MATHEMATICS OF COMPUTATION
Volume 79, Number 270, April 2010, Pages 1209–1229
S 0025-5718(09)02312-6
Article electronically published on September 25, 2009

# AVERAGE PRIME-PAIR COUNTING FORMULA

JAAP KOREVAAR AND HERMAN TE RIELE

ABSTRACT. Taking $r > 0$, let $\pi_{2r}(x)$ denote the number of prime pairs $(p,\, p+2r)$ with $p \le x$. The prime-pair conjecture of Hardy and Littlewood (1923) asserts that $\pi_{2r}(x) \sim 2C_{2r}\,\mathrm{li}_2(x)$ with an explicit constant $C_{2r} > 0$. There seems to be no good conjecture for the remainders $\omega_{2r}(x) = \pi_{2r}(x) - 2C_{2r}\,\mathrm{li}_2(x)$ that corresponds to Riemann's formula for $\pi(x) - \mathrm{li}(x)$. However, there is a heuristic approximate formula for averages of the remainders $\omega_{2r}(x)$ which is supported by numerical results.

## 1. Introduction

For $r \in \mathbb{N}$, let $\pi_{2r}(x)$ denote the number of prime pairs $(p,\, p+2r)$ with $p \le x$. The famous prime-pair conjecture (PPC) of Hardy and Littlewood [12] asserts that for $x \to \infty$,

$$\pi_{2r}(x) \sim 2C_{2r}\mathrm{li}_2(x) = 2C_{2r}\int_2^x \frac{dt}{\log^2 t} \sim 2C_{2r}\,\frac{x}{\log^2 x}. \tag{1.1}$$

Here $C_2$ is the 'twin-prime constant',

$$C_2 = \prod_{p\ \mathrm{prime},\ p>2} \left\{1 - \frac{1}{(p-1)^2}\right\} \approx 0.6601618158, \tag{1.2}$$

and the 'general prime-pair constant' $C_{2r}$ is given by

$$C_{2r} = C_2 \prod_{p\ \mathrm{prime},\ p|r,\ p>2} \frac{p-1}{p-2}. \tag{1.3}$$

Assuming that the PPC is true, let $\omega_{2r}(x)$ denote the remainder

$$\omega_{2r}(x) \stackrel{\mathrm{def}}{=} \pi_{2r}(x) - 2C_{2r}\mathrm{li}_2(x). \tag{1.4}$$

We have not been able to find a good approximation for the remainder $\omega_{2r}(x)$ that corresponds to Riemann's approximate formula for $\pi(x) - \mathrm{li}(x)$ (see (1.9) below). Instead, by complex analysis and heuristic arguments, we obtain the following *plausible approximation* for *averages* $(1/N)\sum_{r=1}^{N} \omega_{2r}(x)$ with large $N$ (cf. Section 5), and we support the formula by extensive numerical results.

Received by the editor February 25, 2009 and, in revised form, June 5, 2009.
2000 *Mathematics Subject Classification.* Primary 11P32; Secondary 65-05.
*Key words and phrases.* Hardy–Littlewood conjecture, prime-pair functions, representation by repeated complex integral, zeta's complex zeros.

**Approximation 1.1.** For $N \geq 1$ and $x \geq N^{2+\delta}$, with $0 < \delta \leq 1$, one has

$$
\begin{aligned}
(1.5) \qquad \frac{1}{N}\sum_{r=1}^{N}\{\pi_{2r}(x) - 2C_{2r}\mathrm{li}_2(x)\} &= -\{4 + \mathcal{O}(N^{-1/2}\log x)\}\sum_{\rho}\rho\,\mathrm{li}_2(x^{\rho}) \\
&\quad - \{1 + \mathcal{O}(N^{-1/2}\log x)\}\mathrm{li}_2(x^{1/2}) + \mathcal{O}(x^{1/(2+\delta)}),
\end{aligned}
$$

with a symmetric sum over the complex zeros $\rho$ of $\zeta(s)$.

To test this conjectured approximation we observe that

$$
(1.6) \qquad \begin{aligned}
&\frac{\sum_{\rho}\rho\,\mathrm{li}_2(x^{\rho})}{\mathrm{li}_2(x^{1/2})} = \left\{\frac{1}{4} + \mathcal{O}\left(\frac{1}{\log x}\right)\right\}T(x) + \mathcal{O}\left(\frac{1}{\log x}\right), \quad \text{where} \\
&T(x) \overset{\text{def}}{=} \sum_{\rho}\frac{x^{\rho-1/2}}{\rho} \quad \text{is real-valued;}
\end{aligned}
$$

cf. Section 6. Neglecting the $\mathcal{O}$-terms in (1.5) and (1.6), dividing by $\mathrm{li}_2(x^{1/2})$, and adding $T(x)+1$, we obtain the error function

$$
(1.7) \qquad \Delta_N(x) \overset{\text{def}}{=} \frac{\sum_{r=1}^{N}\omega_{2r}(x)}{N\mathrm{li}_2(x^{1/2})} + T(x) + 1.
$$

We have evaluated and plotted this function for fixed $x = 10^6$, $10^8$, $10^{10}$, $10^{12}$ and $2 \leq 2N \leq 5000$ (Figures 1–4 in Section 7), and for fixed $N = 400, 2500$ and $6 \leq \log_{10} x \leq 12$ (Figures 5–6 in Section 7). Taking into account the $\mathcal{O}$-terms in (1.5) and (1.6), $\Delta_N(x)$ should have the form

$$
\mathcal{O}\big(N^{-1/2}(\log x) + 1/\log x\big)T(x) + \mathcal{O}\big(N^{-1/2}(\log x) + 1/\log x\big).
$$

Our four plots for fixed $x$, and the two for fixed $N$, show that Approximation 1.1 is good for large $N$, provided $x/N^2$ is large.

When $x$ is comparable to $N^2$, the theory predicts a sizeable deviation, roughly

$$
(1.8) \qquad \Delta_N(x) \approx \overline{\Delta}_N(x) \overset{\text{def}}{=} -\frac{2N\log^2 x}{8x^{1/2}\log^2 2N};
$$

see Section 5. Behavior of this type is seen in the plots for $x = 10^6$ and $10^8$.

In connection with (1.5), we recall Riemann's approximation for the remainder $\pi(x) - \mathrm{li}(x)$. If $\mathrm{Re}\,\rho = 1/2$ for all $\rho$ one has

$$
(1.9) \qquad \omega(x) \overset{\text{def}}{=} \pi(x) - \mathrm{li}(x) = -\sum_{\rho}\mathrm{li}(x^{\rho}) - (1/2)\,\mathrm{li}(x^{1/2}) + \mathcal{O}(x^{b}),
$$

where $b$ may be any number greater than $1/3$. This can be derived from von Mangoldt's formula

$$
(1.10) \qquad \psi(x) \overset{\text{def}}{=} \sum_{n \leq x}\Lambda(n) = x - \sum_{\rho}\frac{x^{\rho}}{\rho} - \frac{\zeta'(0)}{\zeta(0)} + \sum_{k}\frac{x^{-2k}}{2k};
$$

cf. Davenport [8], Edwards [9], and Ivić [14]. Recall that von Mangoldt's function $\Lambda(n)$ is equal to $\log p$ if $n = p^{\alpha}$ with $p$ prime, and equal to 0 if $n$ is not a prime power. Formula (1.10) is exact at all points $x > 1$ where $\psi(x)$ is continuous.

## PART I. HEURISTICS

### 2. First step towards conjectured Approximation 1.1

Let us start by introducing the functions

$$\psi_{2r}(x) \stackrel{\text{def}}{=} \sum_{n\le x} \Lambda(n)\Lambda(n+2r), \quad \theta_{2r}(x) \stackrel{\text{def}}{=} \sum_{p,\, p+2r \text{ prime};\, p\le x} \log^2 p,$$
$$\theta^*_{2r}(x) \stackrel{\text{def}}{=} \sum_{p,\, p^2\pm 2r \text{ prime};\, p\le x} \log^2 p. \tag{2.1}$$

Partial summation or integration by parts shows that the PPC (1.1) is equivalent to each of the asymptotic relations

$$\theta_{2r}(x) \sim 2C_{2r}x, \quad \psi_{2r}(x) \sim 2C_{2r}x \quad \text{as } x\to\infty. \tag{2.2}$$

We have counted the prime pairs $(p, p+2r)$ with $2r \le 5\cdot 10^3$ and $p \le x = 10^3, 10^4, \dots, 10^{12}$. Table 1 is based on this work; cf. also a table in Granville and Martin [11] and one by Fokko van de Bult [7]. The bottom line shows (rounded) values of the comparison function

$$L_2(x) \stackrel{\text{def}}{=} 2C_2\mathrm{li}_2(x) \quad \text{for } \pi_2(x).$$

Computations based on these prime-pair counts make it plausible that for every $r\in\mathbb{N}$ and every $\varepsilon > 0$,

$$\omega_{2r}(x) = \pi_{2r}(x) - 2C_{2r}\mathrm{li}_2(x) = \mathcal{O}(x^{(1/2)+\varepsilon}). \tag{2.3}$$

Equivalently, there would be relations

$$\Omega_{2r}(x) \stackrel{\text{def}}{=} \psi_{2r}(x) - 2C_{2r}x = \mathcal{O}(x^{(1/2)+\varepsilon'}), \tag{2.4}$$

which follow from similar estimates for $\theta_{2r}(x) - 2C_{2r}x$.

Our work requires a good estimate for the difference $\psi_{2r}(x) - \theta_{2r}(x)$. The nonvanishing terms $\Lambda(n)\Lambda(n+2r)$ of $\psi_{2r}(x)$ have the form $\log p \log q$, where $p$ and $q$ are distinct primes. We distinguish four cases: (1) $n = p$, $n+2r = q$; (2) $n = p^2$, $n+2r = q$; (3) $n = p$, $n+2r = q^2$; (4) $n = p^\alpha$, $n+2r = q^\beta$ with $\alpha+\beta \ge 4$. Taking $x > 2r$ we compare the sum for case (1) with $\theta_{2r}(x)$:

$$\sum_{p\le x;\, p+2r \text{ prime}} \{\log p\, \log(p+2r) - \log^2 p\} = \int_2^x \log t\, \log(1+2r/t)\, d\pi_{2r}(t)$$
$$= \int_{2r}^x + \int_2^{2r} = 4rC_{2r}\log\log x + \mathcal{O}(rC_{2r}/\log 2r).$$

The contributions to $\psi_{2r}(x)$ from cases (2) and (3) are, respectively,

$$\sum_{p^2\le x;\, q=p^2+2r} \log p \log q \approx 2 \sum_{p\le x^{1/2};\, p^2+2r \text{ prime}} \log^2 p,$$
$$\sum_{p\le x;\, q^2=p+2r} \log p \log q \approx 2 \sum_{q\le (x+2r)^{1/2};\, q^2-2r \text{ prime}} \log^2 q.$$

Here the factors 2 come from the fact that $q\approx p^2$ in the first formula, and $p\approx q^2$ in the second. The sum of the two contributions is well-approximated by $2\theta^*_{2r}(x^{1/2})$; see (2.1). Finally, we consider case (4). If $\alpha=\beta=2$ so that $q^2-p^2 = 2r$, one has $q+p\le r$ and $\log p \log q < \log^2 r$, while the number of possibilities for $p$ and $q$ is bounded by $d(r) = \mathcal{O}(r^\varepsilon)$. When $\alpha\ge 3$ or $\beta \ge 3$, one of the primes $p$, $q$ is $\mathcal{O}(x^{1/3})$.

The total contribution to $\psi_{2r}(x)$ in case (4) can then be estimated as $\mathcal{O}(r^{\varepsilon}\log^2 r) + \mathcal{O}(x^{1/3}\log^2 x)$. Summarizing, one finds that uniformly in $r$,

$$\begin{aligned}\psi_{2r}(x)-\theta_{2r}(x) &= 2\theta^*_{2r}(x^{1/2}) + 4rC_{2r}\log\log x\\ &\quad + \mathcal{O}(rC_{2r}/\log 2r) + \mathcal{O}(x^{1/3}\log^2 x).\end{aligned}\tag{2.5}$$

Our first goal will be to motivate the following heuristic.

**Approximation 2.1.** Taking $N$ large and $x$ much larger than $N$, one has

$$\begin{aligned}\frac{1}{N}\sum_{r=1}^{N}\{\psi_{2r}(x)-2C_{2r}\,x\} &= -\{4+\mathcal{O}(N^{-1/2}\log x)\}\sum_{\rho}\frac{x^{\rho}}{\rho}\\ &\quad + \mathcal{O}(N^{-1/2}x^{1/2}\log x) - \{1+o(1)\}N.\end{aligned}\tag{2.6}$$

The step from Approximation 2.1 to Approximation 1.1 will be carried out in Section 5.

In support of the conjectured Approximation 2.1 we will derive a related conjecture involving Dirichlet series. For $s=\sigma+i\tau$ with $\sigma>1/2$, set

$$D_{2r}(s) \overset{\text{def}}{=} \sum_{n=1}^{\infty}\frac{\Lambda(n)\Lambda(n+2r)}{n^{2s}} = \int_1^{\infty} x^{-2s}d\psi_{2r}(x) = 2s\int_1^{\infty} x^{-2s-1}\psi_{2r}(x)dx.\tag{2.7}$$

Here we use the denominator $n^{2s}$ (and not $n^s$) because of the function $\Phi^{\lambda}(s)$ in Theorem 3.1 and the corresponding integral in (8.2).

| $2r\backslash x$ | $10^3$ | $10^4$ | $10^6$ | $10^8$ | $10^{10}$ | $10^{12}$ | $C_{2r}/C_2$ |
|---|---|---|---|---|---|---|---|
| 2 | 35 | 205 | 8169 | 440312 | 27412679 | 1870585220 | 1 |
| 4 | 41 | 203 | 8144 | 440258 | 27409999 | 1870585459 | 1 |
| 6 | 74 | 411 | 16386 | 879908 | 54818296 | 3741217498 | 2 |
| 8 | 38 | 208 | 8242 | 439908 | 27411508 | 1870580394 | 1 |
| 10 | 51 | 270 | 10934 | 586811 | 36548839 | 2494056601 | 4/3 |
| 12 | 70 | 404 | 16378 | 880196 | 54822710 | 3741051790 | 2 |
| 14 | 48 | 245 | 9878 | 528095 | 32891699 | 2244614812 | 6/5 |
| 16 | 39 | 200 | 8210 | 441055 | 27414828 | 1870557044 | 1 |
| 18 | 74 | 417 | 16451 | 880444 | 54823059 | 3741063106 | 2 |
| 20 | 48 | 269 | 10972 | 586267 | 36548155 | 2494072774 | 4/3 |
| 22 | 41 | 226 | 9171 | 489085 | 30459489 | 2078443752 | 10/9 |
| 24 | 79 | 404 | 16343 | 880927 | 54823858 | 3741122743 | 2 |
| 30 | 99 | 536 | 21990 | 1173934 | 73094856 | 4988150875 | 8/3 |
| 210 | 107 | 641 | 26178 | 1409150 | 87712009 | 5985825351 | 16/5 |
| $L_2(x)$: | 46 | 214 | 8248 | 440368 | 27411417 | 1870559867 | |

TABLE 1. $\pi_{2r}(x)$ for selected values of $2r$ and $x$

By a two-way Wiener–Ikehara theorem for Dirichlet series with positive coefficients, the PPC in the form (2.2) is true if and only if the difference

$$G_{2r}(s) \stackrel{\text{def}}{=} D_{2r}(s) - \frac{2C_{2r}}{2s-1} = \int_1^\infty x^{-2s} d\Omega_{2r}(x) = 2s \int_1^\infty x^{-2s-1}\Omega_{2r}(x)dx \tag{2.8}$$

has 'good' boundary behavior as $\sigma \searrow 1/2$. That is, $G_{2r}(\sigma + i\tau)$ should tend to a distribution $G_{2r}\{(1/2)+i\tau\}$ which is locally equal to a pseudofunction; see Korevaar [15]. Here, a pseudofunction is the distributional Fourier transform of a bounded function which tends to zero at infinity. It cannot have poles and is locally given by Fourier series whose coefficients tend to zero. In particular, $D_{2r}(s)$ itself would have to show pole-type behavior, with residue $C_{2r}$, for angular approach of $s$ to $1/2$ from the right; there should be no other poles on the line $\{\sigma = 1/2\}$.

In view of the expected estimate (2.4) it is reasonable to suppose that the difference $G_{2r}(s)$ is actually analytic for $\sigma > 1/4$. Where would one expect the first singularities? Assuming Riemann's Hypothesis (RH), we will motivate a conjecture involving *averages* of functions $G_{2r}(s)$:

**Conjecture 2.2.** *For $\sigma > 1/4$ and $N \to \infty$, one has*

$$\begin{aligned}\frac{1}{N}\sum_{r=1}^{N} G_{2r}(s) &= \frac{\mathcal{O}(N^{-1/2})}{(4s-1)^2} + \frac{\mathcal{O}(N^{-1/2})}{4s-1} + \mathcal{O}(N^{-1/2})\sum_\rho \frac{1}{(2s-\rho)^2} \\ &\quad - \{4+\mathcal{O}(N^{-1/2})\}\sum_\rho \frac{1}{2s-\rho} + H^N(s),\end{aligned} \tag{2.9}$$

*with symmetric sums over zeta's complex zeros $\rho$. The remainder $H^N(s)$ has 'good' boundary behavior as $\sigma \searrow 1/4$. For large $N$ its most significant part may be a term $-\{1+o(1)\}N$.*

This conjecture motivates Approximation 2.1 through formal Fourier inversion; cf. (2.8). If $L(c)$ denotes a 'vertical line' given by $\sigma = c > 1/4$, then

$$\Omega_{2r}(x) = \frac{1}{2\pi i}\int_{L(c)} G_{2r}(s)x^{2s}\frac{ds}{s}. \tag{2.10}$$

## 3. The theorem behind Conjecture 2.2

To arrive at (2.9) we start with a result for a weighted sum of functions $D_{2r}(s)$; cf. [16], where there is a less precise result. The weights are derived from an even 'sieving function' $E(\nu)$, with $E(0) = 1$ and support $[-1, 1]$, that can be made to approach 1 on $(-1, 1)$. A minimal smoothness requirement is that $E(\nu)$ be absolutely continuous, with derivative $E'(\nu)$ of bounded variation.

**Theorem 3.1.** *Assume RH. Then for $\lambda > 0$ and $1/2 < \sigma < 1$,*

$$\begin{aligned}\Phi^\lambda(s) &\stackrel{\text{def}}{=} D_0(s) + 2\sum_{0<2r\le\lambda} E(2r/\lambda)D_{2r}(s) \\ &= \frac{2A^E\lambda}{2s-1} - 4A^E\lambda\sum_\rho \frac{1}{2s-\rho} + \Sigma^\lambda(s) + H_0^\lambda(s).\end{aligned} \tag{3.1}$$

*The function $D_0(s)$ is obtained from (2.7) by taking $r = 0$, and the constant $A^E$ is given by $\int_0^1 E(\nu)d\nu$. The function $\Sigma^\lambda(s) = \Sigma^{\lambda,E}(s)$ is given by a sum which will be described below, and the remainder $H_0^\lambda(s) = H_0^{\lambda,E}(s)$ is analytic for $0 < \sigma < 1$.*

*For large $\lambda$ its most significant part may be a term that behaves like $-\lambda^2/2$ when $E(\nu)$ is close to* 1 *on* $(-1,1)$.

The proof of this theorem is described in the Appendix (Section 8). The function $D_0(s)$ above can be written as follows:

$$D_0(s) = \sum_{n=1}^{\infty} \frac{\Lambda^2(n)}{n^{2s}} = \frac{1}{2}\frac{d}{ds}\left\{\frac{\zeta'(2s)}{\zeta(2s)} - \frac{1}{2}\frac{\zeta'(4s)}{\zeta(4s)}\right\} + H_1(s), \tag{3.2}$$

where $H_1(s)$ is analytic for $\sigma > 1/6$. Hence $D_0(s)$ is meromorphic for $\sigma > 1/6$. Its poles there are purely quadratic, and located at $s = 1/2$, $1/4$ and the points $\rho/2$. Thus by (3.1), and under assumption (2.4), the pole of the difference $\Sigma^\lambda(s) - D_0(s)$ at $s = 1/2$ can only be of first order. Under (2.4) the residue $R(1/2,\lambda)$ will be equal to

$$2\sum_{0<2r\le\lambda} E(2r/\lambda)C_{2r} - A^E\lambda, \quad \text{which we call } \mathbb{R}_0^E(1/2,\lambda). \tag{3.3}$$

We need the important fact that the constants $C_{2r}$ have mean value one. Stronger results were obtained by Bombieri–Davenport [5] and Montgomery [19], and these were later improved by Friedlander and Goldston [10] to

$$S_m \overset{\text{def}}{=} \sum_{r=1}^{m} C_{2r} = m - (1/2)\log m + \mathcal{O}\{\log^{2/3}(m+1)\}. \tag{3.4}$$

Partial summation in (3.3) will now show that for our sieving functions $E$, the quantity $\mathbb{R}_0^E(1/2,\lambda)$ is $o(\lambda)$ and in fact, $\mathcal{O}(\log\lambda)$ as $\lambda\to\infty$.

The description of $\Sigma^\lambda(s)$ requires a *Mellin transform* associated with the Fourier transform $\hat{E}^\lambda(t)$ of $E(\nu/\lambda)$. For $z = x + iy$ with $0 < x < 1$ we set

$$\begin{aligned} M^\lambda(z) &\overset{\text{def}}{=} \frac{1}{\pi}\int_0^\infty \hat{E}^\lambda(t)t^{-z}dt \\ &= \frac{2}{\pi}\lambda^z\Gamma(1-z)\sin(\pi z/2)\int_0^1 E(\nu)\nu^{z-1}dz \\ &= \frac{2}{\pi}\lambda^z\Gamma(-z-1)\sin(\pi z/2)\int_0^{1+}\nu^{z+1}dE'(\nu). \end{aligned} \tag{3.5}$$

The function $M^\lambda(z)$ extends to a meromorphic function for $x > -1$ with simple poles at the points $z = 1, 3, \ldots$. The residue of the pole at $z = 1$ is $-2(\lambda/\pi)A^E$ with $A^E = \int_0^1 E(\nu)d\nu$, and $M^\lambda(0) = 1$. Furthermore, the standard order estimates

$$\Gamma(z) \ll |y|^{x-1/2}e^{-\pi|y|/2}, \quad \sin(\pi z/2) \ll e^{\pi|y|/2} \tag{3.6}$$

for $|x| \le C$ and $|y| \ge 1$ (cf. Whittaker and Watson [24]) imply the useful majorization

$$M^\lambda(x+iy) \ll \lambda^x(|y|+1)^{-x-3/2} \quad \text{for} \ -1 < x \le C, \ |y| \ge 1. \tag{3.7}$$

**Example 3.2.** One may take $E(\nu/\lambda)$ equal to the Fejér kernel for $\mathbb{R}$:

$$E(\nu/\lambda) = \frac{1}{\pi}\int_0^\infty \frac{\sin^2(\lambda t/2)}{\lambda(t/2)^2}\cos\nu t\,dt = \begin{cases} 1 - |\nu|/\lambda & \text{for } |\nu| \le \lambda, \\ 0 & \text{for } |\nu| \ge \lambda. \end{cases}$$

In this case one finds

$$M^\lambda(z) = \frac{2}{\pi}\lambda^z\Gamma(-z-1)\sin(\pi z/2).$$

*The function* $\Sigma^\lambda(s)$. For any $\lambda > 0$, the function $\Sigma^\lambda(s)$ is given by the sum

$$\begin{aligned}(3.8)\qquad &\left\{\frac{\zeta'(s)}{\zeta(s)}\right\}^2 + 2\,\frac{\zeta'(s)}{\zeta(s)}\sum_\rho \Gamma(\rho-s)M^\lambda(\rho-s)\cos\{\pi(\rho-s)/2\}\\ &+\sum_{\rho',\,\rho''}\Gamma(\rho'-s)\Gamma(\rho''-s)M^\lambda(\rho'+\rho''-2s)\cos\{\pi(\rho'-\rho'')/2\}.\end{aligned}$$

Here $\rho$, $\rho'$ and $\rho''$ independently run over the complex zeros of $\zeta(s)$. It is convenient to denote the sum of the first two terms by $\Sigma_1^\lambda(s)$; for $0 < \sigma \le 1$ it has poles at $s = 1$ and the points $\rho$. For well-behaved functions $M^\lambda(z)$, the double series defines a function $\Sigma_2^\lambda(s)$ as a limit of square partial sums. Under RH the double series with our normal $M^\lambda(z)$ is absolutely convergent for $1/2 < \sigma < 3/2$. Indeed, setting $\rho' = (1/2) + i\gamma'$, $\rho'' = (1/2) + i\gamma''$ and $s = \sigma + i\tau$, the inequalities (3.6), (3.7) show that the terms in the double series are bounded by

$$C(\lambda,\tau)(|\gamma'|+1)^{-\sigma}(|\gamma''|+1)^{-\sigma}(|\gamma'+\gamma''|+1)^{-1+2\sigma-3/2}.$$

Observing that the number of zeros $\rho = (1/2)\pm i\gamma$ with $n < \gamma \le n+1$ is $\mathcal{O}(\log n)$, cf. Titchmarsh [23], the convergence now follows from a discrete analog of the following simple lemma; cf. [16].

**Lemma 3.3.** *For real constants* $a$, $b$, $c$, *the function*

$$\phi(y,v) = (|y|+1)^{-a}(|v|+1)^{-b}(|y+v|+1)^{-c}$$

*is integrable over* $\mathbb{R}^2$ *if and only if* $a+b > 1$, $a+c > 1$, $b+c > 1$ *and* $a+b+c > 2$. *For integrability over* $\mathbb{R}^2_+$ *the condition* $a+b > 1$ *may be dropped.*

By the lemma, the part of the double sum $\Sigma_2^\lambda(s)$ in (3.8) in which $\gamma' = \operatorname{Im}\rho'$ and $\gamma'' = \operatorname{Im}\rho''$ have the same sign defines a meromorphic function for $0 < \sigma < 1$ whose only poles occur at the complex zeros of $\zeta(\cdot)$. Thus for a study of its pole-type behavior near the point $s = 1/2$, the sum $\Sigma_2^\lambda(s)$ in (3.8) may be reduced to the sum $\Sigma_3^\lambda(s)$ in which $\gamma'$ and $\gamma''$ have opposite sign. Replacing $\gamma''$ by $-\gamma''$ and using standard asymptotics for the Gamma function, it follows that the pole-type behavior of $\Sigma_3^\lambda(s)$ and $\Sigma^\lambda(s)$ as $s \searrow 1/2$ is the same as that of the reduced sum

$$(3.9)\qquad \Sigma_4^\lambda(s) = 2\pi\sum_{\gamma'>0,\,\gamma''>0}(\gamma'\gamma'')^{-s+i(\gamma'-\gamma'')/2}\,M^\lambda\{1-2s+i(\gamma'-\gamma'')\}.$$

Hence in the study of the PPC under RH, the *differences* of zeta's zeros in, say, the upper half-plane, play a key role; cf. Montgomery [20].

Formally, the poles of $\Sigma^\lambda(s)$ at the points $s = \rho$ cancel each other. Under assumption (2.4), the function $\Sigma_2^\lambda(s)$ has a meromorphic continuation to the half-plane $\{\sigma > 1/4\}$, and then there will be real cancellation; see (3.1).

## 4. Motivation of Conjecture 2.2

As we saw, numerical results make it plausible that the functions $G_{2r}(s)$ in (2.8) have an analytic continuation to the half-plane $\{\sigma > 1/4\}$. If this is correct, then by Theorem 3.1 and (3.3), assuming RH, the function

$$\begin{aligned}\Psi^\lambda(s) &\overset{\text{def}}{=} 2\sum_{0<2r\le\lambda}E(2r/\lambda)G_{2r}(s) + 4A^E\lambda\sum_\rho\frac{1}{2s-\rho}\\ (4.1)\qquad &= \Sigma^\lambda(s) - D_0(s) - \frac{2\mathbb{R}_0^E(1/2,\lambda)}{2s-1} + H_0^\lambda(s)\end{aligned}$$

will also have an analytic continuation to the half-plane $\{\sigma > 1/4\}$. In that case the quadratic pole $1/(2s-1)^2$ of $D_0(s)$ at $s = 1/2$ must be cancelled by a pole of $\Sigma^\lambda(s)$ at $s = 1/2$. By (3.8) the latter pole is the same as that of the double sum $\Sigma_2^\lambda(s)$, and hence, of $\Sigma_4^\lambda(s)$ in (3.9).

Computation suggests that the quadratic part of the pole at $s = 1/2$ in $\Sigma_4^\lambda(s)$ comes from the terms with $\gamma'' = \gamma'$. Indeed, the counting function $N(t)$ for zeta's complex zeros $(1/2) + i\gamma$ in the upper half-plane satisfies the relation

$$2\pi dN(t) = \{\log t + c_1 + \mathcal{O}(1/t)\}dt + 2\pi dS(t), \quad S(t) = \mathcal{O}(\log t);$$

cf. Titchmarsh [23]. Thus for $s = (1/2) + \delta$ with small $\delta > 0$, (3.9) with $\gamma' = \gamma'' = \gamma$ and (3.5) lead to the reduced sum

$$\begin{aligned} &2\pi \sum_{\gamma>0} \gamma^{-1-2\delta} M^\lambda(-2\delta) = 2\pi\lambda^{-2\delta} M^1(-2\delta) \int_1^\infty t^{-1-2\delta} dN(t) \\ &= (1 - 2\delta \log\lambda + \cdots)(1 + c_2\delta + \cdots)\left(\frac{1}{4\delta^2} + \frac{c_1}{2\delta} + \mathcal{O}(1)\right) \\ &= \frac{1}{4\delta^2} - \frac{\log\lambda + \mathcal{O}(1)}{2\delta} + \cdots . \end{aligned}$$

Under assumption (2.4), the residue $R(1/2, \lambda)$ of the pole of $\Sigma_2^\lambda(s)$ at $s = 1/2$ is equal to $\mathbb{R}_0^E(1/2, \lambda)$. We know that the latter quantity is $\mathcal{O}(\lambda^\varepsilon)$ as $\lambda \to \infty$. *Independently of* (2.4), the relation $c_{-1}(\lambda) = \mathcal{O}(\lambda^\varepsilon)$ for the coefficient of $1/(s-1/2)$ in the expansion

$$\Sigma_2^\lambda(s) = \frac{c_{-1}(\lambda)}{s - 1/2} + c_0(\lambda) + \cdots$$

is made *highly plausible* by the following fact: $\lambda$ occurs in the terms of the defining series for $\Sigma_2^\lambda(s)$ only as $\lambda^{\rho'+\rho''-2s}$, which is $\mathcal{O}(\lambda^\varepsilon)$ for $s \approx 1/2$.

We now turn to the likely behavior of $\Psi^\lambda(s)$ in (4.1) near the line $L = \{\sigma = 1/4\}$. Since $D_0(s)$ has quadratic poles at the points $s = 1/4$ and $\rho/2$, and no other poles on $L$, cf. (3.2), we assume that the (meromorphic continuation of the) double sum $\Sigma_2^\lambda(s)$ likewise has poles at $1/4$ and the points $\rho/2$, and nowhere else on $L$. This assumption is plausible because it is known to be true for $\lambda \le 2$, when the sum over $r$ in (4.1) is empty, so that the difference $\Sigma_2^\lambda(s) - D_0(s)$ has no poles on $L$ other than first-order poles at the points $\rho/2$. If the *heuristic argument* in the preceding paragraph has general validity, one expects that the coefficients of the pole terms of $\Sigma_2^\lambda(s)$ at $1/4$ and the points $\rho/2$ are $\mathcal{O}(\lambda^{1/2})$, or in any case $\mathcal{O}(\lambda^{(1/2)+\varepsilon})$ for every $\varepsilon > 0$. Indeed, the terms in $\Sigma_2^\lambda(s)$ contain $\lambda$ as a factor $\lambda^{\rho'+\rho''-2s} = \mathcal{O}(\lambda^{(1/2)+\varepsilon})$ for $\sigma \searrow 1/4$.

Hence by (4.1), taking the coefficient bound $\mathcal{O}(\lambda^{1/2})$ for simplicity, the sum $2\sum_{r=1}^N E(r/N)G_{2r}(s)$ should behave like

$$\frac{\mathcal{O}(\lambda^{1/2})}{(4s-1)^2} + \frac{\mathcal{O}(\lambda^{1/2})}{4s-1}$$

near the point $s = 1/4$, and like

$$\frac{\mathcal{O}(\lambda^{1/2})}{(2s-\rho)^2} - \frac{4A^E\lambda + \mathcal{O}(\lambda^{1/2})}{2s-\rho}$$

near the points $s = \rho/2$. Assuming uniformity here relative to $\rho$, and taking $\lambda = 2N$, the singular part of the average

$$\frac{1}{N}\sum_{r=1}^{N} E(r/N)G_{2r}(s)$$

for $\sigma \geq 1/4$ will have the form

$$\frac{1}{N}\sum_{r=1}^{N} G_{2r}(s) = \frac{\mathcal{O}(N^{-1/2})}{(4s-1)^2} + \frac{\mathcal{O}(N^{-1/2})}{4s-1} + \mathcal{O}(N^{-1/2})\sum_{\rho}\frac{1}{(2s-\rho)^2}$$
$$- \{4A^E + \mathcal{O}(N^{-1/2})\}\sum_{\rho}\frac{1}{2s-\rho} + H^N(s).$$

The remainder $H^{N,E}(s)$ will have good boundary behavior as $\sigma \searrow 1/4$, and it contains $1/(2N)$ times the remainder $H_0^{\lambda,E}(s)$ from Theorem 3.1 with $\lambda = 2N$. Now taking $E(\nu)$ close to the function which is equal to 1 on $[-1,1]$ and 0 elsewhere, one is led to Conjecture 2.2.

This conjecture, finally, makes the conjectured Approximation 2.1 plausible through formal Fourier inversion (2.10).

## 5. From Approximation 2.1 to Approximation 1.1 via Approximation 5.2

After motivating Approximation 2.1 for averages of functions $\psi_{2r}(x)$, we turn to a corresponding approximation involving the functions $\theta_{2r}(x)$. For large $N$ and $x > 2N$, cf. (2.5),

$$\frac{1}{N}\sum_{r=1}^{N}\theta_{2r}(x) = \frac{1}{N}\sum_{r=1}^{N}\psi_{2r}(x) - \frac{2}{N}\sum_{r=1}^{N}\theta^*_{2r}(x^{1/2})$$
$$- \frac{1}{N}\sum_{r=1}^{N} 4rC_{2r}\log\log x + \mathcal{O}(x^{1/3}\log^2 x) + o(N). \tag{5.1}$$

According to the Bateman–Horn conjecture [2], [3], applied to the special case of prime pairs $(p, p^2 \pm 2r)$, there should be specific positive constants $2C^*_{2r} = C^{[2]}_{2r} + C^{[-2]}_{2r}$ such that

$$\theta^*_{2r}(x) = \{2C^*_{2r} + o(1)\}x \quad \text{as } x \to \infty. \tag{5.2}$$

Here there is no need to study the Bateman–Horn constants in detail; our only concern will be their mean value (apparently equal to one).

**Conjecture 5.1.** *For $x \to \infty$ one has*

$$\frac{1}{N}\sum_{r=1}^{N}\theta^*_{2r}(x) = \{2 + o(N^{-1/2})\}x. \tag{5.3}$$

The motivation for Conjecture 5.1 is given by Conjecture 5.3 below. Combining Conjecture 5.1 with formula (5.1) and Approximation 2.1, one obtains the (conjectured)

**Approximation 5.2.** For large $N$ and $x$ much larger than $N$, one has

$$\begin{aligned}\frac{1}{N}\sum_{r=1}^{N}\{\theta_{2r}(x)-2C_{2r}x\} &= -\{4+\mathcal{O}(N^{-1/2}\log x)\}\sum_{\rho}\frac{x^{\rho}}{\rho}-\{1+o(1)\}N \\ &- \{4+\mathcal{O}(N^{-1/2}\log x)\}x^{1/2}-\frac{1}{N}\sum_{r=1}^{N}4rC_{2r}\log\log x+\mathcal{O}(x^{1/3}\log^2 x).\end{aligned} \tag{5.4}$$

To go from here to Approximation 1.1 we use the operation represented by $\int_2^x (1/\log^2 t)\, d\cdots$. The sum on the left of (5.4) then becomes the sum on the left of (1.5), cf. (2.1):

$$\frac{1}{N}\sum_{r=1}^{N}\int_2^x \frac{1}{\log^2 t}\, d\big\{\theta_{2r}(t)-2C_{2r}t\big\} = \frac{1}{N}\sum_{r=1}^{N}\big\{\pi_{2r}(x)-2C_{2r}\mathrm{li}_2(x)\big\}.$$

In the application of the same operation to the right-hand side of (5.4) it is assumed that contributions due to derivatives of the $\mathcal{O}$-terms can be neglected. Ignoring the $\log\log x$-term for a moment, the right-hand side of (5.4) then gives the right-hand side of (1.5) for any $\delta \le 1$.

The $\log\log x$-term (with its minus sign) ultimately leads to a contribution

$$-\frac{1}{N}\sum_{r=1}^{N}4rC_{2r}\int_{2r}^{x}\frac{1}{\log^2 t}\, d\log\log t \approx -\frac{N}{\log^2 2N}. \tag{5.5}$$

To assess its effect on $\Delta_N(x)$ in (1.7), one still has to divide by $\mathrm{li}_2(x^{1/2}) \sim 4x^{1/2}/\log^2 x$. The result $\overline{\Delta}_N(x)$ in (1.8) will be small when $x$ is much larger than $N^2$.

In support of Conjecture 5.1 we proceed with a conjecture involving the related Dirichlet series

$$D^*_{2r}(s) = \sum_{p,\, p^2\pm 2r \text{ prime}} \frac{\log^2 p}{p^{4s}} = \int_1^{\infty} x^{-4s}d\theta^*_{2r}(x) \qquad (\sigma > 1/4). \tag{5.6}$$

**Conjecture 5.3.** *For $\sigma > 1/4$ and $N \to \infty$, one has*

$$\frac{1}{N}\sum_{r=1}^{N} D^*_{2r}(s) = \frac{2+o(N^{-1/2})}{4s-1} + H_2^N(s), \tag{5.7}$$

*with an analytic function $H_2^N(s)$ that has good boundary behavior as $\sigma \searrow 1/4$.*

The arguments supporting Conjecture 5.3 are similar to those given for Conjecture 2.2. For $1/4 < \sigma < 1/2$ one may write

$$\begin{aligned}\Phi^{\lambda}_{1,2}(s) &\stackrel{\text{def}}{=} D^*_0(s) + 2\sum_{0<2r\le\lambda} E(2r/\lambda)D^*_{2r}(s) \\ &= \frac{2A^E\lambda}{4s-1} + \frac{\zeta'(2s)}{\zeta(2s)}J(s,s) + \sum_{\rho}\frac{1}{2}\Gamma\{(\rho/2)-s\}J(\rho/2,s) + H_3^{\lambda}(s),\end{aligned} \tag{5.8}$$

where $D^*_0(s) = \sum_p (\log^2 p)/p^{4s}$ and $H_3^{\lambda}(s)$ is analytic for $1/4 \le \sigma < 1/2$. The functions $J(s,s)$ and $J(\rho/2,s)$ are analytic for $1/4 \le \sigma < 1/2$; cf. (8.4) in the Appendix. Hence, formally the poles at the points $s = \rho/2$ in the combination of $J$-terms in (5.8) will cancel each other. However, one constituent of $J(\rho/2,s)$ is an infinite series. It leads to a repeated series $\Sigma^{\lambda}_{2,2}(s)$ when it is substituted into the sum over $\rho$ in (5.8):

$$\Sigma^{\lambda}_{2,2}(s) = \frac{1}{2}\sum_{\rho}\Gamma\{(\rho/2)-s\}\sum_{\rho'}\Gamma(\rho'-s) \times M^{\lambda}\{\rho'+(\rho/2)-2s\}\cos\{\pi(\rho'-\rho/2)/2\}.$$

This series is absolutely convergent only for $3/8 < \sigma < 1/2$; for $1/4 < \sigma \le 3/8$ the sum over $\rho = (1/2) + i\gamma$ has to be interpreted as a limit of partial sums $\sum_{|\gamma|\le B}$ as $B \to \infty$.

In view of the similarity of the Hardy–Littlewood conjecture and our case of the Bateman–Horn conjecture, it is reasonable to suppose that the differences

$$G^*_{2r}(s) = D^*_{2r}(s) - \frac{2C^*_{2r}}{4s-1}$$

have an analytic continuation to the half-plane $\{\sigma \ge 1/4\}$. If that is correct, the combination of the $J$-terms in (5.8) truly has no poles at the points $s = \rho/2$. The repeated sum $\Sigma^{\lambda}_{2,2}(s)$ then would have an analytic continuation to the strip $1/4 \le \sigma < 1/2$, except for a pole at $s = 1/4$. The quadratic pole $1/(4s-1)^2$ of $D^*_0(s)$ at $s = 1/4$ would be cancelled by the quadratic part of the pole of $\Sigma^{\lambda}_{2,2}(s)$ there. Finally, the residue of the pole of $\Sigma^{\lambda}_{2,2}(s)$ at $s = 1/4$ would be $\mathcal{O}(\lambda^{(1/4)+\varepsilon})$ by heuristics as in Section 4. Hence by (5.8), the residue at $s = 1/4$ of

$$\frac{1}{N}\sum_{r=1}^{N} E(r/N)D^*_{2r}(s) \quad \text{would be} \quad \frac{1}{2}A^E + o(N^{-1/2}),$$

thus leading to (5.7) when $E(\nu)$ is taken close to 1 on $(-1, 1)$.

The proof of (5.8) is similar to that of (3.1) described in the Appendix. Here, one would start with the integral obtained from (8.2) through replacement of one of the quotients $\zeta'(\cdot)/\zeta(\cdot)$ by $\zeta'(2\,\cdot)/\zeta(2\,\cdot)$.

# PART II. NUMERICAL RESULTS AND GRAPHS

## 6. Comparing averages of functions $\omega_{2r}(x)$ with $\omega(x)$

Which of the two terms on the right-hand side of (1.5), in the conjectured Approximation 1.1, is larger? One may write

$$\sum_{\rho}\rho\,\mathrm{li}_2(x^{\rho}) = \frac{x^{1/2}}{\log^2 x}\sum_{\rho}\frac{x^{\rho-1/2}}{\rho} + \mathcal{O}\Big(\frac{x^{1/2}}{\log^3 x}\Big),$$

$$\mathrm{li}_2(x^{1/2}) = 4\,\frac{x^{1/2}}{\log^2 x} + \mathcal{O}\Big(\frac{x^{1/2}}{\log^3 x}\Big).$$

Combining the terms with $\rho = (1/2) \pm i\gamma$ for $\gamma > 0$, one obtains

$$T(x) \overset{\text{def}}{=} \sum_{\rho}\frac{x^{\rho-1/2}}{\rho} = \sum_{\gamma>0}\frac{\cos(\gamma\log x) + 2\gamma\sin(\gamma\log x)}{\gamma^2+1/4}. \tag{6.1}$$

Thus relation (1.5) takes the form

$$
\begin{aligned}
Q_N(x) &\stackrel{\text{def}}{=} \frac{\sum_{r=1}^{N} \omega_{2r}(x)}{N\mathrm{li}_2(x^{1/2})} \\
&= -[\{1+\mathcal{O}(N^{-1/2}\log x)\}T(x) + 1 + \mathcal{O}(N^{-1/2}\log x)].
\end{aligned}
\tag{6.2}
$$

It is interesting that Riemann's formula (1.9) leads to a combination similar to the right-hand side of (6.2). Indeed, assuming RH one may write

$$
\frac{2\omega(x)}{\mathrm{li}(x^{1/2})} = -\big[\{1+\mathcal{O}(1/\log x)\}T(x) + 1 + \mathcal{O}(1/\log x)\big]. \tag{6.3}
$$

Littlewood's work [18], cf. Ingham [13], implies that the function $T(x)$ oscillates unboundedly. More precisely, he showed that there are constants $c$, $c' > 0$ and arbitrarily large $x$, $x'$ such that

$$
T(x) < -c\log\log\log x, \quad T(x') > c'\log\log\log x'.
$$

However, $\pi(x)$ becomes larger than $\mathrm{li}(x)$, that is, $\omega(x) > 0$, only for certain very large $x$. The first such number is associated with the name of Skewes; cf. te Riele [22], and Bays and Hudson [4]. Under RH one has $T(x) = \mathcal{O}(\log^2 x)$, and Kotnik [17] made it plausible that $T(x) = \mathcal{O}(\log x)$. He also graphed the function $\omega(x)(\log x)/x^{1/2}$, cf. (6.3), for $x \le 10^{14}$. On a logarithmic scale, his Figure 1 shows rapid oscillations of amplitude greater than 1/2.

The Skewes story seems to have no analog for prime pairs; cf. Brent [6]. Here we focus on the case of twin primes. Nicely [21] has counted prime twins up to $x = 10^{16}$. His table uses steps $10^k$ from $1 \cdot 10^k$ through $9 \cdot 10^k$ for $k = 1, 2, \ldots, 12$. From there on the steps are $10^{12}$. Nicely's table shows that for $x$ going to $10^{16}$, the quantity $|\omega_2(x)|$ often becomes a good deal larger than $\mathrm{li}_2(x^{1/2})$. His table implies 16 *sign changes* of $\omega_2(x)$ [which is minus his entry $\delta_2(x)$]. The first occurs between $10^6$ and $2 \cdot 10^6$, the last between $7.5 \cdot 10^{13}$ and $7.6 \cdot 10^{13}$. Although $\omega_2(x)$ oscillates, it then remains positive until the end of Nicely's table.

In a preprint on a 'Skewes number for twin primes', Marek Wolf [25] analyzed the sign changes in $\omega_2(x)$ up to $2^{42} \approx 4.4 \cdot 10^{12}$. He found the first one at the twin with $p = 1369391$. A table in his preprint lists the number of sign changes up to $2^k$ for $k = 22, 23, \ldots, 42$. Wolf found 90355 sign changes up to $2^{42}$. He found none between $2^{22}$ and $2^{25}$, none between $2^{28}$ and $2^{31}$, and none between $2^{37}$ and $2^{39}$.

In our range of $x$, the values of $|T(x)|$ are smaller than one. In particular,

$$
\begin{aligned}
&T(10^6) \approx 0.41156, \quad T(10^8) \approx 0.17554, \\
&T(10^{10}) \approx -0.42122, \quad T(10^{12}) \approx -0.04014.
\end{aligned}
\tag{6.4}
$$

These values were computed with the aid of von Mangoldt's formula (1.10), by which (for $x > 1$ and $x$ not a prime power)

$$
T(x) = x^{-1/2}\big\{x - \psi(x) - \log(2\pi) - (1/2)\log(1-x^{-2})\big\}. \tag{6.5}
$$

The function $\psi(x) = \sum_{p^m \le x} \log p$ was computed by summing the values of $\lfloor \log_p x \rfloor \log p$ for all the primes $p \le x$ (generated with the sieve of Eratosthenes). Here, $\lfloor \log_p x \rfloor$ is the exponent of $p$ in the highest power of $p$ not exceeding $x$. The values of $T(x)$ given in (6.4) were computed with an accuracy of at least 5 decimal digits. We were using Fortran double precision floating point arithmetic which works with an accuracy of about 15 decimal digits, but precision is lost as $x$ grows when (6.5) is used to compute $T(x)$. To illustrate this, we found that

| $2N$ | $S_N/C_2$ | $\Pi_N(10^6)$ | $\Delta_N(10^6)$ | $\Pi_N(10^8)$ | $\Delta_N(10^8)$ |
|---|---|---|---|---|---|
| 100 | 73.6377551 | 605087 | +0.09722 | 32417440 | -0.08872 |
| 200 | 149.3252708 | 1226667 | -0.02199 | 65739481 | +0.03162 |
| 300 | 225.4407734 | 1851433 | -0.12785 | 99245855 | -0.09833 |
| 400 | 300.3132204 | 2465581 | -0.23344 | 132202659 | -0.23013 |
| 500 | 376.0636735 | 3086695 | -0.32860 | 165551273 | -0.18188 |
| 600 | 452.4693143 | 3714028 | -0.31371 | 199186203 | -0.19507 |
| 700 | 527.3827110 | 4328507 | -0.34805 | 232164862 | -0.18926 |
| 800 | 603.4536365 | 4951873 | -0.42140 | 265651152 | -0.21737 |
| 900 | 679.4011178 | 5574196 | -0.48004 | 299079601 | -0.28690 |
| 1000 | 754.4223630 | 6188960 | -0.52230 | 332105577 | -0.27582 |
| 2000 | 1511.5853400 | 12391586 | -0.78001 | 665435604 | -0.16751 |
| 3000 | 2269.6853566 | 18597363 | -0.95390 | 999175096 | -0.14446 |
| 4000 | 3026.0445409 | 24783891 | -1.11135 | 1332114654 | -0.23565 |
| 5000 | 3783.8474197 | 30975067 | -1.28953 | 1665693721 | -0.28111 |

TABLE 2. Values of $S_N/C_2$, $\Pi_N(10^6)$, $\Delta_N(10^6)$, $\Pi_N(10^8)$, and $\Delta_N(10^8)$

$\psi(10^{12}) = 1000000040136.76$, so that in the difference $10^{12} - \psi(10^{12}) = -40136.76$ only about seven digits are still correct and $T(10^{12}) = -0.04013860$.

Alternative computations based on formula (6.1) and the first two million values of $\gamma$ gave the values $T(10^6) \approx 0.41276$, $T(10^8) \approx 0.17469$, $T(10^{10}) \approx -0.41944$, and $T(10^{12}) \approx -0.04010$, i.e., an accuracy of only about 3 decimal digits.

## 7. TESTING THE CONJECTURED APPROXIMATION 1.1

In the following we will consider the aggregate

$$\Pi_N(x) \stackrel{\text{def}}{=} \pi_2(x) + \pi_4(x) + \cdots + \pi_{2N}(x) \tag{7.1}$$

for certain large values of $N$ and $x$. Setting

$$S_N = C_2 + C_4 + \cdots + C_{2N}, \tag{7.2}$$

cf. (3.4), we compare $\Pi_N(x)$ with $2S_N \operatorname{li}_2(x) = (S_N/C_2)L_2(x)$ (cf. Section 2). In view of the conjectured Approximation 1.1, the difference is divided by $N \operatorname{li}_2(x^{1/2})$ to obtain the quotient

$$\frac{\Pi_N(x) - (S_N/C_2)L_2(x)}{N \operatorname{li}_2(x^{1/2})} = \frac{\sum_{r=1}^N \omega_{2r}(x)}{N \operatorname{li}_2(x^{1/2})} = Q_N(x); \tag{7.3}$$

cf. (6.2). For large $N$ the quotient should have the form

$$-\{1 + \mathcal{O}(N^{-1/2} \log x)\}T(x) - \{1 + \mathcal{O}(N^{-1/2} \log x)\}.$$

Ignoring the $\mathcal{O}$-terms, we will compare $Q_N(x)$ with $-T(x) - 1$, setting

$$Q_N(x) + T(x) + 1 = \Delta_N(x). \tag{7.4}$$

Tables 2, 3 give results for $x = 10^6$, $10^8$, $10^{10}$, $10^{12}$. The values $S_N/C_2$ were obtained by computing $C_{2r}/C_2$ from (1.3) and adding. For the values of $\Pi_N(x)$ we

| $2N$ | $\Pi_N(10^{10})$ | $\Delta_N(10^{10})$ | $\Pi_N(10^{12})$ | $\Delta_N(10^{12})$ |
|---|---|---|---|---|
| 100 | 2018498733 | +0.23101 | 137743459486 | -0.22449 |
| 200 | 4093181354 | +0.19981 | 279320931774 | -0.52374 |
| 300 | 6179575427 | +0.04646 | 421698995095 | -0.60678 |
| 400 | 8231900717 | -0.00307 | 561752066806 | -0.47345 |
| 500 | 10308323520 | +0.09461 | 703447298670 | -0.52336 |
| 600 | 12402663153 | +0.00891 | 846368266787 | -0.46665 |
| 700 | 14456137134 | +0.06512 | 986498011024 | -0.37686 |
| 800 | 16541312091 | +0.03187 | 1128792535379 | -0.48827 |
| 900 | 18623097684 | -0.00710 | 1270856645797 | -0.39850 |
| 1000 | 20679532323 | +0.04311 | 1411187901897 | -0.41454 |
| 2000 | 41434008965 | -0.14700 | 2827502930522 | -0.31142 |
| 3000 | 62214267139 | -0.14273 | 4245571295213 | -0.21865 |
| 4000 | 82946817735 | -0.13473 | 5660383932743 | -0.12392 |
| 5000 | 103718886923 | -0.15324 | 7077896171945 | -0.12569 |

TABLE 3. Values of $\Pi_N(10^{10})$, $\Delta_N(10^{10})$, $\Pi_N(10^{12})$, and $\Delta_N(10^{12})$

added columns of numbers $\pi_{2r}(x)$. We next computed $Q_N(x)$ from (7.3). Here we used the approximations

$$L_2(10^6) \approx 8248.0297, \quad L_2(10^8) \approx 440367.7942,$$
$$L_2(10^{10}) \approx 27411416.53, \quad L_2(10^{12}) \approx 1870559866.82$$

and

$$\mathrm{li}_2(10^3) \approx 34.6851, \quad \mathrm{li}_2(10^4) \approx 162.2412,$$
$$\mathrm{li}_2(10^5) \approx 945.75959, \quad \mathrm{li}_2(10^6) \approx 6246.9757.$$

The table entries $\Delta_N(x)$ are based on (7.4) and the approximations for $T(x)$ in (6.4).

In Figures 1–4 we show plots of $\Delta_N(x)$ as a function of $N$ ($50 \le 2N \le 5000$), for $x = 10^6$, $10^8$, $10^{10}$, and $10^{12}$. We have omitted the function values for $2 \le 2N \le 48$ since they very much dominate (and are atypical for) the other function values. In Figures 1 and 2 we compare $\Delta_N(x)$ with the function $\overline{\Delta}_N(x) = -(2N\log^2 x)/(8x^{1/2}\log^2 2N)$ as defined in (1.8).

We have made, but not given here, plots of $\Delta_N(x)$ for several other values of $x$. E.g., for $x = 10^{11}$ and $2N = 1000, 2000, 3000, 4000, 5000$, we found: $\Delta_N(x) = -0.229, -0.072, -0.034, +0.004$, and $-0.034$, respectively (compare these values with the corresponding values for $x = 10^{10}$ and $x = 10^{12}$ in Table 3).

Figures 5 and 6 show plots of $\Delta_N(x)$ as a function of $x$ ($6 \le \log_{10} x \le 12$), for $N = 400$ and $N = 2500$, respectively. The plots have been constructed by connecting the values of $\Delta_N(x)$ for $x = 10^6$ and for $x = i \times 10^j$, $j = 6, 7, \ldots, 11$ and $i = 1, 2, \ldots, 10$ by straight lines. The different behaviours of the plots of $\Delta_{400}(x)$ and $\Delta_{2500}(x)$ may reflect the influence of the unknown $\mathcal{O}(N^{-1/2}\log x)$–terms, which were neglected in the derivation of the error function $\Delta_N(x)$ from (1.5).

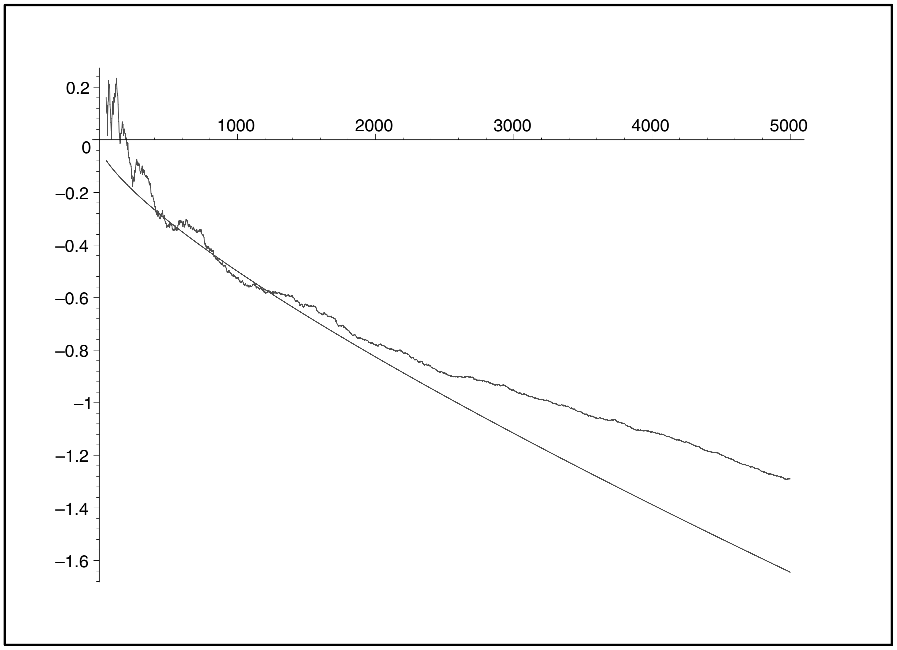


FIGURE 1. $\Delta_N(10^6)$ compared with $\overline{\Delta}_N(10^6)$ for $50 \le 2N \le 5000$

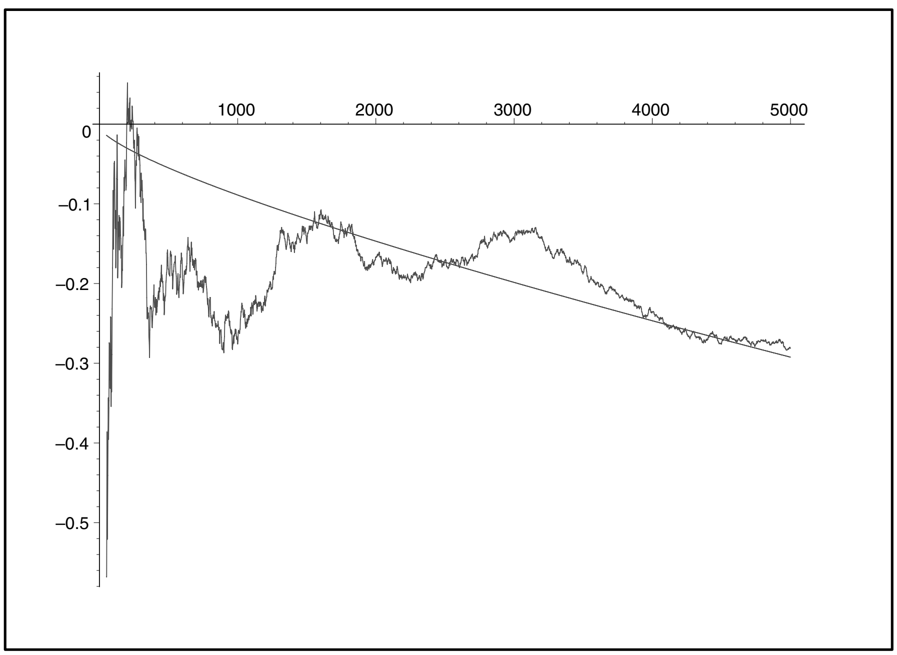


FIGURE 2. $\Delta_N(10^8)$ compared with $\overline{\Delta}_N(10^8)$ for $50 \le 2N \le 5000$

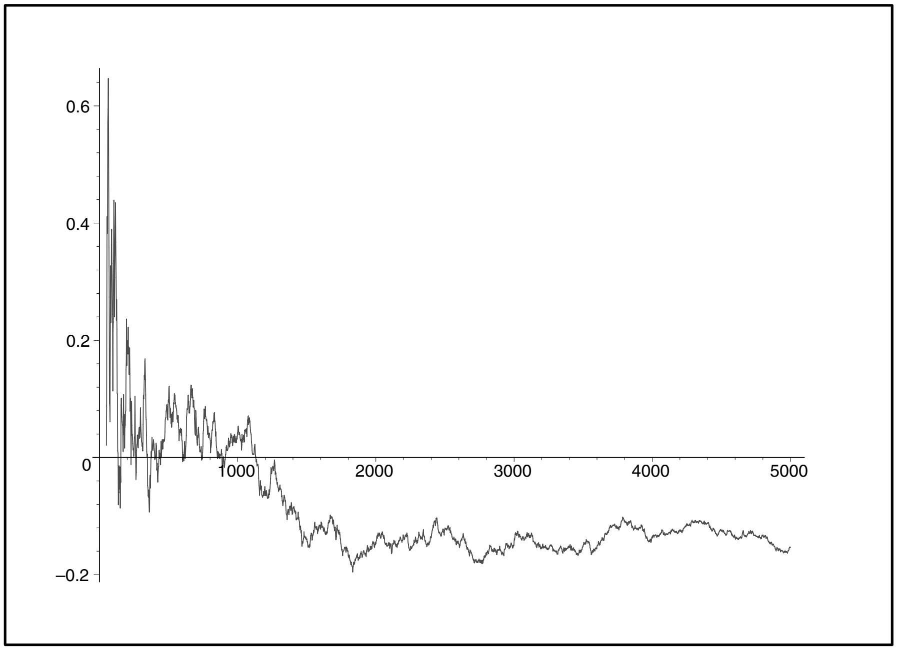


FIGURE 3. $\Delta_N(10^{10})$ for $50 \le 2N \le 5000$

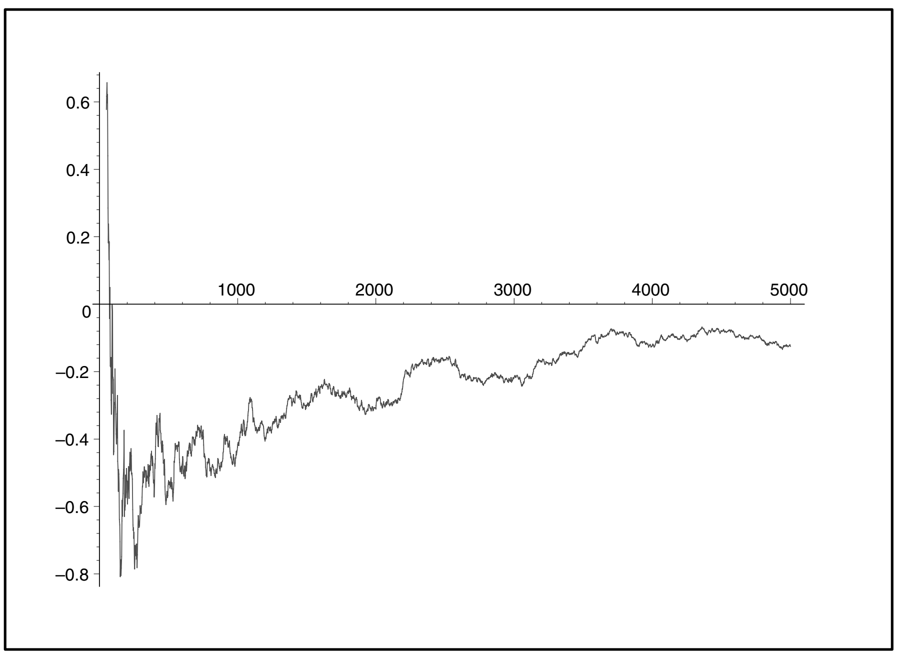


FIGURE 4. $\Delta_N(10^{12})$ for $50 \le 2N \le 5000$

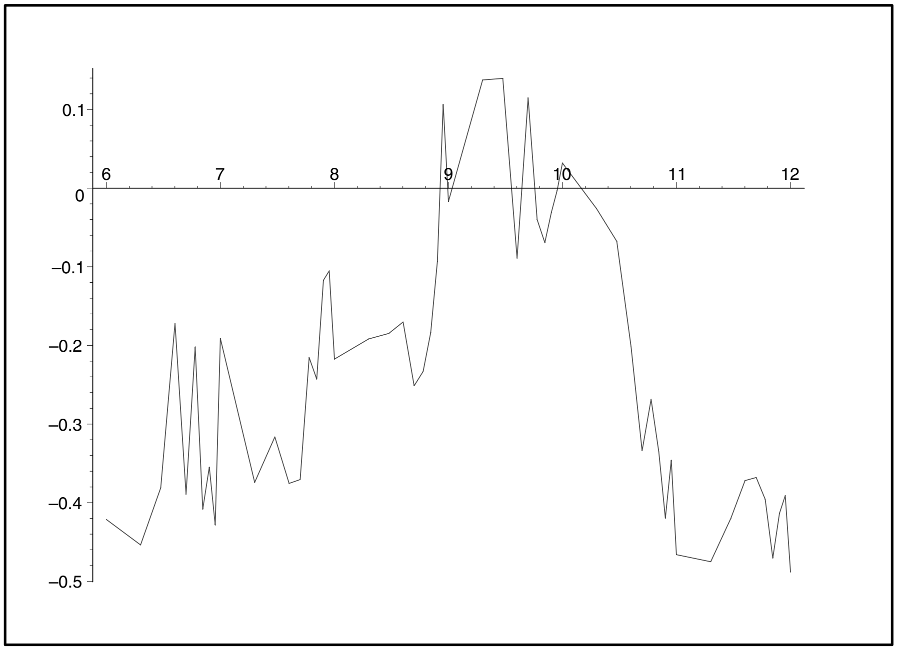


FIGURE 5. $\Delta_{400}(x)$ for $6 \leq \log_{10} x \leq 12$

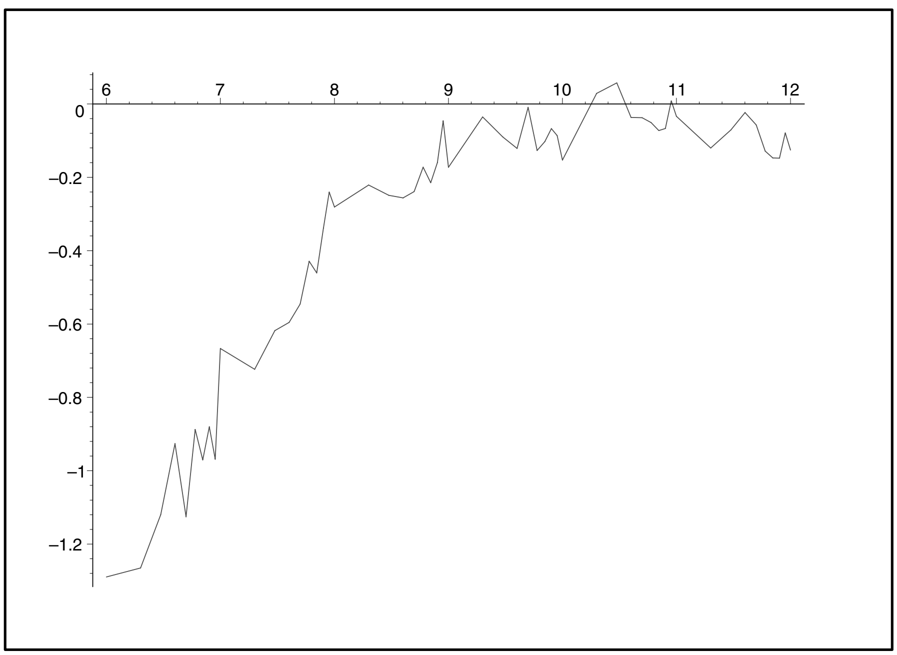


FIGURE 6. $\Delta_{2500}(x)$ for $6 \leq \log_{10} x \leq 12$

## APPENDIX

### 8. Outline of the proof of Theorem 3.1

Extending an idea that goes back to Arenstorf [1], cf. [16], one is led to a representation for $E\{(\alpha-\beta)/\lambda\}$ (Section 3) by an absolutely convergent repeated complex integral in which $\alpha>0$ and $\beta>0$ occur separately:

$$E\{(\alpha-\beta)/\lambda\} = \frac{1}{(2\pi i)^2}\int_{L(c,B)}\Gamma(z)\alpha^{-z}dz\int_{L(c,B)}\Gamma(w)\beta^{-w} \times M^\lambda(z+w)\cos\{\pi(z-w)/2\}\,dw. \tag{8.1}$$

Here the path $L(c,B)=L(c_1,c_2,B)$ in the $z=x+iy$ plane is taken to be of the form

$$L(c,B) = \begin{cases} \text{the half-line} & \{x=c_1,\ -\infty<y\le -B\} \\ +\text{ the segment} & \{c_1\le x\le c_2,\ y=-B\} \\ +\text{ the segment} & \{x=c_2,\ -B\le y\le B\} \\ +\text{ the segment} & \{c_2\ge x\ge c_1,\ y=B\} \\ +\text{ the half-line} & \{x=c_1,\ B\le y<\infty\}, \end{cases}$$

and similarly for the $w=u+iv$ plane. For $-1/2<c_1<0<c_2<1/2$, say, and arbitrary $B>0$, the absolute convergence of the repeated integral in (8.1) follows from (3.6), (3.7) and Lemma 3.3.

For the verification of formula (8.1) one may write $\cos\alpha t$ as a complex (inverse) Mellin integral involving $\Gamma(z)$:

$$\cos\alpha t = \frac{1}{2\pi i}\int_{L(c,B)}\Gamma(z)(\alpha t)^{-z}\cos(\pi z/2)dz,$$

and $\cos\beta t$ as such an integral involving $\Gamma(w)$. Multiplying the two, doing the same with sines and adding, one obtains a repeated complex integral for $\cos(\alpha-\beta)t$:

$$\cos(\alpha-\beta)t = \frac{1}{(2\pi i)^2}\int_{L(c,B)}\Gamma(z)\alpha^{-z}t^{-z}dz \times\int_{L(c,B)}\Gamma(w)\beta^{-w}t^{-w}\cos\{\pi(z-w)/2\}dw.$$

This integral is multiplied by $\hat{E}^\lambda(t)$; integration over $0<t<\infty$ and use of (3.5) then gives the desired result.

Formula (8.1) leads to the following integral for $\Phi^\lambda(s)$ in (3.1), modulo a function $H^\lambda(s)$ that turns out to be analytic for $\sigma>0$:

$$\Phi^\lambda(s) = \frac{1}{(2\pi i)^2}\int_{L(c,B)}\Gamma(z)\,\frac{\zeta'(z+s)}{\zeta(z+s)}\,dz\int_{L(c,B)}\Gamma(w)\,\frac{\zeta'(w+s)}{\zeta(w+s)} \times M^\lambda(z+w)\cos\{\pi(z-w)/2\}dw + H^\lambda(s). \tag{8.2}$$

For verification one introduces the Dirichlet series $-\sum\Lambda(k)k^{-z-s}$ for the quotient $(\zeta'/\zeta)(z+s)$ and $-\sum\Lambda(l)l^{-w-s}$ for $(\zeta'/\zeta)(w+s)$. One then integrates term by term, initially taking $\sigma>1+|c_1|$ to ensure uniform convergence. The result

$$\sum_{k,l=1}^{\infty}\frac{\Lambda(k)\Lambda(l)}{k^s l^s}\,E\{(k-l)/\lambda\}$$

differs from $\Phi^\lambda(s)$ in (3.1) by

$$H^\lambda(s) \stackrel{\text{def}}{=} 2 \sum_{0<2r\le\lambda} \sum_{n=1}^{\infty} \Lambda(n)\Lambda(n+2r)\left\{\frac{1}{n^{2s}} - \frac{1}{n^s(n+2r)^s}\right\}E(2r/\lambda)$$
$$- 2 \sum_{0<2r-1\le\lambda} \sum_{n=1}^{\infty} \frac{\Lambda(n)\Lambda(n+2r-1)}{n^s(n+2r-1)^s}\, E\{(2r-1)/\lambda\}.$$

The first expression on the right is analytic for $\sigma > 0$, and so is the second: if $n$ and $n+2r-1$ are both prime powers, one of them must be a power of 2. One may verify that $H^\lambda(s)$ is the Mellin transform of a function $h^\lambda(x)$ which is $\mathcal{O}(\lambda^3 x^{-1}\log^2 x + \lambda \log^2 x)$ for $x > \lambda$.

Analytic continuation shows that under RH, one may take paths $L(c,B)$ in (8.2) with $c_1 = -\eta$ and $c_2 = (1/2) - \eta$, where $0 < \eta < 1/2$. Thus the integral representation may be used for $s = \sigma + i\tau$ with $\sigma > (1/2) + \eta$ and $|\tau| < B$; cf. [16]. Additionally requiring $\sigma < 1$, we now move the paths $L(c,B)$ across the poles at the points $1-s$, 0 and $\rho - s$ to lines $L(d)$, given by $x$ or $u$ equal to $d = -(1/2) + \eta$. Here $\rho$ runs over the complex zeros of $\zeta(\cdot)$. The moves may be justified by Cauchy's theorem and the estimates in Lemma 3.3. On the relevant vertical lines, $(\zeta'/\zeta)(Z)$ only grows logarithmically in $Y$, and auxiliary horizontal segments can be suitably chosen between zeta's complex zeros. Cf. Titchmarsh [23].

First moving the $w$-path one obtains a new repeated integral, along with a single 'residue-integral'. It is convenient to write the latter in the form

$$\frac{1}{2\pi i}\int_{L(c,B)} \Gamma(z)\,\frac{\zeta'(z+s)}{\zeta(z+s)}\, J(z+s,s)dz, \tag{8.3}$$

where by the residue theorem,

$$\begin{aligned} J(z+s,s) &= -\Gamma(1-s)M^\lambda(z+1-s)\cos\{\pi(z+s-1)/2\} \\ &\quad + \frac{\zeta'(s)}{\zeta(s)}\, M^\lambda(z)\cos(\pi z/2) \\ &\quad + \sum_\rho \Gamma(\rho - s)M^\lambda(z+\rho-s)\cos\{\pi(z+s-\rho)/2\}. \end{aligned} \tag{8.4}$$

Next move the $z$-path $L(c,B)$ in the new repeated integral and the $z$-path in the single integral to the line $L(d)$. Thus we obtain another repeated integral, now involving two paths $L(d)$, and a single integral with path $L(d)$, where $d = -(1/2)+\eta$. Varying $\eta \in (0,1/2)$, one sees that the new integrals represent analytic functions for $-1/2 < \sigma < 1$. The operation on the repeated integral produces a harmless residue, namely, another copy of the single integral with path $L(d)$. However, the operation on the single integral yields the following residue:

$$-\Gamma(1-s)J(1,s) + \frac{\zeta'(s)}{\zeta(s)}\, J(s,s) + \sum_{\rho'} \Gamma(\rho'-s)J(\rho',s), \tag{8.5}$$

where $\rho'$ runs over the complex zeros of $\zeta(\cdot)$. Working out this residue with the aid of (8.4), one obtains nine terms. Four of these supply the sum $\Sigma^\lambda(s)$ of (3.8) in

(3.1). The remaining five terms combine into the sum

$$(8.6)\quad \begin{aligned} V^{\lambda}(s) &\stackrel{\text{def}}{=} \Gamma^2(1-s)M^{\lambda}(2-2s) - 2\Gamma(1-s)\,\frac{\zeta'(s)}{\zeta(s)}\,M^{\lambda}(1-s)\sin(\pi s/2) \\ &\quad - 2\Gamma(1-s)\sum_{\rho}\Gamma(\rho-s)M^{\lambda}(1+\rho-2s)\sin(\pi\rho/2). \end{aligned}$$

Here, the apparent poles at the points $s=\rho$ cancel each other. The first term provides the important pole-term at the point $s=1/2$ in (3.1). Indeed, by the pole-type behavior of $M^{\lambda}(Z)$ at the point $Z=1$ (Section 2),

$$(8.7)\quad \Gamma^2(1-s)M^{\lambda}(2-2s) = \frac{2A^E\lambda}{2s-1} + H_4^{\lambda}(s),$$

where $H_4^{\lambda}(s)$ is analytic for $-1/2<\sigma<1$. The final term in (8.6) generates simple poles at the points $s=\rho/2$. A short computation shows that the residues at those poles are all equal to $-2A^E\lambda$, thus leading to the term $-4A^E\lambda\sum_{\rho}1/(2s-\rho)$ in (3.1).

To round out the proof of Theorem 3.1 we evaluate the inverse Mellin transform of $V^{\lambda}(s)$ in (8.6). Taking $c=(1/2)+\eta$ and $x>\lambda$, and moving the path to the left, one finds that

$$\frac{1}{2\pi i}\int_{L(c)} V^{\lambda}(s)\,x^{2s}\,\frac{ds}{s} = 2A^E\lambda\,x - 4A^E\lambda\sum_{\rho}\frac{x^{\rho}}{\rho} + g(x,\lambda).$$

Here for large $\lambda$ and $x/\lambda\to\infty$,

$$g(x,\lambda) = -\{1+o(1)\}\lambda^2\int_0^1 E(\nu)\nu\,d\nu \approx -\lambda^2/2$$

when $E(\nu)$ is close to 1 on $(-1,1)$.

By our hypothesis the double sum $\Sigma_2^{\lambda}(s)$ in (3.8) generates a pole at $s=1/2$ with residue $R^E(1/2,\lambda)$ given by (3.3). Through Mellin inversion this becomes $2R^E(1/2,\lambda)x$; when added to $2A^E\lambda x$, it gives the principal part

$$2\sum_{0<2r\le\lambda}E(2r/\lambda)\cdot 2C_{2r}x \quad\text{of}\quad 2\sum_{0<2r\le\lambda}E(2r/\lambda)\psi_{2r}(x).$$

Other conjectural contributions of $\Sigma_2^{\lambda}(s)$ were discussed in Section 4.

## References


[1] R.F. Arenstorf, *There are infinitely many prime twins.* Available on the internet, at http://arxiv.org/abs/math/0405509v1. Article posted May 26, 2004; withdrawn June 9, 2004.

[2] P. T. Bateman and R. A. Horn, *A heuristic asymptotic formula concerning the distribution of prime numbers.* Math. Comp. **16** (1962), 363–367. MR0148632 (26:6139)

[3] P. T. Bateman and R. A. Horn, *Primes represented by irreducible polynomials in one variable.* Proc. Sympos. Pure Math., vol. VIII, pp. 119–132. Amer. Math. Soc., Providence, R.I., 1965. MR0176966 (31:1234)

[4] C. Bays and R. Hudson, *A new bound for the smallest $x$ with $\pi(x) > \mathrm{li}(x)$.* Math. Comp. **69** (2000), 1285–1296. MR1752093 (2001c:11138)

[5] E. Bombieri and H. Davenport, *Small differences between prime numbers.* Proc. Roy. Soc. Ser. A **293** (1966), 1–18. MR0199165 (33:7314)

[6] R. P. Brent, *Irregularities in the distribution of primes and twin primes.* Math. Comp. **29** (1975), 43–56. MR0369287 (51:5522)

[7] F. J. van de Bult, *Counts of prime pairs.* Private communication including spreadsheet attachment tot10-tot.csv, February 2007.

[8] H. Davenport, *Multiplicative number theory.* (Third edition, revised by H. L. Montgomery.) Graduate Texts in Math., 74. Springer, New York, 2000. MR1790423 (2001f:11001)

[9] H. M. Edwards, *Riemann's zeta function.* Academic Press, New York, 1974. Reprinted by Dover Publications, Mineola, N.Y., 2001. MR1854455 (2002g:11129)

[10] J. B. Friedlander and D. A. Goldston, *Some singular series averages and the distribution of Goldbach numbers in short intervals.* Illinois J. Math. **39** (1995), 158–180. MR1299655 (95i:11115)

[11] A. Granville and G. Martin, *Prime number races.* Amer. Math. Monthly **113** (2006). 1–33. MR2202918 (2006h:11112)

[12] G. H. Hardy and J. E. Littlewood, *Some problems of 'partitio numerorum'. III: On the expression of a number as a sum of primes.* Acta Math. **44** (1923), 1–70. MR1555183

[13] A. E. Ingham, *The distribution of prime numbers.* Reprint of the 1932 original, with a foreword by R. C. Vaughan. Cambridge Univ. Press, 1990. MR1074573 (91f:11064)

[14] A. Ivić, *The Riemann zeta-function.* Wiley, New York, 1985. Reprinted by Dover Publ., Mineola, N.Y., 2003. MR792089 (87d:11062)

[15] J. Korevaar, *Distributional Wiener–Ikehara theorem and twin primes.* Indag. Math. (N.S.) **16** (2005), 37–49. MR2138049 (2006d:11105)

[16] J. Korevaar, *Prime pairs and the zeta function.* J. Approx. Theory **158** (2009), 69–96.

[17] T. Kotnik, *The prime-counting function and its analytic approximations.* Adv. Comput. Math. **29** (2008), 55–70. MR2420864 (2009c:11209)

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

[19] H. L. Montgomery, *Topics in multiplicative number theory.* Lecture Notes in Math., vol. 227, Springer-Verlag, Berlin-New York, 1971. MR0337847 (49:2616)

[20] H. L. Montgomery, *The pair correlation of zeros of the zeta function*, in: Analytic Number Theory (Proc. Symp. Pure Math., vol. 24), pp 181–193, Amer. Math. Soc., Providence, RI, 1973. MR0337821 (49:2590)

[21] T. R. Nicely, *Enumeration of the twin-prime pairs to* $10^{16}$. See the internet, http://www.trnicely.net, September 2008.

[22] H. J. J. te Riele, *On the sign of the difference* $\pi(x) - \mathrm{li}(x)$. Math. Comp. **48** (1987), 323–328. MR866118 (88a:11135)

[23] E.C. Titchmarsh, *The theory of the Riemann zeta-function.* First edition 1951, second edition edited by D.R. Heath-Brown. Clarendon Press, Oxford, 1986. MR882550 (88c:11049)

[24] E.T. Whittaker and G.N. Watson, *A Course of Modern Analysis.* Cambridge Univ. Press, 1927; reprinted 1996. MR1424469 (97k:01072)

[25] Marek Wolf, *An analog of the Skewes number for twin primes.* Undated attachment to private communication, July 2008.

KdV Institute of Mathematics, University of Amsterdam, Science Park 904, P.O. Box 94248, 1090 GE Amsterdam, The Netherlands

*E-mail address*: `J.Korevaar@uva.nl`

CWI: Centrum Wiskunde en Informatica, Science Park 123, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

*E-mail address*: `Herman.te.Riele@cwi.nl`