^2}}\right), \tag{3.2}$$

$$\big(\pi(x) - \mathrm{Li}(x)\big)\frac{\log x}{\sqrt{x}} = -1 - \sum_{|\gamma|<X} \frac{x^{i\gamma}}{\frac{1}{2}+i\gamma} + O\left(\frac{\sqrt{x}\log(Xx)^2}{X} + \frac{1}{\log x}\right), \tag{3.3}$$

where for the first equation χ_1 is the quadratic character modulo q, and for the second $X \geq 1$ and $x \geq 2$, and γ refers to the imaginary parts of the zeroes of $\zeta(s)$.

Their first application of the above formulas is equation (3.2) to characterize $\hat{\mu}_{q;R,N}(\xi)$ in terms of its symmetries and from that derive various implications about prime number races. Note that in this case q is odd prime power, 2 times an odd prime power, or 4, so χ_1 is a primitive character.

Using the well-known property that J_0 is an even function they determine $\prod_{\gamma_{\chi_1}>0} J_0\Big(\frac{2\xi}{\sqrt{\frac{1}{4}+\gamma^2}}\Big)$ is also even. Thus it follows that the density function $\mu_{q;R,N}(t)$ is symmetric about $t = -1$. From this Rubinstein and Sarnak conclude that $\delta(P_{q;R,N}) = \int_0^\infty d\mu_{q;R,N}(t) < \frac{1}{2}$. Under the assumption of GRH and LI, this shows that quadratic residues lose the race against quadratic nonresidues more often than not, for any modulus q which is an odd prime power, 2 times an odd prime power, or 4.

Rubinstein and Sarnak next move to prove Theorem 1.4. They expand $\hat{\mu}_{q;R,N}(\xi/\sqrt{\log q})$ using (3.2) and power series expansions of J_0 and exp. This yields for $|\xi| \leq A$ with $A > 0$ a fixed constant that

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

where γ refers to the zeroes of $L(s,\chi_1)$. They then use an argument of Littlewood's to express the sum over the zeroes in terms of $\frac{L'}{L}(1,\chi_1)$, and then estimate the size of the error term, ultimately yielding that

$$\lim_{q\to\infty} \log \mu_{q;R,N}\left(\frac{\xi}{\sqrt{\log q}}\right) = -\frac{1}{2}\xi^2$$

uniformly for $|\xi| \leq A$. Rubinstein and Sarnak then apply Lévy's Theorem to this result to show the measures $\tilde{\mu}_{q;N,R}$ converge in measure to the standard multivariate normal distribution as q goes to infinity. An immediate consequence is that $\delta(P_{q;N,R}) \to \frac{1}{2}$ as $q \to \infty$, which proves Theorem 1.4.

Theorem 1.3 is proven by similar means, but more work is required to deal with nonprimitive characters. This is mostly handled by elementary methods (such as the observation that the number $\beta(q)$ of primitive characters a modulus is such that under Dirichlet convolution $\beta * 1 = \varphi$) and some basic properties relating $L(s,\chi)$ and $L(s,\chi^*)$.

Ultimately, Rubinstein and Sarnak show by an application of Lévy's Theorem that $\tilde{\mu}_{q;a_1,\ldots,a_r}$ converges to the standard multivariate Gaussian on $\mathbb{R}^r$, where $\tilde{\mu}_{q;a_1,\ldots,a_r}$ is the measure on $\mathbb{R}^r$ with Fourier transform $\hat{\mu}_{q;a_1,\ldots,a_r}\big(\frac{\xi}{\sqrt{\varphi(q)\log q}}\big)$. They then deduce that for any $D \subseteq \mathbb{R}^r$, and for any permutation σ of the r coordinates, that

$$\lim_{q\to\infty} |\tilde{\mu}_{q;a_1,\ldots,a_r}(D) - \tilde{\mu}_{q;a_1,\ldots,a_r}(D^\sigma)| = 0.$$

From this equation it is trivial to show that each ordering of the prime number race result must converge to a fair one, which requires the limiting density $\lim_{q\to\infty}\delta(P_{q;a_1,\ldots,a_r}) = \frac{1}{r!}$, proving Theorem 1.3.

Next, Rubinstein and Sarnak prove the following proposition, which will ultimately be used to prove Theorem 1.2.

Proposition 3.1. [9, Proposition 3.1] *The density function $\mu_{q;a_1,\dots,a_r}$ is symmetric in $(x_1,\dots,x_r)$ if and only if*

- $r = 2$ *and* $c(q, a_1) = c(q, a_2)$, *or*
- $r = 3$ *and there exists* $\rho \neq 1$ *such that* $\rho^3 \equiv 1 \pmod{q}$, $a_1\rho \equiv a_2 \pmod{q}$, $a_1\rho^2 = a_3 \pmod{q}$.

The proof of the proposition is done in two steps. The first is to observe that by (3.1) that the term $\exp\left(i\sum_{m=1}^{r} c(q,a_m)\xi_m\right)$ yields by taking the derivative of $\hat{\mu}$ that the distribution of $\mu_{q;a_1,\dots,a_r}$ has mean value $(c(q;a_1),\dots,c(q;a_r))$. In order for the distribution to be symmetric, its mean value must be invariant under permutation, which requires $c(q;a_1) = \dots = c(q;a_r)$.

That each $c(a,a_j)$ be equal is thus a necessary condition of the density function being symmetric. Thus to prove the proposition, the symmetries of the terms inside the Bessel function J_0 must be characterized. This is done by the following Lemma:

Lemma 3.2. [9, Lemma 3.2] *The expression $B_\chi(\xi_1,\dots,\xi_r) = \left|\sum_{j=1}^{r}\chi(a_j)\xi_j\right|$ is symmetric in $(\eta_1,\dots,\eta_r)$ if and only if one of the two conditions of the above proposition hold.*

When $r = 2$, simply observe

$$\begin{aligned} B_\chi(\xi_1,\xi_2) = \left|\chi(a_1)\xi_1 + \chi(a_2)\xi_2\right| &= \left|\bar{\chi}(a_1)\bar{\chi}(a_2)(\chi(a_1)\xi_1 + \chi(a_2)\xi_2)\right| \\ &= |\chi(a_2)\xi_1 + \chi(a_1)\xi_2| = B_\chi(\eta_2,\eta_1). \end{aligned}$$

When $r = 3$ and ρ as described above exists, if $\tau = \chi(\rho)$, it follows that since $|\tau| = 1$,

$$B_\chi(\xi_1,\xi_2,\xi_3) = \left|\xi_1 + \tau\xi_2 + \tau^2\xi_3\right| = \left|\tau^k(\xi_1 + \tau\xi_2 + \tau^2\xi_3)\right|$$

which freely permutes the sum in a cycle, since $\tau^3 = \chi(\tau^3) = \chi(1) = 1$. By taking the conjugate of the interior of the sum, it becomes clear that one can exchange the positions of ξ_2 and ξ_3. Combined, this produces every possible permutation, showing $B_\chi(\xi_1,\xi_2,\xi_3)$ is symmetric.

An argument showing the converse is then demonstrated for $r = 3$. The case where $r \geq 4$ is excluded because of this converse argument, since it will imply the above relation for any three residues, which contradicts the uniqueness of a_1, a_2, a_3.

The remainder of this section formally proves the rest of the proposition, demonstrating that if the condition in the lemma fails, a contradiction results if the limiting distribution is symmetric.

4. A Survey of the Numerical Investigations

In the fourth section of their paper, Rubinstein and Sarnak give a detailed description of the computations that lead to the following numbers.

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

This involves evaluating the integrals $\delta(P_{q;N,R}) = \int_{-\infty}^{1} d\omega_{q;R,N}(t)$ and $\delta(P_1^{\text{comp}}) = \int_{-\infty}^{1} d\omega_1(t)$.

To show how these integrals were evaluated, some notation must first be defined. Let $f_{q;N,R}(t)$ and $f_1(t)$ be the density functions of $\mu_{q;N,R}$ and μ_1 respectively. Here, the authors observe that it is more convenient to work with the distribution ω whose density function is given by $g(t) := f(t-1)$, with Fourier transform

$$\hat{\omega}(\xi) = \prod_{\gamma>0} J_0\left(\frac{2\xi}{\sqrt{1/4+\gamma^2}}\right). \tag{4.1}$$

Since $g_{q;N,R}$ is symmetric about 0,

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

where the second equality above is obtained using the inversion formula for characteristic functions. The evaluation of the integrals in the above expression essentially involves three steps. Firstly, the integral is replaced with a sum. The second step involves replacing the infinite summation by a large but finite sum. Lastly, the third step involves replacing the infinite product for $\hat{\omega}$ by a finite product. The first step will be explained in detail and an outline of the second and the third steps will be given.

Set

$$\phi(u) = \frac{1}{2\pi}\frac{\sin u}{u}\hat{\omega}(u) \text{ and } \hat{\phi}(x) = \frac{1}{2}\big(\chi_{[-1,1]} * g\big)(x) = \frac{1}{2}\int_{x-1}^{x+1} g(u)\,du = \frac{1}{2}\int_{x-1}^{x+1} d\omega(u). \tag{4.3}$$

By applying the Poisson summation formula to equation (**??**) to equations (4.3), and using a well-known bound on the Bessel function J_0, Rubinstein and Sarnak determine that

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\sin u}{u}\hat{\omega}(u)du = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}}\epsilon\frac{\sin(n\epsilon)}{n\epsilon}\hat{\omega}(n\epsilon) - \sum_{\substack{n\in\mathbb{Z}\\ n\neq 0}}\hat{\phi}(n/\epsilon) \tag{4.4}$$

Hence, the error of replacing the integral in equation (4.2) with the sum as in equation (4.4) can be estimated if an upper bound for $\hat{\phi}(n/\epsilon)$ is found. Using an upper bound for ω from [7] and computed values for $L(1,\chi_1)$, $L'(1,\chi_1)$, and $\sum_{\gamma_{\chi_1}>0}\frac{1}{\gamma_{\chi_1}^2+1/4}$ collected in Table 2 [9, p. 193], for $q = 1,3,4,5,7,11,13$ and $\lambda \geq 2$ the authors conclude that $\omega[\lambda,\infty) \leq \exp\big(-\frac{1}{6}(\lambda-2)^2\big)$.

Hence, for $n \geq 1$ and $n/\epsilon - 1 \geq 2$, it follows that

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

Using the fact that $\hat{\phi}$ is symmetric about 0 and choosing $\epsilon = \frac{1}{20}$, it follows

$$\sum_{\substack{n\in\mathbb{Z}\\ n\neq 0}}\hat{\phi}(n/\epsilon) = 2\sum_{n=1}^{\infty}\hat{\phi}(n/\epsilon) \leq \sum_{n=1}^{\infty}\exp(-\frac{1}{6}(20n-3)^2) < 2\exp(-\frac{1}{6}17^2) = 10^{-20.617\cdots}. \tag{4.5}$$

Therefore, using equation (4.5) in equation (4.4), equation (4.2) becomes

$$\delta(P_{q;N,R}) = \frac{1}{2} + \frac{1}{2\pi}\sum_{n\in\mathbb{Z}}\epsilon\frac{\sin(n\epsilon)}{n\epsilon}\hat{\omega}(n\epsilon) + E, \tag{4.6}$$

where $\epsilon = \frac{1}{20}$ and the error term E satisfies $|E| < 10^{-20}$.

The second step involves replacing the sum in the second tern of equation (4.6) above by a finite sum over $-C < n\epsilon < C$, where C is chosen to be sufficiently large such that the contribution of the tail ends of the sum is small. Subsequently, the authors derive an upper bound for $\frac{1}{2\pi}\sum_{n\in\mathbb{Z}} \epsilon\frac{\sin(n\epsilon)}{n\epsilon}\hat{\omega}(n\epsilon)$ and together with the calculations in Table 1 [9], they conclude that

$$\delta(P_{q;N,R}) = \frac{1}{2\pi}\sum_{-25\leq n\epsilon 25} \epsilon\frac{\sin(n\epsilon)}{n\epsilon}\prod_{\gamma_\chi>0} J_0\left(\frac{n\epsilon}{\sqrt{1/4+\gamma_{\chi_1}^2}}\right) + 1/2 + \text{error term.} \tag{4.7}$$

Lastly, the infinite product in equation (4.7) is replaced by a finite product and an approximating polynomial. This leads to

$$\delta(P_{q;N,R}) = \frac{1}{2\pi}\sum_{-25\leq n\epsilon\leq 25} \epsilon\frac{\sin(n\epsilon)}{n\epsilon}(1+b_1(n\epsilon)^2)\prod_{0<\gamma\leq 9999} J_0\left(\frac{2n\epsilon}{\sqrt{1/4+\gamma^2}}\right) + 1/2 + \text{error}$$

where

$$b_1 = T_1(0) + \sum_{0<\gamma\leq 9999}\frac{1}{1/4+\gamma^2}.$$

Therefore, for different values of q, the corresponding δ values are obtained as was mentioned in the beginning of this section. Similar treatment and calculations can be done to derive the corresponding formula for $\delta(P_1^{\text{comp}})$ and $\delta(P_1^{\text{comp}}) = 0.99999973\cdots$ is obtained.

References

[1] Carter Bays and Richard H. Hudson. Details of the first region of integers x with $\pi_{3,2}(x) < \pi_{3,1}(x)$. *Math. Comp.*, 32(142):571–576, 1978.

[2] Harold Davenport. *Multiplicative number theory*, volume 74 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2000. Revised and with a preface by Hugh L. Montgomery.

[3] J. Kaczorowski. The boundary values of generalized Dirichlet series and a problem of Chebyshev. Number 209, pages 14, 227–235. 1992. Journées Arithmétiques, 1991 (Geneva).

[4] Youness Lamzouri. Prime number races with three or more competitors. *Math. Ann.*, 356(3):1117–1162, 2013.

[5] John Leech. Note on the distribution of prime numbers. *J. London Math. Soc.*, 32:56–58, 1957.

[6] J.E. Littlewood. Distribution des nombres premiers. *C. R. Acad. Sci. Paris*, 158:1869–1862, 1914.

[7] H. L. Montgomery. The zeta function and prime numbers. In *Proceedings of the Queen's Number Theory Conference, 1979 (Kingston, Ont., 1979)*, volume 54 of *Queen's Papers in Pure and Appl. Math.*, pages 14–24. Queen's Univ., Kingston, Ont., 1980.

[8] H. L. Montgomery and R. C. Vaughan. *Multiplicative Number Theory I. Classical Theory*. Cambridge University Press, 2007.

[9] Michael Rubinstein and Peter Sarnak. Chebyshev's bias. *Experiment. Math.*, 3(3):173–197, 1994.

[10] Aurel Wintner. Asymptotic distributions and infinite convolutions, 1938. Notes distributed by the Institute for Advanced Study (Princeton).

[11] Aurel Wintner. On the distribution function of the remainder term of the prime number theorem. *Amer. J. Math.*, 63:233–248, 1941.