ON THE FIRST SIGN CHANGE OF $\theta(x) - x$

D. J. PLATT AND T. S. TRUDGIAN

Abstract. Let $\theta(x) = \sum_{p \leq x} \log p$. We show that $\theta(x) < x$ for $2 < x < 1.39 \cdot 10^{17}$. We also show that there is an $x < \exp(727.951332668)$ for which $\theta(x) > x$.

1. Introduction

Let $\pi(x)$ denote the number of primes not exceeding $x$. The prime number theorem is the statement that

$$\pi(x) \sim \text{li}(x) = \int_2^x \frac{dt}{\log t}.$$  \hfill (1)

One often deals not with $\pi(x)$ but with the less obstinate Chebyshev functions $\theta(x) = \sum_{p \leq x} \log p$ and $\psi(x) = \sum_{p^m \leq x} \log p$. The relation (1) is equivalent to $\psi(x) \sim x$ and $\theta(x) \sim x$.

Littlewood [10], showed that $\pi(x) - \text{li}(x)$ and $\psi(x) - x$ change sign infinitely often. Indeed, (see, e.g., [7, Thms. 34 and 35]) he showed more than this, namely that

$$\pi(x) - \text{li}(x) = \Omega_\pm \left( \frac{x^{1/2}}{\log x} \log \log \log x \right),$$  \hfill (2)

$$\psi(x) - x = \Omega_\pm \left( x^{1/2} \log \log \log x \right).$$

By [16] (3.36) we have

$$\psi(x) - \theta(x) \leq 1.427 \sqrt{x} \quad (x > 1),$$  \hfill (3)

which, together with the second relation in (2), shows that $\theta(x) - x$ changes sign infinitely often.

Littlewood’s proof that $\pi(x) - \text{li}(x)$ changes sign infinitely often was ineffective: the proof did not furnish a number $x_0$ such that one could guarantee that $\pi(x) - \text{li}(x)$ changes sign for some $x \leq x_0$. Skewes [19] made Littlewood’s theorem effective; the best known result is that there must be a sign change less that $1.3972 \cdot 10^{316}$ [17]. On the other hand, Kotnik [8] showed that $\pi(x) < \text{li}(x)$ for all $2 < x \leq 10^{14}$.

We turn now to the question of sign changes in $\psi(x) - x$ and $\theta(x) - x$. There is nothing of much interest to be said about the first change of sign of $\psi(x) - x$: for $x \in [0, 100]$ there are 24 sign changes. The problem of determining values of $C$ such that $\psi(x) - x$ changes sign in every interval $[x, Cx]$, for all sufficiently large $x$,
is much more interesting (as examined in [11]) but it is not something we consider here. As for the first change of sign in \(\theta(x) - x\), Schoenfeld [18, p. 360] showed that \(\theta(x) < x\) for all \(0 < x \leq 10^{11}\). This range appears to have been improved by Dusart, [2] p. 4] to \(0 < x \leq 8 \cdot 10^{11}\). We increase this in

**Theorem 1.** For \(0 < x \leq 1.39 \cdot 10^{17}\), \(\theta(x) < x\).

A result of Rosser [15, Lemma 4] is

**Lemma 1 (Rosser).** If \(\theta(x) < x\) for \(e^{2.4} \leq x \leq K\) for some \(K\), then \(\pi(x) < \text{li}(x)\) for \(e^{2.4} \leq x \leq K\).

This enables us to extend Kotnik’s result by proving

**Corollary 1.** \(\pi(x) < \text{li}(x)\) for all \(2 < x \leq 1.39 \cdot 10^{17}\).

Rosser and Schoenfeld [16, (3.38)], proved

\[
\psi(x) - \theta(x) - \theta(x^{\frac{3}{5}}) < 3x^{\frac{3}{5}}, \quad (x > 0).
\]

Table 3 in [6] gives us the bound \(|\psi(x) - x| \leq 7.5 \cdot 10^{-7}x\), which is valid for all \(x \geq e^{35} > 1.5 \cdot 10^{15}\). This, together with [11] and Theorem [11] enables us to make the following improvement to two results of Schoenfeld [18, (5.1*) and (5.3*)].

**Corollary 2.** For \(x > 0\),

\[
\theta(x) < (1 + 7.5 \cdot 10^{-7})x, \quad \psi(x) - \theta(x) < (1 + 7.5 \cdot 10^{-7})\sqrt{x} + 3x^{\frac{3}{5}}.
\]

We now turn to the question of sign changes in \(\theta(x) - x\). In [3,1] we prove

**Theorem 2.** There is some \(x \in [\exp(727.951332642), \exp(727.951332668)]\) for which \(\theta(x) > x\).

Throughout this article we make use of the following notation. For functions \(f(x)\) and \(g(x)\) we say that \(f(x) = O^*(g(x))\) if \(|f(x)| \leq g(x)\) for the range of \(x\) under consideration.

**2. OUTLINE OF ARGUMENT**

The explicit formula for \(\psi(x)\) is [7, Thm. 29]

\[
\psi_0(x) = \frac{\psi(x + 0) + \psi(x - 0)}{2} = x - \sum_{\rho} \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2} \log \left(1 - \frac{1}{x^2}\right).
\]

Since

\[
\psi(x) = \theta(x) + \theta(x^{\frac{1}{2}}) + \theta(x^{\frac{1}{3}}) + \ldots,
\]

we can manufacture an explicit formula for \(\theta(x)\). Using (4) and (5) we find that

\[
\theta(x) - x > -\theta(x^{\frac{3}{5}}) - \sum_{\rho} \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - 3x^{\frac{3}{5}}.
\]

One can see why \(\theta(x) < x\) ‘should’ happen often. On the Riemann hypothesis, \(\rho = \frac{1}{2} + i\gamma\); since \(\gamma \geq 14\), one expects the dominant term on the right side of (6) to be \(-\theta(x^{\frac{3}{5}})\).

We proceed in a manner similar to that in Lehman [9]. Let \(\alpha\) be a positive number. We shall make frequent use of the Gaussian kernel \(K(y) = \sqrt{\frac{2\pi}{\alpha}} \exp(-\frac{1}{2}\alpha y^2)\), which has the property that \(\int_{-\infty}^{\infty} K(y) dy = 1\).
Divide both sides of (6) by $x^{\frac{1}{2}}$, make the substitution $x \mapsto e^u$ and integrate against $K(u - \omega)$. This gives
\[
\int_{\omega - \eta}^{\omega + \eta} K(u - \omega) e^{\frac{u}{2}} \{ \theta(e^u) - e^u \} \, du > - \int_{\omega - \eta}^{\omega + \eta} K(u - \omega) \theta \left( e^{\frac{u}{2}} \right) e^{-\frac{u}{2}} \, du
\]
(7)
\[- \frac{1}{\rho} \int_{\omega - \eta}^{\omega + \eta} K(u - \omega) e^{u(\rho - \frac{1}{2})} \, du - \frac{\zeta'(0)}{\zeta(0)} \int_{\omega - \eta}^{\omega + \eta} K(u - \omega) e^{-\frac{u}{2}} \, du
\]
\[- 3 \int_{\omega - \eta}^{\omega + \eta} K(u - \omega) e^{-\frac{u}{2}} \, du = -I_1 - I_2 - I_3 - I_4,
\]
say. The interchange of summation and integration may be justified by noting that the sum over the zeroes of $\zeta(s)$ in (6) converges boundedly in $u \in [\omega - \eta, \omega + \eta]$. Noting that $\zeta'(0)/\zeta(0) = \log 2\pi$, we proceed to estimate $I_3$ and $I_4$ trivially to show that
\[0 < I_3 < e^{-\frac{\omega}{2}} \log 2\pi, \quad 0 < I_4 < 3e^{-\frac{\omega}{2}}.
\]
It will be shown in §3 that the contributions of $I_3$ and $I_4$ to (7) are negligible — this justifies our cavalier approach to their approximation.

We now turn to $I_2$. Let $A$ be the height to which the Riemann hypothesis has been verified, and let $T \leq A$ be the height to which we can reasonably compute zeroes to a high degree of accuracy — we make this notion precise in §3. Write $I_2 = S_1 + S_2$, where
\[S_1 = \sum_{|\gamma| \leq A} \frac{1}{\rho} \int_{\omega - \eta}^{\omega + \eta} K(u - \omega) e^{i\gamma u} \, du,
\]
\[S_2 = \sum_{|\gamma| > A} \frac{1}{\rho} \int_{\omega - \eta}^{\omega + \eta} K(u - \omega) e^{(\rho - \frac{1}{2})u} \, du.
\]
Our $S_1$ is the same as that used by Lehman in [9, pp. 402-403]. Using (4.8) and (4.9) of [9] shows that
\[S_1 = \sum_{|\gamma| \leq T} \frac{e^{i\gamma \omega}}{\rho} e^{-\gamma^2/2\alpha} + E_1,
\]
where
\[|E_1| < 0.08\sqrt{\alpha} e^{-\alpha \gamma^2/2} + e^{-T^2/2\alpha} \left\{ \frac{\alpha}{\pi T^2} \log \frac{T}{2\pi} + 8 \log \frac{T}{\alpha} + 4/3 \right\}.
\]
Lehman considers
\[f_\rho(s) = \rho s e^{-\rho s} \mathrm{li}(e^{\rho s}) e^{-\alpha (s - w)^2/2},
\]
whence he writes his analogous version of $S_2$ as a function of $f_\rho(s)$ and then estimates this using integration by parts, Cauchy’s theorem, and the bound
\[|f_\rho(s)| \leq 2 \exp(-\frac{1}{2} \alpha (s - w)^2).
\]
We consider the simpler function $f_\rho(s) = \exp(-\frac{1}{2} \alpha (s - w)^2)$, which clearly satisfies §3. We may proceed as in §5 of [9] to deduce that
\[|S_2| \leq A \log A e^{-A^2/(2\alpha) + (w + \eta)/2} \left\{ 4\alpha^{-\frac{1}{2}} + 15\eta \right\},
\]
provided that
\[4A/w \leq \alpha \leq A^2, \quad 2A/\alpha \leq \eta < w/2.
\]
All that remains is for us to estimate
\[ I_1 = \int_{\omega - \eta}^{\omega + \eta} \theta \left( e^{\frac{u}{2}} \right) e^{-\frac{u}{2}} K(u - \omega) \, du. \]

Table 3 in \cite{6} and \cite{3} give us
\[ |\theta(x) - x| \leq 1.5423 \cdot 10^{-9} x, \quad x \geq e^{200}, \]
which gives
\[ I_1 < 1 + 1.5423 \cdot 10^{-9}, \quad (\omega - \eta) \geq 400. \]

Thus, we have

**Theorem 3.** Let \( A \) be the height to which the Riemann hypothesis has been verified, and let \( T \) satisfy \( 0 < T \leq A \). Let \( \alpha, \eta \) and \( \omega \) be positive numbers for which \( \omega - \eta \geq 400 \) and for which
\[ 4A/\omega \leq \alpha \leq A^2, \quad 2A/\alpha \leq \eta \leq \omega/2. \]
Define \( K(y) = \sqrt{\alpha/(2\pi)} \exp(-\frac{1}{2} \alpha y^2) \) and
\[ I(\omega, \eta) = \int_{\omega - \eta}^{\omega + \eta} K(u - \omega) e^{-u/2} \{ \theta(u) - e^u \} \, du. \]
Then
\[ I(\omega, \eta) \geq -1 - \sum_{|\gamma| \leq T} \frac{e^{i \gamma \omega}}{\rho} e^{-\gamma^2/(2\alpha)} R_1 - R_2 - R_3 - R_4, \]
where
\[ R_1 = 1.5423 \cdot 10^{-9}, \]
\[ R_2 = 0.08 \sqrt{\alpha} e^{-\alpha \gamma^2/2} + e^{-T^2/2\alpha} \left\{ \frac{\alpha}{\pi T^2} \log \frac{T}{2\pi} + 8 \log \frac{T}{T} + \frac{4\alpha}{T^3} \right\}, \]
\[ R_3 = e^{-(\omega - \eta)/2} \log 2\pi + 3e^{-(\omega - \eta)/6}, \]
\[ R_4 = A(\log A) e^{-A^2/(2\alpha) + (\omega + \eta)/2} \left\{ 4\alpha^{-\frac{1}{2}} + 15\eta \right\}. \]

If one were to assume the Riemann hypothesis one could reduce the term \( R_4 \). This would give greater freedom in the choice of \( \alpha \); see \cite{3.1.3}

Approximations different from \cite{9} are available. For example, one could use Lemma 1 in \cite{20} to obtain \( |\theta(x) - x| \leq 0.0045x/(\log x)^2 \). One could also restrict the conditions in Theorem 3 to \( \omega - \eta \geq 600 \) using the slightly improved results from \cite{6} that are applicable thereto. Neither of these improves significantly the bounds in Theorem 2.

We now need to search for values of \( \omega, \eta, A, T \) and \( \alpha \) for which the right side of (11) is positive.

3. Computations

3.1. **Locating a crossover.** Consider the sum \( \Sigma_1 = \sum_{|\gamma| \leq T} \frac{e^{i \gamma \omega}}{\rho} \). We wish to find values of \( T \) and \( \omega \) for which this sum is small, that is, close to \(-1\); for such values the sum that appears in (11) should also be small. Bays and Hudson \cite{2}, when considering the problem of the first sign change of \( \pi(x) - \text{li}(x) \), identified some values of \( \omega \) for which \( \Sigma_1 \) is small. We investigated their values: \( \omega = 405, 412, 437, 599, 686 \) and 728.
For $\omega$ in this range, we have $R_1 = 1.5423 \cdot 10^{-9}$ so we endeavour to choose the parameters $A, T, \alpha$ and $\eta$ to make the other error terms comparable.

3.1.1. Choosing $A$. We relied on the rigorous verification of the Riemann hypothesis for $A = 3.0610046 \cdot 10^{10}$ by the second author [13]. This computation also produced a database of the zeroes below this height computed to an absolute accuracy of $\pm 2 \cdot 10^{-12}$ [3].

3.1.2. Choosing $T$. As already observed, we have sufficient zeroes to set $T = A \approx 3 \cdot 10^{10}$ but, since summing over roughly the $10^{11}$ zeroes below this height is too computationally expensive, we settled for $T = 6,970,346,000$ (about $2 \cdot 10^{10}$ zeroes). Even then, computing the sum using multiple precision interval arithmetic (see §3.1.4) takes about 40 hours on an 8 core platform.

3.1.3. Choosing the other parameters. To get the finest granularity on our search (i.e. to be able to detect narrow regions where $\theta(x) > x$) we aim at setting $\eta$ as small as possible. This in turn means setting $\alpha$ (which controls the width of the Gaussian) as large as possible. However, to ensure that $R_4$ is manageable, we need $A^2/(2\alpha) > \omega/2$ or $\alpha < A^2/\omega$. A little experimentation led us to

$$\alpha = 1, 153, 308, 722, 614, 227, 968, \quad \eta = \frac{933831}{2^{44}},$$

both of which are exactly representable in IEEE double precision.

3.1.4. Summing over the zeroes. Since

$$\exp(i\gamma \omega) + \exp(-i\gamma \omega) = \frac{\cos(\gamma \omega) + 2\gamma \sin(\gamma \omega)}{1 + \gamma^2},$$

the dominant term in $\Sigma_4$ is roughly $2\sin(\gamma \omega)/\gamma$. Though one might expect a relative accuracy of $2^{-53}$ when computing this in double precision, the effect of reducing $\gamma \omega \mod 2\pi$ degrades this to something like $2^{-17}$ when $\gamma = 10^9$ and $\omega = 400$. We are therefore forced into using multiple precision, even though that entails a performance penalty perhaps as high as a factor of 100. To avoid the need to consider rounding and truncation errors at all, we use the MPFI [14] multiple precision interval arithmetic package for all floating point computations. Making the change from scalar to interval arithmetic probably costs us another factor of 4 in terms of performance.

3.1.5. Results. We initially searched the regions around $\omega = 405, 412, 437, 599, 686$ and 728 using only those zeroes $\frac{1}{2} + i\gamma$ with $0 < \gamma < T = 5,000$. Although these results were not rigorous, it was hoped that a sum approaching $-1$ would indicate a potential crossover worth investigating with full rigour. As an example, Figure 1 shows the results for a region near $\omega = 437.7825$. This is some way from dipping below the $-1$ level and indeed a rigorous computation using the full set of zeroes and with $\omega = 437.78249$ fails to get over the line. The same pattern repeats for $\omega$ near 405, 412, 599 and 686.

In contrast, we expected the region near 728 to yield a point where $\theta(x) > x$. The lowest published interval containing an $x$ such that $\pi(x) > \text{li}(x)$ is

$$x \in [\exp(727.951335231), \exp(727.951335621)]$$
in [17]. Since the error terms for $\theta(x) - x$ are tighter than those for $\pi(x) - \text{li}(x)$ this necessarily means that the same $x$ will satisfy $\theta(x) > x$. In fact, we can do better.

Using $\omega = 727.951332655$ we get

$$\sum_{|\gamma| \leq 5000} \frac{e^{i\omega\gamma}}{\rho} \exp \left( -\frac{\gamma^2}{2\alpha} \right) \in [-1.0013360278, -1.0013360277].$$

We also have $R_1 + R_2 + R_3 + R_4 < 1.7 \cdot 10^{-9}$, so that

$$\int_{\omega - \eta}^{\omega + \eta} K(u - \omega)e^{u/2} \left\{ \theta(e^u) - e^u \right\} du > 0.0013360261.$$ 

3.1.6. Sharpening the region. Using the same argument as [17, §9], we can analyse the tails of the integral (10) and sharpen the region considerably. Consider, for $\eta_0 \in (0, \eta]$,

$$T_1 = \int_{\omega + \eta_0}^{\omega + \eta} K(u - \omega)e^{u/2} \left\{ \theta(e^u) - e^u \right\} du$$

and

$$T_2 = \int_{\omega - \eta_0}^{\omega - \eta} K(u - \omega)e^{u/2} \left\{ \theta(e^u) - e^u \right\} du.$$ 

Another appeal to Table 3 in [6], and (3), gives us

$$|\theta(x) - x| \leq 1.3082 \cdot 10^{-9} x, \quad x \geq e^{700}.$$ 

Thus for $\omega - \eta > 700$ we have

$$|T_1| + |T_2| \leq 1.3082 \cdot 10^{-9} (\eta - \eta_0)K(\eta_0) \left[ e^{\frac{\omega + \eta}{2}} + e^{\frac{\omega - \eta_0}{2}} \right].$$
Applying (13) to (12), we find we can take \( \eta_0 = \eta / 4.2867 \) so that
\[
\int_{\omega - \eta_0}^{\omega + \eta_0} K(u - \omega)e^{-u/2} \{ \theta(e^u) - e^u \} \, du > 2.75 \cdot 10^{-6},
\]
which proves Theorem 2. Therefore, there is at least one \( u \in (\omega - \eta_0, \omega + \eta_0) \) with \( \theta(e^u) - e^u > 0 \). Owing to the positivity of the kernel \( K(u - \omega) \) we deduce that there is at least one such \( u \) with
\[
\theta(e^u) - e^u > 2.75 \cdot 10^{-6} e^{u/2} > 10^{152}.
\]
Since \( \theta(x) \) is nondecreasing this proves

**Corollary 3.** There are more than \( 10^{152} \) successive integers \( x \) satisfying
\[
x \in [\exp(727.951332642), \exp(727.951332668)],
\]
for which \( \theta(x) > x \).

3.2. A lower bound. Having established an upper bound for the first \( x \) for which \( \theta(x) \) exceeds \( x \), we now turn to a lower bound. A simple method would be to sieve all the primes \( p \) less than some bound \( B \), sum \( \log p \) starting at \( p = 2 \), and compare the running total each time to \( p \). We set \( B = 1.39 \cdot 10^{17} \) since this was required by the second author for another result in [1]. By the prime number theorem we would expect to find about \( 3.5 \cdot 10^{15} \) primes below this bound. Since this is far too many for a single thread computation we must look for some way of computing in parallel.

3.2.1. A parallel algorithm. We divide the range \([0,B]\) into contiguous segments. For each segment \( S_j = [x_j, y_j] \) we set \( T = \Delta = \Delta_{\min} = 0 \). We look at each prime \( p_i \) in this segment, compute \( i = \log p_i \), and add it to \( T \). We set \( \Delta = \Delta + i - p_i + p_{i-1} \) and \( \Delta_{\min} = \min(\Delta_{\min}, \Delta) \). Thus, at any \( p_i \), \( \Delta_{\min} \) is the maximum amount by which \( \theta(p) \) has caught up with or gone further ahead of \( p \) within this segment. After processing all the primes within a segment, we output \( T \) and \( \Delta_{\min} \).

Now, for each segment \( S_j = [x, y] \) the value of \( \theta(x) \) is simply the sum of \( T_k \) with \( k < j \) and \( \theta(y) = \theta(x) + T_j \). Furthermore, if \( \theta(x) < x \) and \( \theta(x) + \Delta_{\min} > 0 \), then \( \theta(w) < w \) for all \( w \in [x, y] \).

3.2.2. Results. We implemented this algorithm in C++ using Kim Walisch’s “primesieve” [21] to enumerate the primes efficiently, and the second author’s double precision interval arithmetic package to manage rounding errors.

We split \( B \) into 10,000 segments of width \( 10^{13} \) followed by 390 segments of width \( 10^{14} \). This pattern was chosen so that we could use Oliveira e Silva’s tables of \( \pi(x) \) [12] as an independent check of the sieving process.

We used the 16 core nodes of the University of Bristol Bluecrystal Phase III cluster [1] and we were able to utilise each core fully. In total we used about 78,000 node hours. This established Theorem 1.

We plot \((x - \theta(x))/\sqrt{x}\) measured at the end of each segment in Figure 2. As one would expect, this appears to be a random walk around the line 1.
Figure 2. Plot of $\frac{x-\Theta(x)}{\sqrt{x}}$.

References

[9] R. S. Lehman, On the difference $\pi(x) - \text{li}(x)$, Acta Arith. 11 (1966), 397–410. MR0202686 (34 #2546)
ON THE FIRST SIGN CHANGE OF $\theta(x) - x$


[18] L. Schoenfeld, Sharper bounds for the Chebyshev functions $\theta(x)$ and $\psi(x)$. II, Math. Comp. 30 (1976), no. 134, 337–360. MR0457374 (56 #15581b)


Heilbronn Institute for Mathematical Research University of Bristol, Bristol, United Kingdom
E-mail address: dave.platt@bris.ac.uk

Mathematical Sciences Institute, The Australian National University, ACT 0200, Australia
E-mail address: timothy.trudgian@anu.edu.au