CHEBYSHEV'S BIAS FOR PRODUCTS OF TWO PRIMES

KEVIN FORD AND JASON SNEED

ABSTRACT. Under two assumptions, we determine the distribution of the difference between two functions each counting the numbers $\leqslant x$ that are in a given arithmetic progression modulo q and the product of two primes. The two assumptions are (i) the Extended Riemann Hypothesis for Dirichlet L-functions modulo q, and (ii) that the imaginary parts of the nontrivial zeros of these L-functions are linearly independent over the rationals. Our results are analogs of similar results proved for primes in arithmetic progressions by Rubinstein and Sarnak.

1. INTRODUCTION

1.1. **Prime number races.** Let $\pi(x; q, a)$ denote the number of primes in the progression $a \bmod q$. For fixed q, the functions $\pi(x; q, a)$ (for $a \in A_q$, the set of residues coprime to q) all satisfy

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

where φ is Euler's totient function [Da]. There are, however, curious inequities. For example $\pi(x; 4, 3) \geqslant \pi(x; 4, 1)$ seems to hold for most x, an observation of Chebyshev from 1853 [Ch]. In fact, $\pi(x; 4, 3) < \pi(x; 4, 1)$ for the first time at $x = 26,861$ [Le]. More generally, one can ask various questions about the behavior of

$$\Delta(x; q, a, b) := \pi(x; q, a) - \pi(x; q, b) \tag{1.2}$$

for distinct $a, b \in A_q$. Does $\Delta(x; q, a, b)$ change sign infinitely often? Where is the first sign change? How many sign changes with $x \leqslant X$? What are the extreme values of $\Delta(x; q, a, b)$? Such questions are colloquially known as *prime race problems*, and were studied extensively by Knapowski and Turán in a series of papers beginning with [KT]. See the survey articles [FK] and [GM] and references therein for an introduction to the subject and summary of major findings. Properties of Dirichlet L-functions lie at the heart of such investigations.

Despite the tendency for the function $\Delta(x; 4, 3, 1)$ to be negative, Littlewood [Li] showed that it changes sign infinitely often. Similar results have been proved for other q, a, b (see [S] and references therein). Still, in light of Chebyshev's observation, we can ask how frequently $\Delta(x; q, a, b)$ is positive and how often it is negative. These questions are best addressed in the context of *logarithmic density*. A set S of positive integers has logarithmic density

$$\delta(S) = \lim_{x\to\infty} \frac{1}{\log x} \sum_{\substack{n \leqslant x \\ n \in S}} \frac{1}{n}$$

provided the limit exists. Let $\delta(q, a, b) = \delta(P(q, a, b))$, where $P(q, a, b)$ is the set of integers n with $\Delta(n; q, a, b) > 0$. In 1994, Rubinstein and Sarnak [RS] showed that $\delta(q; a, b)$ exists, assuming two hypotheses (i) the Extended Riemann Hypothesis for Dirichlet L-functions modulo q (ERH_q), and (ii) the imaginary parts of zeros of each Dirichlet L-function are linearly independent over the rationals (GSH_q - Grand Simplicity Hypothesis). The authors also gave methods to accurately estimate the "bias", for example showing that $\delta(4; 3, 1) \approx 0.996$ in Chebyshev's case. More generally, $\delta(q; a, b) = \frac{1}{2}$ when a and b are either

---

Date: January 6, 2010.
2000 Mathematics Subject Classification:11M06, 11N13, 11N25.
The research of K. F. was supported in part by National Science Foundation grants DMS-0555367 and DMS-0901339.

both quadratic residues modulo q or both quadratic nonresidues (unbiased prime races), but $\delta(q; a, b) > \frac{1}{2}$ whenever a is a quadratic non-residue and b is a quadratic residue. A bit later we will discuss the reasons behind these phenomena. Sharp asymptotics for $\delta(q; a, b)$ have recently been given by Fiorilli and Martin [FM], which explain other properties of these densities.

1.2. **Quasi-prime races.** In this paper we develop a parallel theory for comparison of functions $\pi_2(x; q, a)$, the number of integers $\leqslant x$ which are in the progression $a \bmod q$ and which are the product of two primes p_1p_2 ($p_1 = p_2$ allowed). Put

$$\Delta_2(x; q, a, b) := \pi_2(x; q, a) - \pi_2(x; q, b),$$

let $P_2(q, a, b)$ be the set of integers n with $\Delta_2(n; q, a, b) > 0$, and set $\delta_2(q, a, b) = \delta(P_2(q, a, b))$. The table below shows all such quasi-primes up to 100 grouped in residue classes modulo 4.

$pq \equiv 1 \pmod 4$	$pq \equiv 3 \pmod 4$
9	15
21	35
25	39
33	51
49	55
57	87
65	91
69	95
77	
85	
93	

Observe that $\Delta_2(x; 4, 3, 1) \leqslant 0$ for $x \leqslant 100$, and in fact the smallest x with $\Delta_2(x; 4, 3, 1) > 0$ is $x = 26747$ (amazingly close to the first sign change of $\Delta(x; 4, 3, 1)$). Some years ago Richard Hudson conjectured that the bias for products of two primes is always reversed from that of primes; i.e., $\delta_2(q; a, b) < \frac{1}{2}$ when a is a quadratic non-residue modulo q and b is a quadratic residue. Under the same assumptions as [RS], namely ERH_q and GSH_q, we confirm Hudson's conjecture and also show that the bias is less pronounced.

Theorem 1. *Let a, b be distinct elements of A_q. Assuming ERH_q and GSH_q, $\delta_2(q; a, b)$ exists. Moreover, if a and b are both quadratic residues modulo q or both quadratic non-residues, then $\delta_2(q; a, b) = \frac{1}{2}$. Otherwise, if a is a quadratic nonresidue and b is a quadratic residue, then*

$$1 - \delta(q; a, b) < \delta_2(q; a, b) < \frac{1}{2}.$$

We can accurately estimate $\delta_2(q; a, b)$ borrowing methods from [RS, §4]. In particular we have

$$\delta_2(4; 3, 1) \approx 0.10572.$$

We deduce Theorem 1 by connecting the distribution of $\Delta_2(x; q, a, b)$ with the distribution of $\Delta(x; q, a, b)$. Although the relationship is "simple", there is no elementary way to derive it, say by writing

$$\pi_2(x; q, a) = \frac{1}{2} \sum_{p \leqslant x} \pi\left(\frac{x}{p}; q, ap^{-1} \mod q\right) + \frac{1}{2} \sum_{\substack{p \leqslant \sqrt{x} \\ p^2 \equiv a \pmod q}} 1.$$

In particular, our result depends strongly on the assumption that the zeros of the L-functions modulo q have only simple zeros. Let $N(q, a)$ be the number of $x \in A_q$ with $x^2 \equiv a \pmod q$, and let $C(q)$ be the set of nonprincipal Dirichlet characters modulo q.

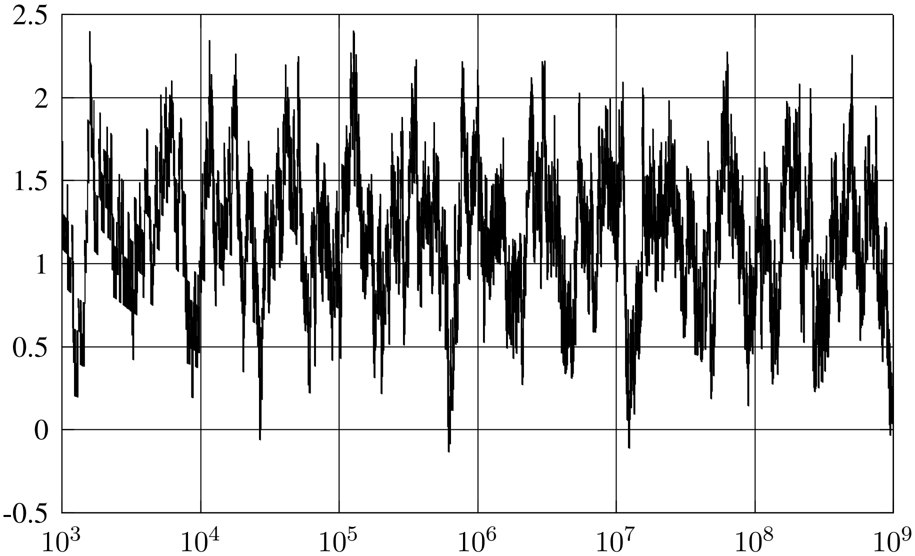

FIGURE 1. $\frac{\log x}{\sqrt{x}}\Delta(x;4,3,1)$

Theorem 2. *Assume ERH$_q$ and for each $\chi \in C(q)$, $L(\frac12,\chi)\ne 0$ and the zeros of $L(s,\chi)$ are simple. Then*

$$\frac{\Delta_2(x;q,a,b)\log x}{\sqrt{x}\log\log x} = \frac{N(q,b)-N(q,a)}{2\phi(q)} - \frac{\log x}{\sqrt{x}}\Delta(x;q,a,b) + \Sigma(x;q,a,b),$$

where $\frac{1}{Y}\int_1^Y |\Sigma(e^y;q,a,b)|^2\,dy = o(1)$ as $Y\to\infty$.

The expression for Δ_2 given in Theorem 2 must be modified if some $L(s,\chi)$ has multiple zeros; see §3 for details.

Figures 1,2 and 3 show graphs corresponding to $(q,a,b)=(4,3,1)$, plotted on a logarithmic scale from $x=10^3$ to $x=10^9$. While $\Sigma(x;4,3,1)$ appears to be oscillating around -0.2, this is caused by some terms in $\Sigma(x;4,3,1)$ of order $1/\log\log x$, and $\log\log 10^9 \approx 3.03$. By Theorem 2, $\Sigma(x;4,3,1)$ will (assuming ERH$_4$ and GSH$_4$) eventually settle down to oscillating about 0.

It is not immediate that Theorem 1 follows from Theorem 2. One first needs more precise information about the distribution of $\Delta(x;q,a,b)$ from [RS].

Theorem RS. [RS, §1] *Assume ERH$_q$ and GSH$_q$. For any distinct $a,b\in A_q$, the function*

$$\frac{u\Delta(e^u;q,a,b)}{e^{u/2}} \tag{1.3}$$

has a probabilistic distribution. This distribution (i) has mean $(N(q,b)-N(q,a))/\phi(q)$, (ii) is symmetric with respect to its mean, and (iii) has a continuous density function.

Assume a is a quadratic nonresidue modulo q and b is a quadratic residue. Then $N(q,b)-N(q,a)>0$. Let f be the density function for the distribution of (1.3), that is,

$$f(t) = \frac{d}{dt}\lim_{U\to\infty}\frac{1}{U}\text{meas}\{0\leqslant u\leqslant U : ue^{-u/2}\Delta(e^u;q,a,b)\leqslant t\}.$$

We see from Theorem RS that

$$\delta(q,a,b) = \int_0^\infty f(t)\,dt > \frac12$$

and from Theorem 2 that

$$\delta_2(q,a,b) = \int_{-\infty}^{\frac{N(q,b)-N(q,a)}{2\phi(q)}} f(t)\,dt,$$

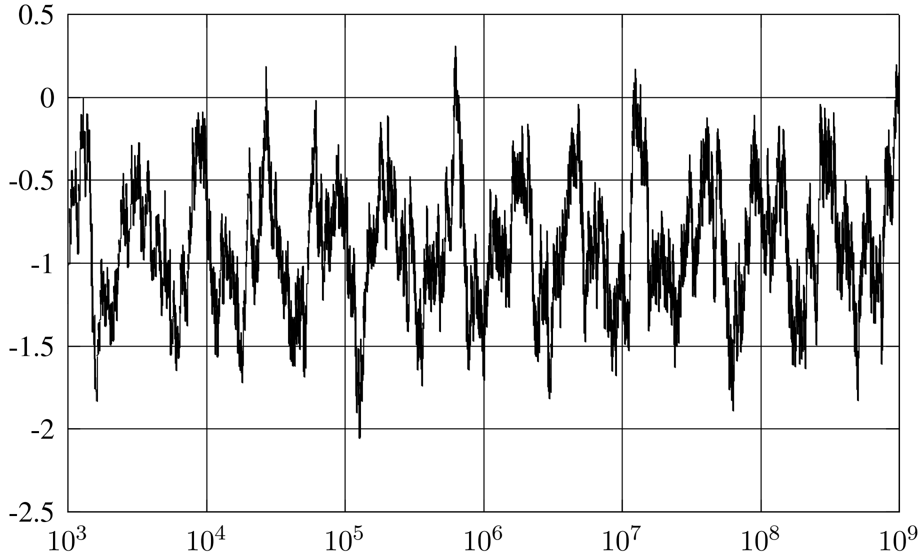

FIGURE 2. $\frac{\log x}{\sqrt{x}\log\log x}\Delta_2(x;4,3,1)$

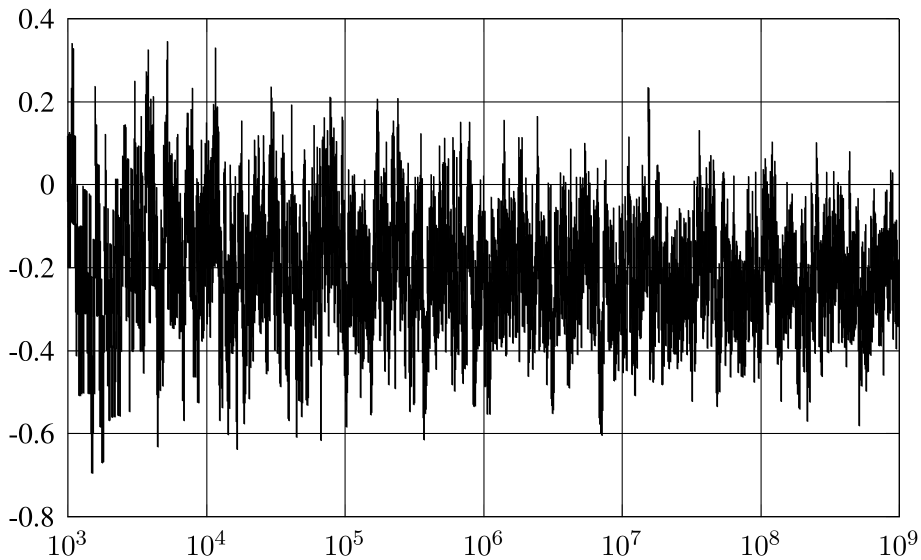

FIGURE 3. $\Sigma(x;4,3,1)$

from which Theorem 1 follows.

Theorem 2 also determines the joint distribution of any vector function

$$\frac{u}{e^{u/2}\log u}\left(\Delta_2(e^u;q,a_1,b_1),\dots,\Delta_2(e^u;q,a_r,b_r)\right). \tag{1.4}$$

Theorem 3. *If $f(x_1,\dots,x_r)$ is the density function of*

$$\frac{u}{e^{u/2}}\left(\Delta(e^u;q,a_1,b_1),\dots,\Delta(e^u;q,a_r,b_r)\right),$$

then the joint density function of (1.4) *is*

$$f\left(\frac{N(q,b_1)-N(q,a_1)}{2\phi(q)}-x_1,\dots,\frac{N(q,b_r)-N(q,a_r)}{2\phi(q)}-x_r\right).$$

1.3. **Origin of Chebyshev's bias.** From an analytic point of view (L-functions), the weighted sum

$$\Delta^*(x;q,a,b)=\sum_{\substack{n\leqslant x\\ n\equiv a \bmod q}}\Lambda(n)-\sum_{\substack{n\leqslant x\\ n\equiv b \bmod q}}\Lambda(n), \tag{1.5}$$

where Λ is the von Mangoldt function, is more natural than (1.2). Expressing $\Delta^*(x;q,a,b)$ in terms of sums over zeros of L-functions in the standard way (§19 of [Da]), we obtain, on ERH_q,

$$e^{-u/2}\phi(q)\Delta^*(e^u;q,a,b)=-\sum_{\chi\in C(q)}(\overline{\chi}(a)-\overline{\chi}(b))\sum_{\gamma}\frac{e^{i\gamma u}}{1/2+i\gamma}+O(u^2e^{-u/2}),$$

where γ runs over imaginary parts of nontrivial zeros of $L(s,\chi)$ (counted with multiplicity). Hypothesis GSH_q implies, in particular, that $L(1/2,\chi)\neq 0$. Each summand $e^{i\gamma u}/(1/2+i\gamma)$ is thus a harmonic with mean zero as $u\to\infty$, and GSH_q implies that the harmonics behave independently. Hence, we expect that $e^{-u/2}\phi(q)\Delta^*(e^u;q,a,b)$ will behave like a mean zero random variable. On the other hand, the right side of (1.5) contains not only terms corresponding to prime n but terms corresponding to powers of primes. Applying the prime number theorem for arithmetic progressions (1.1) to the terms $n=p^2$ in (1.5) gives

$$\Delta^*(x;q,a,b)=\sum_{\substack{p\leqslant x\\ p\equiv a \bmod q}}\log p-\sum_{\substack{p\leqslant x\\ p\equiv b \bmod q}}\log p+\frac{x^{1/2}}{\phi(q)}\left(N(q,a)-N(q,b)\right)+O(x^{1/3}).$$

Hence, on ERH_q and GSH_q, we expect the expression

$$\frac{1}{\sqrt{x}}\Bigg(\sum_{\substack{p\leqslant x\\ p\equiv a \bmod q}}\log p-\sum_{\substack{p\leqslant x\\ p\equiv b \bmod q}}\log p\Bigg) \tag{1.6}$$

to behave like a random variable with mean $(N(q,b)-N(q,a))/\phi(q)$. Finally, the distribution of $\Delta(x;q,a,b)$ is obtained from the distribution of (1.6) and partial summation.

1.4. **Analyzing** $\Delta_2(x;q,a,b)$**.** A natural analog of $\Delta^*(x;q,a,b)$ is

$$\sum_{\substack{mn\leqslant x\\ mn\equiv a \bmod q}}\Lambda(m)\Lambda(n)-\sum_{\substack{mn\leqslant x\\ mn\equiv b \bmod q}}\Lambda(m)\Lambda(n). \tag{1.7}$$

As with $\Delta^*(x;q,a,b)$, the expression in (1.7) can be easily written as a sum over zeros of L-functions plus a small error. The main problem now is that the principal summands, namely $\log p_1\log p_2$ for primes p_1,p_2, are very irregular as a function of p_1p_2, and thus estimates for $\Delta_2(x;q,a,b)$ cannot be recovered by partial summation. We get around this problem using a double integration, a method which goes back to Landau [La, §88]. We have

$$\begin{aligned}\Delta_2(x;q,a,b)&=\frac{1}{\phi(q)}\sum_{\chi\in C(q)}(\overline{\chi}(a)-\overline{\chi}(b))\sum_{\substack{n=p_1p_2\leqslant x\\ p_1\leqslant p_2}}\chi(n)\\ &=\frac{1}{2\phi(q)}\sum_{\chi\in C(q)}(\overline{\chi}(a)-\overline{\chi}(b))\int_0^\infty\!\!\int_0^\infty G(x,u,v;\chi)\,du\,dv+O\left(\frac{\sqrt{x}}{\log x}\right),\end{aligned} \tag{1.8}$$

where

$$G(x,u,v;\chi)=\sum_{p_1p_2\leqslant x}\frac{\chi(p_1p_2)\log p_1\log p_2}{p_1^u p_2^v}. \tag{1.9}$$

The related functions

$$G^*(x,u,v;\chi)=\sum_{mn\leqslant x}\frac{\chi(mn)\Lambda(m)\Lambda(n)}{m^u n^v}$$

are more "natural" from an analytic point of view, being easily expressed in terms of zeros of Dirichlet L-functions. By the reasoning of the previous subsection, each $G^*(x,u,v;\chi)$ is expected to be unbiased, the bias in $\Delta_2(x;q,a,b)$ originating from the summands in $G^*(x,u,v;\chi)$ where m is not prime or n is not prime.

1.5. **A heuristic argument for the bias in $\Delta_2(x;q,a,b)$.** We conclude this introduction with a heuristic evaluation of the bias in $\Delta_2(x;q,a,b)$, which originates from the difference between functions $G(x;u,v;\chi)$ and $G^*(x,u,v;\chi)$. For simplicity of exposition, we'll concentrate on the special case $(q,a,b)=(4,3,1)$. In this case, the bias arises from terms $p_1p_2^2$ and $p_1^2p_2^2$ which appear in $G^*(x;u,v;\chi)$ but not in $G(x,u,v;\chi)$. Let χ be the non-principal character modulo 4, so that

$$\frac12\int_0^\infty\int_0^\infty\big(G^*(x,u,v;\chi)-G(x,u,v;\chi)\big)\,du\,dv=\frac12\sum_{\substack{p_1^ap_2^b\leqslant x\\ \max(a,b)\geqslant 2}}\frac{\chi(p_1^ap_2^b)}{ab}.$$

There are $O(x^{1/2}/\log x)$ terms with $\min(a,b)\geqslant 2$ and $\max(a,b)\geqslant 3$. By the prime number theorem and partial summation,

$$\frac12\sum_{p_1^2p_2^2\leqslant x}\frac14=\frac18\sum_{p\leqslant\sqrt{x}}\pi\left(\sqrt{x/p^2}\right)\sim\frac{x^{1/2}\log\log x}{2\log x}.$$

Thus,

$$\Delta_2(x;4,3,1)=-\frac12\sum_{mn\leqslant x}\frac{\chi(mn)\Lambda(m)\Lambda(n)}{\log m\log n}-\left(\sum_{k=2}^\infty\frac1k\sum_{p_1^k\leqslant x}\chi(p_1^k)\Delta(x/p_1^k;4,3,1)\right)$$
$$+\left(\frac12+o(1)\right)\frac{x^{1/2}\log\log x}{\log x}.$$

By Theorem RS, $\Delta(y;4,3,1)=y^{1/2}/\log y+E(y)$, where $E(y)$ oscillates with mean 0. Thus,

$$\sum_{k=2}^\infty\frac1k\sum_{p_1^k\leqslant x}\chi(p_1^k)\Delta(x/p_1^k;4,3,1)=\sum_{k=2}^\infty\frac2k\sum_{p_1^k\leqslant x}\chi(p_1^k)\frac{\sqrt{x/p_1^k}}{\log(x/p_1^k)}+E'(x),$$

where $E'(x)$ is expected to oscillate with mean zero. The $k=2$ terms are

$$\sum_{p_1^2\leqslant x}\frac{\sqrt{x/p_1^2}}{\log(x/p_1^2)}\sim\frac{\sqrt{x}\log\log x}{\log x},$$

while the terms corresponding to $k\geqslant 3$ contribute

$$\ll\sum_{k=3}^\infty\frac1k\sum_{p_1^k\leqslant x}\frac{\sqrt{x/p_1^k}}{\log(x/p_1^k)}\ll\frac{\sqrt{x}}{\log x}.$$

Thus, we find that

$$\Delta_2(x;4,3,1)=-\frac12\sum_{mn\leqslant x}\frac{\chi(mn)\Lambda(m)\Lambda(n)}{\log m\log n}-\left(\frac12+o(1)\right)\frac{x^{1/2}\log\log x}{\log x}+E'(x).$$

1.6. **Further problems.** It is natural to consider the distribution, in arithmetic progressions, of numbers composed of exactly k prime factors, where $k \geqslant 3$ is fixed. As with the cases $k = 1$ and $k = 2$, we expect there to be no bias if we count all numbers $p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$ with weight $(a_1 \cdots a_k)^{-1}$. If, however, we count terms which are the product of precisely k primes (that is, numbers $p_1^{a_1} \cdots p_j^{a_j}$ with $a_1 + \cdots + a_j = k$), then there will be a bias. Hudson has conjectured that the bias will be in the same direction as for primes when k is odd, and in the opposite direction for even k. We conjecture that, in addition, the bias becomes less pronounced as k increases.

2. Preliminaries

With χ fixed, the letter γ, with or without subscripts, denotes the imaginary part of a zero of $L(s, \chi)$ inside the critical strip. In sums over γ, each term appears with its multiplicity $m(\gamma)$ unless we specify that we sum over distinct γ. Constants implied by $O-$ and $\ll -$symbols depend only on χ (and hence, on q) unless additional dependence is indicated with a subscript. Let

$$A(\chi) = \begin{cases} 1 & \chi^2 = \chi_0 \\ 0 & \text{else} \end{cases},$$

where χ_0 is the principal character modulo q. That is, $A(\chi) = 1$ if and only if χ is a real character. For $\chi \in C(q)$, define

$$F(s, \chi) = \sum_p \frac{\chi(p) \log p}{p^s}.$$

The following estimates are standard; see e.g. [Da, §15,16].

Lemma 2.1. *Let $\chi \in C(q)$, assume ERH$_q$ and fix $c > \frac{1}{3}$. Then $F(s, \chi) = -\frac{L'}{L}(s, \chi) + A(\chi)\frac{\zeta'}{\zeta}(2s) + H(s, \chi)$, where $H(s, \chi)$ is analytic and uniformly bounded in the half-plane $\Re s \geqslant c$.*

Lemma 2.2. *Let χ be a Dirichlet character modulo q. Let $N(T, \chi)$ denote the number of zeros of $L(s, \chi)$ with $0 < \Re s < 1$ and $|\Im s| < T$. Then*

(1) $N(T, \chi) = O(T \log(qT))$ *for* $T \geqslant 1$.
(2) $N(T, \chi) - N(T-1, \chi) = O(\log(qT))$ *for* $T \geqslant 1$.
(3) *Uniformly for $s = \sigma + it$ and $\sigma \geqslant -1$,*

$$\frac{L'(s, \chi)}{L(s, \chi)} = \sum_{|\gamma - t| < 1} \frac{1}{s - \rho} + O(\log q(|t| + 2)).$$

(4) $-\frac{\zeta'}{\zeta}(\sigma) = \frac{1}{\sigma - 1} + O(1)$ *uniformly for $\sigma \geqslant \frac{1}{2}$, $\sigma \neq 1$.*
(5) $\left|\frac{\zeta'}{\zeta}(\sigma + iT)\right| \leqslant -\frac{\zeta'}{\zeta}(\sigma)$ *for $\sigma > 1$.*

For a suitably small, fixed $\delta > 0$, we say that a number $T \geqslant 2$ is *admissible* if for all $\chi \in C(q) \cup \{\chi_0\}$ and all zeros $\frac{1}{2} + i\gamma$ of $L(s, \chi)$, $|\gamma - T| \geqslant \delta(\log T)^{-1}$. By Lemma 2.2, we can choose δ small enough, depending on q, so that there is an admissible T in $[U, U+1]$ for all $U \geqslant 2$. From Lemma 2.2 we obtain

Lemma 2.3. *Uniformly for $\sigma \geqslant \frac{2}{5}$ and admissible $T \geqslant 2$,*

$$|F(\sigma + iT, \chi)| = O(\log^2 T).$$

Lemma 2.4. *Fix $\chi \in C(q)$ and assume $L(\frac{1}{2}, \chi) \neq 0$. For $A \geqslant 0$ and real $k \geqslant 0$,*

$$\sum_{\substack{|\gamma_1|, |\gamma_2| \geqslant A \\ |\gamma_1 - \gamma_2| \geqslant 1}} \frac{\log^k(|\gamma_1| + 3) \log^k(|\gamma_2| + 3)}{|\gamma_1||\gamma_2||\gamma_1 - \gamma_2|} \ll_k \frac{(\log(A+3))^{2k+3}}{A+1}.$$

Proof. The sum in question is at most twice the sum of terms with $|\gamma_2| \geqslant |\gamma_1|$, which is

$$\ll \sum_{|\gamma_2|\geqslant A} \frac{\log^{2k}(|\gamma_2|+3)}{|\gamma_2|}\Bigg(\frac{1}{|\gamma_2|}\sum_{|\gamma_1|<\frac{|\gamma_2|}{2}} \frac{1}{|\gamma_1|} + \frac{1}{|\gamma_2|} \sum_{\substack{\frac{|\gamma_2|}{2}\leqslant|\gamma_1|\leqslant|\gamma_2| \\ |\gamma_2-\gamma_1|\geqslant 1}} \frac{1}{|\gamma_2-\gamma_1|}\Bigg).$$

By Lemma 2.2 (1), the two sums over γ_1 are $O(\log^2(|\gamma_2|+3))$. A further application of Lemma 2.2 (1) completes the proof. □

We conclude this section with a truncated version of the Perron formula for $G(x,u,v;\chi)$.

Lemma 2.5. *Uniformly for $x \leqslant T \leqslant 2x^2$, $x \geqslant 2$, $u \geqslant 0$ and $v \geqslant 0$, we have*

$$G(x,u,v;\chi) = \frac{1}{2\pi i}\int_{c-iT}^{c+iT} F(s+u,\chi)F(s+v,\chi)\frac{x^s}{s}ds + O(\log^3 x), \tag{2.1}$$

where $c = 1 + \frac{1}{\log x}$.

Proof. For $\Re s > 1$, we have

$$F(s+u,\chi)F(s+v,\chi) = \sum_{n=1}^{\infty} f(n)n^{-s}, \qquad f(n) = \sum_{p_1p_2=n} \frac{\chi(p_1p_2)\log p_1 \log p_2}{p_1^u p_2^v}.$$

Using the trivial estimate $|f(n)| \leqslant \log^2 n$ and a standard argument [Da, §17, (3) and (5)], we obtain the desired bounds. □

3. Outline of the proof of Theorem 2

Throughout the remainder of this paper, fix q, assume ERH$_q$ and that $L(\frac12,\chi) \neq 0$ for each $\chi \in C(q)$. Let

$$\varepsilon = \frac{1}{100}.$$

We next define a function $T(x)$ as follows. For each positive integer n, let T_n be an admissible value of T satisfying $\exp(2^{n+1}) \leqslant T_n \leqslant \exp(2^{n+1}) + 1$ and set $T(x) = T_n$ for $\exp(2^n) < x \leqslant \exp(2^{n+1})$. In particular, we have

$$x \leqslant T(x) \leqslant 2x^2 \qquad (x \geqslant e^2).$$

Our first task is to express the double integrals in (1.8) in terms of sums over zeros of $L(s,\chi)$. This is proved in Section 4.

Lemma 3.1. *Let $\chi \in C(q)$ and let $T = T(x)$. Then*

$$\begin{aligned} x^{-1/2}&\int_0^\infty\int_0^\infty G(x,u,v;\chi)\,du\,dv \\ &= 2\int_0^{2\varepsilon}\int_0^{2\varepsilon} \sum_{|\gamma|\leqslant T} \frac{F(\frac12+u-v+i\gamma,\chi)x^{-v+i\gamma}}{\frac12 - v + i\gamma} du\,dv + \frac{A(\chi)\log\log x + \Sigma_1(x;\chi) + O(1)}{\log x}, \end{aligned}$$

where $\int_1^Y |\Sigma_1(e^y;\chi)|^2 dy = O(Y)$.

The aggregate of terms $A(\chi)\log\log x/\log x$ account for the bias for products of two primes. As with the Chebyshev bias for primes, these terms arise from poles of $F(s)$ at $s=\frac12$ when $A(\chi)=1$ (see Lemma 2.1) and correspond to the contribution to $F(s)$ from squares of primes. The double integral on the right side in Lemma 3.1 is complicated to analyze. In Section 5 we prove the following.

Lemma 3.2. *Let $\chi \in C(q)$. Let n be a positive integer, $2^n < \log x \leqslant 2^{n+1}$ and $T = T(x)$. Then*

$$2\int_0^{2\varepsilon}\int_0^{2\varepsilon}\sum_{|\gamma|\leqslant T}\frac{F(\frac{1}{2}+u-v+i\gamma,\chi)x^{-v+i\gamma}}{\frac{1}{2}-v+i\gamma}du\,dv = \frac{\Sigma_2(x;\chi)}{\log x}$$
$$+ 2\sum_{\substack{|\gamma|\leqslant T\\ \gamma \text{ distinct}}} m^2(\gamma)x^{i\gamma}(\frac{1}{2}+i\gamma)\int_0^{2\varepsilon-2^{-n}}\frac{x^{-v}}{\frac{1}{2}-v+i\gamma}\int_{v+2^{-n}}^{2\varepsilon}\frac{du}{(u-v)(\frac{1}{2}-u+i\gamma)}dv,$$

where $\int_1^Y |\Sigma_2(e^y;\chi)|^2 dy = o(Y\log^2 Y)$.

The terms on the right in Lemma 3.2 with small $|\gamma|$ will give the main term, and terms with larger $|\gamma|$ are considered as error terms. The next lemma is proved in Section 6.

Lemma 3.3. *Let $\chi \in C(q)$. Let n be a positive integer, $2^n < \log x \leqslant 2^{n+1}$, $T = T(x)$ and $2 \leqslant T_0 \leqslant T$. Then*

$$2\sum_{\substack{|\gamma|\leqslant T\\ \gamma \text{ distinct}}} m^2(\gamma)x^{i\gamma}(\frac{1}{2}+i\gamma)\int_0^{2\varepsilon-2^{-n}}\frac{x^{-v}}{\frac{1}{2}-v+i\gamma}\int_{v+2^{-n}}^{2\varepsilon}\frac{du}{(u-v)(\frac{1}{2}-u+i\gamma)}dv$$
$$= \frac{2\log\log x}{\log x}\sum_{\substack{|\gamma|\leqslant T_0\\ \gamma \text{ distinct}}}\frac{m^2(\gamma)x^{i\gamma}}{1/2+i\gamma} + O\left(\frac{\log^3 T_0}{\log x}\right) + \frac{\Sigma_3(x,T_0;\chi)}{\log x},$$

where

$$\frac{1}{Y}\int_1^Y |\Sigma_3(e^y,T_0;\chi)|^2\,dy \ll \frac{\log^5 T_0}{T_0}\log^2 Y.$$

Combining Lemmas 3.1, 3.2 and 3.3 with (1.8) yields (for fixed, large T_0)

$$\Delta_2(x;q,a,b) = \frac{\sqrt{x}}{2\phi(q)}\sum_{\chi\in C(q)}(\overline{\chi}(a)-\overline{\chi}(b))\left[\frac{\log\log x}{\log x}\left(A(\chi)+2\sum_{\substack{|\gamma|\leqslant T_0\\ \gamma \text{ distinct}}}\frac{m^2(\gamma)x^{i\gamma}}{1/2+i\gamma}\right)\right.$$
$$\left. + \frac{\Sigma_1(x;\chi)+\Sigma_2(x;\chi)+\Sigma_3(x,T_0;\chi)+O(\log^3 T_0)}{\log x}\right],$$

where

$$\lim_{T_0\to\infty}\left(\limsup_{Y\to\infty}\frac{1}{Y\log^2 Y}\sum_{\chi\in C(q)}\int_1^Y |\Sigma_1(e^y;\chi)+\Sigma_2(e^y;\chi)+\Sigma_3(e^y;T_0;\chi)|^2\,dy\right) = 0.$$

On the other hand (cf. [RS]),

$$\Delta(x;q,a,b) = \frac{\sqrt{x}}{\log x}\left(\frac{N(q,b)-N(q,a)}{\phi(q)} - \sum_{\chi\in C(q)}(\overline{\chi}(a)-\overline{\chi}(b))\sum_{|\gamma|\leqslant T_0}\frac{x^{i\gamma}}{1/2+i\gamma} + \Sigma_4(x;T_0)\right),$$

where

$$\lim_{T_0\to\infty}\left(\limsup_{Y\to\infty}Y^{-1}\int_1^Y |\Sigma_4(e^y;T_0)|^2\,dy\right) = 0.$$

Now assume $m(\gamma) = 1$ for all γ, and note that

$$\sum_{\chi\in C(q)}(\overline{\chi}(a)-\overline{\chi}(b))\,A(\chi) = N(q,a)-N(q,b).$$

Letting $T_0 \to \infty$ finishes the proof of Theorem 2.

4. PROOF OF LEMMA 3.1

Assume ERH_q throughout. We first estimate $G(x,u,v;\chi)$ for different ranges of u,v.

Lemma 4.1. *Let $\chi \in C(q)$, $\chi \neq \chi_0$. For $x \geqslant 4$, the following hold:*

(1) *For $u, v \geqslant \varepsilon$, $G(x,u,v;\chi) \ll x^{\frac{1}{2}-\frac{\varepsilon}{2}} \log^5 x$.*

(2) *For $u \geqslant 2\varepsilon$, $v \leqslant \varepsilon$ and $T = T(x)$,*

$$x^{-1/2}G(x,u,v;\chi) = \sum_{|\gamma|\leqslant T} \frac{F(\frac12+u-v+i\gamma,\chi)x^{-v+i\gamma}}{\frac12 - v + i\gamma} - A(\chi)\frac{F(\frac12+u-v,\chi)x^{-v}}{1-2v} + O(x^{-\frac{3\varepsilon}{2}}\log^5 x).$$

(3) *For $u \leqslant 2\varepsilon$, $v \leqslant 2\varepsilon$, $u \neq v$ and $T = T(x)$,*

$$x^{-1/2}G(x,u,v;\chi) = \sum_{|\gamma|\leqslant T} \frac{F(\frac12+u-v+i\gamma,\chi)x^{-v+i\gamma}}{\frac12 - v + i\gamma} + \frac{F(\frac12-u+v+i\gamma,\chi)x^{-u+i\gamma}}{\frac12 - u + i\gamma}$$
$$- A(\chi)\left(\frac{F(\frac12+u-v,\chi)x^{-v}}{1-2v} + \frac{F(\frac12-u+v,\chi)x^{-u}}{1-2u}\right) + O(x^{-3\varepsilon}\log^5 x).$$

Proof. Assume $u \geqslant \varepsilon$ and $v \geqslant \varepsilon$. Start with the approximation of $G(x,u,v;\chi)$ given by Lemma 2.5, then deform the segment of integration to the contour consisting of three straight segments connecting $c-iT$, $b-iT$, $b+iT$ and $c+iT$, where $b = \frac12 - \frac{\varepsilon}{2}$ and $T = T(x)$. The rectangle formed by the new and old contours does not contain any poles of $F(s+u,\chi)F(s+v,\chi)s^{-1}$. On the three new segments, by Lemmas 2.1, 2.2 and 2.3, we have $|F(s+u,\chi)F(s+v,\chi)| \ll \log^4 T$. Hence the integral of $F(s+u,\chi)F(s+v,\chi)x^s s^{-1}$ over the three segments is

$$\ll (\log^4 x)\Big(\int_b^c \frac{x^\sigma}{|\sigma+iT|}\,d\sigma + \int_{-T}^T \frac{x^b}{|b+it|}\,dt\Big) \ll x^b \log^5 x.$$

This proves (1).

We now consider the case $v \leqslant \varepsilon$ and $u \geqslant 2\varepsilon$. We set $b = \frac12 - \frac{3\varepsilon}{2}$ and deform the contour of integration as in the previous case. Since $u + b \geqslant \frac12 + \frac{\varepsilon}{2}$ and $v + b \leqslant \frac12 - \frac{\varepsilon}{2}$, we have by Lemma 2.3 that $|F(s+u,\chi)F(s+v,\chi)| \ll \log^4 T \ll \log^4 x$ on all three new segments. As in the proof of (1), the integral over the new contour is $\ll x^b \log^5 x$. We pick up residue terms from poles of $F(s+v,\chi)$ inside the rectangle coming from the nontrivial zeros of $L(s,\chi)$, plus a pole at $s = \frac12 - v$ from the $\frac{\zeta'}{\zeta}(2s+2v)$ term if $\chi^2 = \chi_0$. The sum of the residues is

$$\sum_{|\gamma|\leqslant T} \frac{F(\frac12+u-v+i\gamma,\chi)x^{\frac12-v+i\gamma}}{\frac12 - v + i\gamma} - A(\chi)\frac{F(\frac12+u-v,\chi)x^{\frac12-v}}{1-2v},$$

and (2) follows.

Finally, consider the case $0 \leqslant u, v \leqslant 2\varepsilon$. Let $b = \frac12 - 3\varepsilon$ and deform the contour as in the previous cases. As before, the integral over the new contour is $O(x^b \log^5 x)$. This time, we pick up residues from poles of both $F(s+u,\chi)$ and $F(s+v,\chi)$. The sum of the residues is

$$\sum_{|\gamma|\leqslant T} \Big(\frac{F(\frac12+u-v+i\gamma,\chi)x^{\frac12-v+i\gamma}}{\frac12 - v + i\gamma} + \frac{F(\frac12-u+v+i\gamma,\chi)x^{\frac12-u+i\gamma}}{\frac12 - u + i\gamma}\Big)$$
$$-A(\chi)\left(\frac{F(\frac12+u-v,\chi)x^{\frac12-v}}{1-2v} + \frac{F(\frac12-u+v,\chi)x^{\frac12-u}}{1-2u}\right),$$

and (3) follows. □

Proof of Lemma 3.1. Begin with

$$\int_0^\infty \int_0^\infty G(x,u,v;\chi)\,du\,dv = I_1 + I_2 + 2I_3 + I_4,$$

where I_1 is the integral over $\max(u,v) \geqslant \log x$, I_2 is the integral over $2\varepsilon \leqslant \max(u,v) \leqslant \log x$ and $\min(u,v) \geqslant \varepsilon$, I_3 is the integral over $0 \leqslant v \leqslant \varepsilon$, $2\varepsilon \leqslant u \leqslant \log x$, and I_4 is the integral over $0 \leqslant u, v \leqslant 2\varepsilon$. For $\max(u,v) \geqslant \log x$,

$$|G(x,u,v;\chi)| \leqslant \sum_{p \leqslant x} \frac{\log p}{p^u} \sum_{p \leqslant x} \frac{\log q}{q^v} \ll \frac{x}{2^{\max(u,v)}},$$

whence $I_1 \ll x^{1-\log 2}$. By Lemma 4.1 (1), $I_2 \ll x^{1/2-\varepsilon/2}\log^7 x$.

By Lemma 4.1 (2),

$$I_3 = x^{1/2}\int_0^\varepsilon \int_{2\varepsilon}^{\log x} \sum_{|\gamma| \leqslant T} \frac{F(\frac12 + u - v + i\gamma, \chi)x^{-v+i\gamma}}{\frac12 - v + i\gamma} - A(\chi)\frac{F(\frac12 + u - v, \chi)x^{-v}}{1-2v}\,du\,dv + O(x^{1/2-\frac{3\varepsilon}{2}}\log^6 x). \tag{4.1}$$

By Lemmas 2.2 and 2.3,

$$\int_0^\varepsilon \int_{2\varepsilon}^{\log x} \frac{F(\frac12 + u - v, \chi)x^{-v}}{1-2v} du\,dv \ll \int_0^\varepsilon x^{-v}\,dv \ll \frac{1}{\log x}. \tag{4.2}$$

Let

$$\Sigma_1(x) = (\log x)\int_0^\varepsilon \int_{2\varepsilon}^{\log x} \sum_{0<|\gamma|<T} \frac{F(\frac12 + u - v + i\gamma, \chi)x^{-v+i\gamma}}{\frac12 - v + i\gamma}\,du\,dv.$$

Since $\frac12 + u - v \geqslant \frac12 + \varepsilon$ for $0 \leqslant v \leqslant \varepsilon$ and $2\varepsilon \leqslant u \leqslant \log x$, by Lemmas 2.1, 2.2, and 2.3,

$$F(\frac12 + u - v + i\gamma, \chi) = -\frac{L'}{L}(\frac12 + u - v + i\gamma, \chi) + O(1) \ll \log(|\gamma| + 3).$$

We also have $F(1/2 + u - v + i\gamma, \chi) \ll 2^{-u}$ for $u \geqslant 2$. Thus, for positive integers n,

$$\int_{2^n}^{2^{n+1}} |\Sigma_1(e^y)|^2 dy \ll 2^{2n} \sum_{|\gamma_1|,|\gamma_2| \leqslant T} \frac{\log(|\gamma_1| + 3)\log(|\gamma_2| + 3)}{|\gamma_1\gamma_2|} \times \int_0^\varepsilon \int_0^\varepsilon \left| \int_{2^n}^{2^{n+1}} e^{y(-v_1 + i\gamma_1 - v_2 - i\gamma_2)} dy \right| dv_1 dv_2.$$

The summands with $|\gamma_1 - \gamma_2| < 1$ contribute, by Lemma 2.2,

$$\ll 2^{2n} \sum_{\substack{|\gamma_1|,|\gamma_2| \leqslant T \\ |\gamma_1 - \gamma_2| < 1}} \frac{\log(|\gamma_1| + 3)\log(|\gamma_2| + 3)}{|\gamma_1||\gamma_2|} \int_{2^n}^{2^{n+1}} \Big(\int_0^\varepsilon e^{-vy} dv\Big)^2 dy$$

$$\ll 2^n \sum_{|\gamma| \leqslant T} \frac{\log^3(|\gamma| + 3)}{|\gamma|^2} \ll 2^n.$$

The summands with $|\gamma_1 - \gamma_2| \geqslant 1$ contribute, by Lemma 2.4,

$$\ll \sum_{\substack{|\gamma_1|,|\gamma_2| < T \\ |\gamma_1 - \gamma_2| \geqslant 1}} \frac{2^{2n}\log(|\gamma_1| + 3)\log(|\gamma_2| + 3)}{|\gamma_1||\gamma_2||\gamma_1 - \gamma_2|} \Big(\int_0^\varepsilon e^{-v2^n} dv\Big)^2 \ll 1.$$

Thus, $\int_{2^n}^{2^{n+1}} |\Sigma_1(e^y)|^2 dy = O(2^n)$. Summing over $n \leqslant \frac{\log Y}{\log 2} + 1$ yields $\int_1^Y |\Sigma_1(e^y)|^2\,dy = O(Y)$.

Finally, using Lemma 4.1 (3) gives

$$I_4 = x^{1/2}\int_0^{2\varepsilon}\int_0^{2\varepsilon}\sum_{|\gamma|\leqslant T}\frac{F(\frac12+u-v+i\gamma,\chi)x^{-v+i\gamma}}{\frac12-v+i\gamma}+\frac{F(\frac12-u+v+i\gamma,\chi)x^{-u+i\gamma}}{\frac12-u+i\gamma}$$
$$(4.3)\qquad -A(\chi)\left(\frac{F(\frac12+u-v,\chi)x^{-v}}{1-2v}+\frac{F(\frac12-u+v,\chi)x^{-u}}{1-2u}\right)du\,dv+O(x^{\frac12-3\varepsilon}\log^3 x).$$

Now assume $\chi^2=\chi_0$. We will show that

$$-\int_0^{2\varepsilon}\int_0^{2\varepsilon}\frac{F(\frac12+u-v,\chi)x^{-v}}{1-2v}+\frac{F(\frac12-u+v,\chi)x^{-u}}{1-2u}\,du\,dv=\frac{\log\log x+O(1)}{\log x}. \tag{4.4}$$

Together with (4.1), (4.2) and (4.3), this completes the proof of Lemma 3.1.

Note that $F(\frac12+w)=-\frac{1}{2w}+O(1)$ by Lemmas 2.1 and 2.3. Replacing x with e^y, the left side of (4.4) is

$$=\frac12\int_0^{2\varepsilon}\int_0^{2\varepsilon}\frac{e^{-yv}}{(u-v)(1-2v)}+\frac{e^{-yu}}{(v-u)(1-2u)}du\,dv+O\Big(\int_0^{2\varepsilon}\int_0^{2\varepsilon}e^{-yv}du\,dv\Big).$$

The error term above is $O(1/y)$. In the main term, when $|u-v|<1/y$, the integrand is $O(ye^{-vy})$ and the corresponding part of the double integral is $O(1/y)$. When $u\geqslant v+1/y$, the integrand is

$$\frac{e^{-vy}}{u-v}+O\left(\frac{ve^{-vy}+e^{-uy}}{u-v}\right)$$

and the corresponding part of the double integral is

$$\int_0^{2\varepsilon}e^{-vy}\log\left(\frac{y}{2\varepsilon-v}\right)dv+O\left(\frac1y\right)=\frac{\log y+O(1)}{y}.$$

The contribution from $u\leqslant v-1/y$ is, by symmetry, also $\frac{\log y+O(1)}{y}$. The asymptotic (4.4) follows. □

5. Proof of Lemma 3.2

Lemma 5.1. *Uniformly for $y\geqslant 1$, $0<|\xi|\leqslant 1$, $|w|\geqslant\frac12$ and $a\geqslant 0$ we have*

$$\left|\int_0^{2\varepsilon}\int_0^{2\varepsilon}\frac{v^a e^{-vy}}{(u-v+i\xi)(w-v)}du\,dv\right|\ll\frac{(4\varepsilon)^a\log\min(2y,\frac{2}{|\xi|})}{y|w|}.$$

Proof. Let I denote the double integral in the Lemma. If $|\xi|\geqslant\frac1y$, then

$$I\ll\frac{1}{|w|}\int_0^{2\varepsilon}v^a e^{-vy}\int_0^{2\varepsilon}\min\left(\frac{1}{|u-v|},\frac{1}{|\xi|}\right)du\,dv$$
$$\ll\frac{(2\varepsilon)^a}{|w|}\left(1+\log\frac{2}{|\xi|}\right)\int_0^{2\varepsilon}e^{-vy}\,dv\ll\frac{(2\varepsilon)^a\log(\frac{2}{|\xi|})}{y|w|}.$$

If $|\xi|<\frac1y$, let $I=I_1+I_2+I_3$, where I_1 is the part of I coming from $|u-v|\leqslant|\xi|$, I_2 is the part of I coming from $|\xi|<|u-v|\leqslant\frac1y$, and I_3 is the part of I coming from $|u-v|>\frac1y$. We have

$$I_1\ll\frac{1}{|w\xi|}\iint\limits_{\substack{0\leqslant u,v\leqslant 2\varepsilon\\|u-v|\leqslant|\xi|}}v^a e^{-vy}du\,dv\ll\frac{(2\varepsilon)^a}{y|w|}.$$

and

$$I_3\ll\frac{(2\varepsilon)^a}{|w|}\iint\limits_{\substack{0\leqslant u,v\leqslant 2\varepsilon\\|u-v|\geqslant\frac1y}}\frac{e^{-vy}}{|u-v|}du\,dv\ll\frac{(2\varepsilon)^a}{|w|}\int_0^{2\varepsilon}e^{-vy}(\log y+1)dv\ll\frac{(2\varepsilon)^a\log(2y)}{y|w|}.$$

By symmetry,

$$I_2 = \frac{1}{2} \iint\limits_{|\xi|<|u-v|\leqslant 1/y} \frac{v^a e^{-vy}}{(u-v+i\xi)(w-v)} + \frac{u^a e^{-uy}}{(v-u+i\xi)(w-u)}\, du\, dv.$$

Since, $|u^a - v^a| \leqslant a|u-v|(2\varepsilon)^{a-1}$,

$$\begin{aligned} u^a e^{-uy} - v^a e^{-vy} &= e^{-vy} v^a \left(e^{(v-u)y} - 1\right) + e^{-vy}(u^a - v^a)e^{(v-u)y} \\ &\ll e^{-vy} y|u-v|(4\varepsilon)^a. \end{aligned} \tag{5.1}$$

We deduce that

$$\begin{aligned} I_2 &= \iint\limits_{\substack{0\leqslant u,v\leqslant 2\varepsilon \\ |\xi|<|u-v|\leqslant 1/y}} \frac{(w-u)(u-v)(u^a e^{-uy} - v^a e^{-vy}) + u^a e^{-uy}(u-v)^2 + O(|\xi w|(2\varepsilon)^a e^{-vy})}{2(u-v+i\xi)(v-u+i\xi)(w-u)(w-v)} du\, dv \\ &\ll \frac{(4\varepsilon)^a}{|w|} \iint\limits_{\substack{0\leqslant u,v\leqslant 2\varepsilon \\ |\xi|<|u-v|\leqslant 1/y}} ye^{-vy} + \frac{|\xi| e^{-vy}}{|u-v|^2}\, du\, dv \ll \frac{(4\varepsilon)^a}{y|w|}. \end{aligned}$$

□

Proof of Lemma 3.2. Let $y = \log x$. We first note by Lemmas 2.1 and 2.2,

$$F(\frac{1}{2} + u - v + i\gamma, \chi) = \frac{m(\gamma)}{u-v} + R(\gamma, u-v) + R'(\gamma, u-v),$$

where

$$R(\gamma, w) = \sum_{0<|\gamma'-\gamma|\leqslant 1} \frac{1}{w + i(\gamma - \gamma')}, \qquad R'(\gamma, u-v) = O(\log(|\gamma|+3)).$$

Then, the double integral in Lemma 3.2 is

$$= \sum_{i=1}^{4} \Sigma_{2,i}(y) + 2 \sum_{\substack{|\gamma|\leqslant T \\ \gamma \text{ distinct}}} m^2(\gamma) e^{iy\gamma} (\frac{1}{2} + i\gamma) \int_0^{2\varepsilon - 2^{-n}} \frac{e^{-yv}}{\frac{1}{2} - v + i\gamma} \int_{v+2^{-n}}^{2\varepsilon} \frac{du}{(u-v)(\frac{1}{2} - u + i\gamma)} dv,$$

where

$$\Sigma_{2,1}(y) = 2 \int_0^{2\varepsilon} \int_0^{2\varepsilon} \sum_{|\gamma|\leqslant T} \frac{R(\gamma, u-v) e^{y(-v+i\gamma)}}{\frac{1}{2} - v + i\gamma} du\, dv,$$

$$\Sigma_{2,2}(y) = 2 \int_0^{2\varepsilon} \int_0^{2\varepsilon} \frac{R'(\gamma, u-v) e^{y(-v+i\gamma)}}{\frac{1}{2} - v + i\gamma}\, du\, dv,$$

$$\Sigma_{2,3}(y) = \sum_{\substack{|\gamma|\leqslant T \\ \gamma \text{ distinct}}} m^2(\gamma) e^{iy\gamma} (\frac{1}{2} + i\gamma) \iint\limits_{\substack{0\leqslant u,v\leqslant 2\varepsilon \\ |u-v|\leqslant 2^{-n}}} \frac{e^{-yv} - e^{-uy}}{(u-v)(\frac{1}{2} - v + i\gamma)(\frac{1}{2} - u + i\gamma)} dv\, du,$$

$$\Sigma_{2,4}(y) = 2 \sum_{\substack{|\gamma|\leqslant T \\ \gamma \text{ distinct}}} m^2(\gamma) e^{iy\gamma} (\frac{1}{2} + i\gamma) \int_{2^{-n}}^{2\varepsilon} \int_0^{v-2^{-n}} \frac{e^{-yv}}{(u-v)(\frac{1}{2} - v + i\gamma)(\frac{1}{2} - u + i\gamma)} du\, dv.$$

We show that $\sum_{j=1}^4 \Sigma_{2,j}(y)$ is small in mean square. Note that for $2^n < y \leqslant 2^{n+1}$, $T = T(e^y)$ is constant. Also, by Lemma 2.2, we have

$$m(\gamma) \ll \log(|\gamma| + 3). \tag{5.2}$$

First, by Lemmas 2.2 and 2.4,

$$\int_{2^n}^{2^{n+1}} |\Sigma_{2,2}(y)|^2\,dy = 4 \iiiint\limits_{[0,2\varepsilon]^4} \sum_{\substack{|\gamma_1|\leqslant T\\|\gamma_2|\leqslant T}} \frac{R'(\gamma_1,u_1-v_1)\overline{R'(\gamma_2,u_2-v_2)}}{(\frac12-v_1+i\gamma_1)(\frac12-v_2-i\gamma_2)}$$

$$\times \int_{2^n}^{2^{n+1}} e^{y(-v_1-v_2+i\gamma_1-i\gamma_2)}\,dy\,du_jdv_j$$

$$\ll \sum_{|\gamma_1-\gamma_2|>1} \frac{\log(|\gamma_1|+3)\log(|\gamma_2|+3)}{|\gamma_1\gamma_2|\cdot|\gamma_1-\gamma_2|} \iiiint\limits_{[0,2\varepsilon]^4} e^{-2^n(v_1+v_2)}\,du_jdv_j \tag{5.3}$$

$$+ \sum_{|\gamma_1-\gamma_2|\leqslant 1} \frac{\log(|\gamma_1|+3)\log(|\gamma_2|+3)}{|\gamma_1\gamma_2|} \int_{2^n}^{2^{n+1}} \iiiint\limits_{[0,2\varepsilon]^4} e^{-y(v_1+v_2)}\,du_jdv_jdy$$

$$\ll 2^{-n}.$$

For the remaining sums, for brevity we define

$$\rho_1 = \frac12 + i\gamma_1, \qquad \rho_2 = \frac12 - i\gamma_2.$$

Next,

$$\int_{2^n}^{2^{n+1}} |\Sigma_{2,3}(y)|^2dy = \int_{2^n}^{2^{n+1}} \sum_{|\gamma_1|,|\gamma_2|\leqslant T} m(\gamma_1)m(\gamma_2)e^{iy(\gamma_1-\gamma_2)}\rho_1\rho_2$$

$$\times \iiiint\limits_{\substack{[0,2\varepsilon]^4\\|u_j-v_j|\leqslant 2^{-n}}} \frac{(e^{-v_1y}-e^{-u_1y})(e^{-v_2y}-e^{-u_2y})}{\prod_{j=1}^2(u_j-v_j)(\rho_j-v_j)(\rho_j-u_j)}dv_jdv_j\,dy.$$

By (5.1), the integrand in the quadruple integral is $\ll y^2e^{-uy-u_1y}|\rho_1\rho_2|^{-2}$. By Lemma 2.2, for a given γ_1, there are $\ll \log(|\gamma_1|+3)$ zeros γ_2 with $|\gamma_1-\gamma_2|<1$. Hence, the contribution from terms with $|\gamma_1-\gamma_2|<1$ is

$$\ll 2^{-n}\sum_{|\gamma_1-\gamma_2|<1} \frac{m(\gamma_1)m(\gamma_2)}{|\rho_1\rho_2|} \ll 2^{-n}\sum_{\gamma_1} \frac{\log^3(|\gamma_1|+3)}{|\gamma_1|^2} \ll 2^{-n}.$$

Using integration by parts, we have

$$\int_{2^n}^{2^{n+1}} e^{iy(\gamma_1-\gamma_2)}(e^{-v_1y}-e^{-u_2y})(e^{-v_1y}-e^{-u_2y})\,dy \ll \frac{2^{3n}|u_1-v_1|\,|u_2-v_2|e^{-2^n(u_1+u_2)}}{|\gamma_1-\gamma_2|}$$

uniformly in u_1,v_1,u_2,v_2. Thus, by (5.2) and Lemma 2.4, the contribution from terms with $|\gamma_1-\gamma_2|\geqslant 1$ is

$$\ll 2^{-n}\sum_{|\gamma_1-\gamma_2|\geqslant 1} \frac{m(\gamma_1)m(\gamma_2)}{|\rho_1\rho_2|\cdot|\gamma_1-\gamma_2|} \ll 2^{-n}.$$

Combining these estimates, we have

$$\int_{2^n}^{2^{n+1}} |\Sigma_{2,3}(y)|^2dy \ll 2^{-n}. \tag{5.4}$$

In the same manner, we have

$$\int_{2^n}^{2^{n+1}} |\Sigma_{2,4}(y)|^2dy = \sum_{\substack{|\gamma_1|\leqslant T\\|\gamma_2|\leqslant T}} m(\gamma_1)m(\gamma_2)\rho_1\rho_2 \int_{2^n}^{2^{n+1}} \iiiint\limits_{\substack{[0,2\varepsilon]^4\\u_j\leqslant v_j-2^{-n}}} \frac{e^{y(-v_1-v_2+i(\gamma_1-\gamma_2))}du_jdv_j}{\prod_{j=1}^2(u_j-v_j)(\rho_j-v_j)(\rho_j-u_j)}dy.$$

The contribution to the right side from terms with $|\gamma_1-\gamma_2|<1$ is

$$\ll \sum_{|\gamma_1-\gamma_2|<1} \frac{m(\gamma_1)m(\gamma_2)}{|\gamma_1\gamma_2|} \int_{2^n}^{2^{n+1}} \left(\int_{2^{-n}}^{2\varepsilon} \int_0^{v-2^{-n}} \frac{e^{-yv}}{(v-u)}\, du\, dv \right)^2$$
$$\ll \sum_{\gamma_1} \frac{\log^3(|\gamma_1|+3)}{|\gamma_1|^2} \int_{2^n}^{2^{n+1}} \left(\int_{1/y}^{\infty} e^{-yv} \log(yv)\, dv \right)^2 \ll 2^{-n}.$$

The terms with $|\gamma_1-\gamma_2|>1$ contribute

$$\ll \sum_{\substack{|\gamma_1|,|\gamma_2|<T \\ |\gamma_1-\gamma_2|>1}} \frac{m(\gamma_1)m(\gamma_2)}{|\gamma_1\gamma_2|\cdot|\gamma_1-\gamma_2|} \left(\int_{2^{-n}}^{2\varepsilon} \int_0^{v-2^{-n}} \frac{e^{-2^n v}}{v-u}\, du\, dv \right)^2$$
$$\ll \sum_{|\gamma_1-\gamma_2|>1} \frac{\log(|\gamma_1|+3)\log(|\gamma_2|+3)}{|\gamma_1\gamma_2|\cdot|\gamma_1-\gamma_2|} \left(\frac{1}{2^n}\right)^2 \ll \frac{1}{2^{2n}}.$$

Therefore,

$$\int_{2^n}^{2^{n+1}} |\Sigma_{2,4}(y)|^2 dy \ll 2^{-n}. \tag{5.5}$$

Estimating an average of $\Sigma_{2,1}(y)$ is more complicated, since $R(\gamma,w)$ could be very large if $|w|$ is small and there is another γ' very close to γ. We get around the problem by noticing that $R(\gamma,w)+R(\gamma,-w)$ is always small. We first have, by (5.1) and Lemma 2.2,

$$\int_{2^n}^{2^{n+1}} |\Sigma_{2,1}(y)|^2 dy \ll \sum_{\gamma_1,\gamma_2} \log^2(|\gamma_1|+3)\log^2(|\gamma_2|+3) \max_{\substack{0<|\gamma_1-\gamma_1'|\leqslant 1 \\ 0<|\gamma_2-\gamma_2'|\leqslant 1}} \int_{2^n}^{2^{n+1}} e^{iy(\gamma_1-\gamma_2)}$$
$$\times \iiiint_{[0,2\varepsilon]^4} \frac{e^{y(-v_1-v_2)}}{(u_1-v_1+i\xi_1)(\rho_1-v_1)(u_2-v_2+i\xi_2)(\rho_2-v_2)} du_j dv_j\, dy, \tag{5.6}$$

where $\xi_1=\gamma_1-\gamma_1'$ and $\xi_2=-(\gamma_2-\gamma_2')$. Let

$$M(\gamma) = \max_{\substack{|\gamma-\gamma_1|\leqslant 1 \\ 0<|\gamma_1-\gamma_1'|<1}} \frac{2}{|\gamma_1-\gamma_1'|}.$$

By Lemmas 2.3 and 5.1, the terms with $|\gamma_1-\gamma_2|<1$ contribute

$$\ll \sum_{|\gamma_1-\gamma_2|<1} \frac{\log^2(|\gamma_1|+3)\log^2(|\gamma_2|+3)}{|\gamma_1\gamma_2|} \int_{2^n}^{2^{n+1}} \frac{1}{y^2} \prod_{j=1}^{2} \log\left(\min\left(2y, \frac{2}{|\gamma_j-\gamma_j'|}\right)\right) dy$$
$$\ll \frac{1}{2^n} \sum_{\gamma_1} \frac{\log^5(|\gamma_1|+3)}{|\gamma_1|^2} \log^2\left(\min(2^{n+2}, M(\gamma))\right) = o\left(\frac{n^2}{2^n}\right) \qquad (n\to\infty).$$

Now suppose $|\gamma_1-\gamma_2|>1$. With $\gamma_1,\gamma_2,\gamma_1',\gamma_2'$ all fixed, let $\Delta=\gamma_1-\gamma_2$. Fixing u_1,v_1,u_2,v_2, we integrate over y first. The quintuple integral in (5.6) is $J(2^{n+1})-J(2^n)$, where

$$J(y) = e^{iy\Delta} \iiiint_{[0,2\varepsilon]^4} \frac{e^{-y(v_1+v_2)}}{(i\Delta-v_1-v_2)\prod_{j=1}^{2}(u_j-v_j+i\xi_j)(\rho_j-v_j)} du_j dv_j.$$

Using

$$\frac{1}{i\Delta - v_1 - v_2} = \frac{1}{i\Delta}\sum_{k=0}^{\infty}\left(\frac{v_1+v_2}{i\Delta}\right)^k = \sum_{a,b\geqslant 0}\binom{a+b}{a}\frac{v_1^a v_2^b}{(i\Delta)^{a+b}},$$

together with Lemma 5.1, yields

$$|J(y)| \ll \frac{\log^2 y}{|\rho_1\rho_2\Delta|y^2}\sum_{a,b\geqslant 0}\binom{a+b}{a}\left(\frac{4\varepsilon}{|\Delta|}\right)^{a+b} \ll \frac{\log^2 y}{|\rho_1\rho_2\Delta|y^2}.$$

Therefore, by Lemma 2.4,

$$\sum_{\gamma_1,\gamma_2}\log^2(|\gamma_1|+3)\log^2(|\gamma_2|+3)\max_{\substack{0<|\gamma_1-\gamma_1'|\leqslant 1\\ 0<|\gamma_2-\gamma_2'|\leqslant 1}}|J(2^{n+1})-J(2^n)| \ll \frac{n^2}{2^{2n}},$$

and hence

$$\int_{2^n}^{2^{n+1}}|\Sigma_{2,1}(y)|^2 = o(n^2 2^{-n}). \tag{5.7}$$

Define

$$\Sigma_2(x;\chi) = (\log x)\sum_{j=1}^{4}\Sigma_{2,j}(\log x).$$

By (5.3), (5.4), (5.5) and (5.7),

$$\int_2^Y |\Sigma_2(e^y;\chi)|^2\,dy \leqslant 4\sum_{j=1}^{4}\sum_{n\leqslant\frac{\log Y}{\log 2}+1}2^{2n}\int_{2^n}^{2^{n+1}}|\Sigma_{2,j}(y)|^2\,dy = o(Y\log^2 Y)\qquad (Y\to\infty).$$

This completes the proof of Theorem 3.2. □

6. Proof of Lemma 3.3

Put $y=\log x$. For any γ we have

$$\begin{aligned}
&\int_0^{2\varepsilon-2^{-n}}\frac{e^{-yv}}{\frac12 - v + i\gamma}\int_{v+2^{-n}}^{2\varepsilon}\frac{du}{(u-v)(\frac12-u+i\gamma)}dv\\
&\quad=\int_0^{2\varepsilon-2^{-n}}e^{-yv}\Big(\frac{1}{\frac12+i\gamma}+O(\frac{v}{\frac14+\gamma^2})\Big)\int_{v+2^{-n}}^{2\varepsilon}\Big(\frac{1}{\frac12+i\gamma}+O(\frac{u}{\frac14+\gamma^2})\Big)\frac{du}{u-v}dv\\
&\quad=\frac{M+E}{(1/2+i\gamma)^2},
\end{aligned}$$

where

$$M=\int_0^{2\varepsilon-2^{-n}}e^{-yv}\left(\log(2\varepsilon-v)+\log 2^n\right)\,dv = \frac{\log y+O(1)}{y}$$

and

$$\begin{aligned}
E &\ll \int_0^{2\varepsilon-2^{-n}}e^{-yv}\int_{v+2^{-n}}^{2\varepsilon}\frac{u}{u-v}\,du\,dv\\
&\ll \int_0^{2\varepsilon-2^{-n}}e^{-yv}\left(1+v\log 2^n+v\log(2\varepsilon-v)\right)\,dv \ll \frac{1}{y}.
\end{aligned}$$

Hence, the zeros with $|\gamma| \leqslant T_0$ contribute

$$\frac{2\log\log x}{\log x} \sum_{\substack{|\gamma|\leqslant T_0 \\ \gamma \text{ distinct}}} \frac{m^2(\gamma)x^{i\gamma}}{1/2+i\gamma} + O\left(\frac{\log^3 T_0}{\log x}\right).$$

Next, let $\Sigma_3(x;T_0)$ be the sum over zeros with $T_0 < |\gamma| \leqslant T$. We have

$$\int_{2^n}^{2^{n+1}} |\Sigma_3(e^y,T_0)|^2 dy \leqslant \sum_{T_0\leqslant|\gamma_1|,|\gamma_2|\leqslant T} 2^{2n+2} m(\gamma_1)m(\gamma_2)\left(\frac{1}{2}+i\gamma_1\right)\left(\frac{1}{2}-i\gamma_2\right) \tag{6.1}$$

$$\int_{2^n}^{2^{n+1}} e^{yi(\gamma_1-\gamma_2)} \iiiint_{u_j\geqslant v_j+2^{-n}} \frac{e^{-yv_1-yv_2}}{\prod_{j=1}^2 (u_j-v_j)(\frac{1}{2}-v_j+i\gamma_j)(\frac{1}{2}-u_j+i\gamma_j)} du_j dv_j\, dy.$$

The sum over $|\gamma_1-\gamma_2|<1$ on the right side of (6.1) is

$$\ll \sum_{\substack{T_0\leqslant|\gamma_1|,|\gamma_2|\leqslant T \\ |\gamma_1-\gamma_2|<1}} \frac{2^{2n}m(\gamma_1)m(\gamma_2)}{|\gamma_1||\gamma_2|} \int_{2^n}^{2^{n+1}} \iiiint_{u_j\geqslant v_j+2^{-n}} \frac{e^{-yv_1-yv_2}}{(u_1-v_1)(u_2-v_2)} du_j dv_j\, dy$$

$$\ll \sum_{\substack{T_0\leqslant|\gamma_1|,|\gamma_2|\leqslant T \\ |\gamma_1-\gamma_2|<1}} \frac{n^2 2^n m(\gamma_1)m(\gamma_2)}{|\gamma_1||\gamma_2|} \ll n^2 2^n \sum_{|\gamma|\geqslant T_0} \frac{\log^3(|\gamma|+3)}{|\gamma|} \ll \frac{n^2 2^n \log^5 T_0}{T_0},$$

applying Lemma 2.2. The terms where $|\gamma_1-\gamma_2|>1$ on the right hand side of (6.1) total

$$\ll \sum_{\substack{T_0\leqslant|\gamma_1|,|\gamma_2|\leqslant T \\ |\gamma_1-\gamma_2|>1}} \frac{2^{2n}m(\gamma_1)m(\gamma_2)}{|\gamma_1||\gamma_2||\gamma_1-\gamma_2|} \iiiint_{u_j\geqslant v_j+2^{-n}} \frac{e^{-2^n v_1-2^n v_2}}{(u_1-v_1)(u_2-v_2)} du_j dv_j$$

$$\ll \sum_{\substack{T_0\leqslant|\gamma_1|,|\gamma_2| \\ |\gamma_1-\gamma_2|>1}} \frac{n^2\log(|\gamma_1|+3)\log(|\gamma_2|+3)}{|\gamma_1||\gamma_2||\gamma_1-\gamma_2|} \ll n^2 \frac{\log^5 T_0}{T_0}.$$

by Lemma 2.4. Summing over n proves the lemma.

REFERENCES

[Ch] P. L. Chebyshev, *Lettre de M. le professeur Tchébyshev á M. Fuss, sur un nouveau théoreme rélatif aux nombres premiers contenus dans la formes* $4n+1$ *et* $4n+3$, Bull. de la Classe phys.-math. de l'Acad. Imp. des Sciences St. Petersburg **11** (1853), 208.

[Da] H. Davenport, *Multiplicative Number Theory, 3rd ed., Graduate Texts in Mathematics vol. 74*, Springer-Verlag, New York-Berlin, 2000.

[FM] D. Fiorilli and G. Martin, *Inequities in the Shanks-Rényi Prime Number Race: An asymptotic formula for the densities* , (pre-print, 2009, arXiv:0912.4908).

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

[GM] A. Granville and G. Martin, *Prime number races*. Amer. Math. Monthly **113** (2006), no. 1, 1–33.

[KT] S. Knapowski and P. Turán, *Comparative Prime Number Theory I.*, Acta. Math. Sci. Hungar. **13** (1962), 315-342.

[La] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen, 3rd ed.*, Chelsea, New York, 1974.

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

[Li] J. E. Littlewood, *Sur la distribution des nombres premiers*, C. R. Acad. des Sciences Paris **158** (1914), 1869-1872.

[RS] M. Rubinstein and P. Sarnak, *Chebyshev's Bias*, J. Exper. Math. **3** (1994), 173-197.

[S] J. Sneed, *Lead changes in the prime number race*, Math. Comp. (to appear).

E-mail address: ford@math.uiuc.edu, jpsneed@uiuc.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, 1409 WEST GREEN ST., URBANA, IL 61801