An asymptotic formula for the Chebyshev theta function

Aditya Ghosh

Indian Statistical Institute, Kolkata, India
e-mail: ghoshadi26@gmail.com

Received: 11 March 2019 Revised: 21 October 2019 Accepted: 23 October 2019

Abstract: Let \( \{p_n\}_{n \geq 1} \) be the sequence of primes and \( \vartheta(x) = \sum_{p \leq x} \log p \), where \( p \) runs over the primes not exceeding \( x \), be the Chebyshev \( \vartheta \)-function. In this note, we derive lower and upper bounds for \( \vartheta(p_n)/n \), by comparing it with \( \log p_{n+1} \) and deduce the asymptotic formula
\[
\vartheta(p_n)/n = \log p_{n+1} \left( 1 - \frac{1}{\log n} + \frac{\log \log n}{\log^2 n} \left( 1 + o(1) \right) \right).
\]

Keywords: Chebyshev theta function, Geometric mean of first \( n \) primes, Prime numbers.

2010 Mathematics Subject Classification: 11A41, 11A25.

1 Introduction

Let \( \{p_n\}_{n \geq 1} \) be the sequence of the prime numbers and \( \vartheta(x) = \sum_{p \leq x} \log p \), where \( p \) runs over the primes not exceeding \( x \), be the Chebyshev \( \vartheta \)-function. The type of bounds that we shall discuss here was introduced by Bonse [2], who showed that \( \vartheta(p_n) > 2 \log p_{n+1} \) holds for every \( n \geq 4 \) and \( \vartheta(p_n) > 3 \log p_{n+1} \) holds for every \( n \geq 5 \). Thereafter, Pósa [8] showed that, given any \( k > 1 \), there exists \( n_k \) such that \( \vartheta(p_n) > k \log p_{n+1} \) holds for all \( n \geq n_k \). Panaitopol [7] showed that in Pósa’s result we can have \( n_k = 2k \) and also gave the bound
\[
\frac{\vartheta(p_n)}{\log p_{n+1}} > n - \pi(n) \quad (n \geq 2),
\]
where \( \pi(n) \) is equal to the number of primes less or equal to \( n \). Hassani [5] refined Panaitopol’s inequality to the following
\[
\frac{\vartheta(p_n)}{\log p_{n+1}} > n - \pi(n) \left( 1 - \frac{1}{\log n} \right) \quad (n \geq 101).
\] (1)
Recently, Axler [1, Propositions 4.1 and 4.5] showed that
\[ 1 + \frac{1}{\log p_n} + \frac{2.7}{\log^2 p_n} < \log p_n - \frac{\vartheta(p_n)}{n} < 1 + \frac{1}{\log p_n} + \frac{3.84}{\log^2 p_n}, \]
where the left-hand side inequality is valid for every integer \( n \geq 218 \) and the right-hand side inequality holds for every \( n \geq 74004585 \). This provides the following asymptotic formula
\[ \frac{\vartheta(p_n)}{n} = \log p_n - 1 - \frac{1}{\log p_n} + \Theta \left( \frac{1}{\log^2 p_n} \right). \]

For further terms, see Axler [1, Proposition 2.1].

In the present note, we show the following result, which is a refinement of (1).

**Theorem 1.1.** For all \( n \geq 6 \), we have
\[ n \left( 1 - \frac{1}{\log n} + \frac{\log \log n}{4 \log^2 n} \right) \leq \frac{\vartheta(p_n)}{\log p_{n+1}} \leq n \left( 1 - \frac{1}{\log n} + \frac{\log \log n}{\log^2 n} \right) \]
(2)
The left-hand side inequality also holds for \( 2 \leq n \leq 6 \).

We can, in fact, generalise the left-hand side of (2) to have the following result.

**Theorem 1.2.** For every \( 0 < \varepsilon < 1 \), there exists \( n_\varepsilon \in \mathbb{N} \) such that for every \( n \geq n_\varepsilon \) it holds that
\[ n \left( 1 - \frac{1}{\log n} + (1 - \varepsilon) \frac{\log \log n}{\log^2 n} \right) \leq \frac{\vartheta(p_n)}{\log p_{n+1}} \leq n \left( 1 - \frac{1}{\log n} + \frac{\log \log n}{\log^2 n} \right). \]
(3)

**Corollary 1.2.1.** We have
\[ \frac{\vartheta(p_n)}{n} = \log p_{n+1} \left( 1 - \frac{1}{\log n} + \frac{\log \log n}{\log^2 n} (1 + o(1)) \right). \]

2 Preliminaries

Define \( G(n, a) = \log n + \log \log n - 1 + \frac{\log \log n - a}{\log n} \). We shall use the following bounds for \( \vartheta(p_n)/n \).

**Lemma 2.1.** For every \( n \geq 3 \), we have
\[ \frac{\vartheta(p_n)}{n} \geq G(n, 2.1454), \]
(4)
and for every \( n \geq 198 \), we have
\[ \frac{\vartheta(p_n)}{n} \leq G(n, 2). \]
(5)

**Proof.** The inequality (4) is due to Robin [9], and the inequality (5) was given by Massias and Robin [6].
Lemma 2.2. For every $n \geq 227$, we have

$$p_n \leq n(\log n + \log \log n - 0.8),$$

and for every $n \geq 2$,

$$p_n \geq n(\log n + \log \log n - 1).$$

Proof. For $n \geq 8602$, we have the following stronger bound

$$p_n \leq n(\log n + \log \log n - 0.9385)$$

given by Massias and Robin [6]. For $227 \leq n \leq 8601$ we verify the inequality (6) by direct computation. The inequality (7) is due to Dusart [4]. ∎

For the sake of brevity, we define $F(n, \lambda) = 1 - \frac{1}{\log n} + \frac{\lambda \log \log n}{\log n}$ and rewrite (2) as

$$F(n, 0.25) \log p_{n+1} \leq \vartheta(p_n)/n \leq F(n, 1) \log p_{n+1}$$

and rewrite (3) as

$$F(n, 1 - \varepsilon) \log p_{n+1} \leq \vartheta(p_n)/n \leq F(n, 1) \log p_{n+1}.$$  

3 Proof of Theorem 1.1

The proof of Theorem 1.1 is split into two lemmas. In the first lemma, we give lower and upper bounds for $\log p_{n+1}$.

Lemma 3.1. For every $n \geq 140$, we have

$$\log p_{n+1} < \log n + \log \log n + \frac{\log \log n - 0.8 + 0.018}{\log n} = U(n),$$

and for every $n \geq 2$, we have

$$\log p_{n+1} > \log n + \log \log n + \frac{\log \log n - 1}{\log n + 0.5(\log \log n - 1)} = V(n).$$

Proof. First, we show that for every $x \geq 1$

$$\frac{1}{x + 0.4} > \log \left(1 + \frac{1}{x}\right) > \frac{1}{x + 0.5}.$$  

In order to prove this, we set $f_a(x) = \log(1 + x) - \frac{x}{1 + ax}$. Note that, $f'_a(x) = \frac{x(a^2x + 2a - 1)}{(1 + x)(1 + ax)^2}$. Hence, $f'_a(x) < 0$ for every $x \in (0, 1.25)$ which yields $f_a(1/x) < f_a(0) = 0$ for every $x \geq 1$.  

3
On the other hand, $f'_{0.5}(x) > 0$ for all positive $x$, which gives $f_{0.5}(1/x) > f_{0.5}(0) = 0$ for every $x \geq 1$.

Next, we give a proof of (11). By (6), we have for $n \geq 227$,
\[
\log p_{n+1} \leq \log(n + 1) + \log(n + 1) + \log \log(n + 1) - 0.8. \tag{14}
\]
The left-hand side inequality of (13) gives $\log(n + 1) < \log n + \frac{1}{n + 0.4}$, which in turn implies that
\[
\log \log(n + 1) < \log \log n \log \left(1 + \frac{1}{(n + 0.4) \log n}\right) < \log \log n + \frac{1}{(n + 0.4) \log n}.
\]
Applying this to (14), we obtain for $n \geq 227$,
\[
\log p_{n+1} \leq \log n + \frac{1}{n + 0.4} + \log \left(\log n + \log \log n - 0.8 + \frac{1 + 1/ \log n}{n + 0.4}\right) < \log n + \log \log n + \frac{\log \log n - 0.8}{\log n} + \frac{1}{\log n} \cdot \frac{\log n + 1 + 1/ \log n}{n + 0.4}. \tag{15}
\]
Now, $g(x) = \frac{\log x + 1 + 1/ \log x}{x + 0.4}$ is a decreasing function for $x \geq 2$ with $g(e^{5.99}) \leq 0.018$. Hence $g(x) \leq 0.018$ holds for every $x \geq 400 > e^{5.99}$. Combined with (15), it yields that $\log p_{n+1} < U(n)$ for every $n \geq 400$. For every $140 \leq n \leq 399$ we check the inequality (11) with a computer. This completes the proof of (11).

To prove the inequality (12), first note that (7) gives for every $n \geq 1$,
\[
\log p_{n+1} \geq \log(n + 1) + \log(n + 1) + \log \log(n + 1) - 1. \tag{16}
\]
The right-side inequality of (13) gives $\log(n + 1) > \log n + \frac{1}{n + 0.5}$. Using (13) once again, we get, for $n \geq 2$,
\[
\log \log(n + 1) - \log \log n > \log \left(1 + \frac{1}{(n + 0.5) \log n}\right) > \frac{1}{(n + 0.5) \log n + 0.5}.
\]
Applying this to (16), we arrive at
\[
\log p_{n+1} > \log n + \log \left(\log n + \frac{1}{n + 0.5} + \log \log n + \frac{1}{(n + 0.5) \log n + 0.5} - 1\right) > \log n + \log \log n + \log \left(1 + \frac{\log \log n - 1}{\log n}\right).
\]
Invoking (13) one more time, we get $\log p_{n+1} < V(n)$ for every $n \geq 2$. \hfill \square

Lemma 3.2. For every $n \geq 396$, we have
\[
G(n, 2.1454) \geq F(n, 0.25) \cdot U(n), \tag{17}
\]
and for every $n \geq 2$, we have
\[
G(n, 2) \leq F(n, 1) \cdot V(n). \tag{18}
\]
Here $U(n)$ and $V(n)$ are defined as in Lemma 3.1.
Proof. We start with the proof of (17). Setting \( x = \log n \), the inequality (17) can be rewritten as
\[
x + \log x - 1 + \frac{\log x - 2.1454}{x} \geq \left(1 - \frac{1}{x} + \frac{\log x}{4x^2}\right) \left(x + \log x + \frac{\log x - 0.8 + 0.018}{x}\right),
\]
which is equivalent to
\[
\left(\frac{3}{4} \log x + \frac{\log x}{x}\right) + \left(-2.1454 - \frac{\log^2 x}{2x} - \frac{\log^2 x}{4x^2}\right) + (0.8 - 0.018) \left(1 - \frac{1}{x} + \frac{\log x}{4x^2}\right) \geq 0.
\]
The left-hand side is a sum of three increasing functions on the interval \([5.7, \infty)\) and at \( x = 5.99 \) the left-hand side is positive. So the last inequality holds for every \( x \geq 5.99 \); i.e., for every \( n \geq 400 \). A direct computation shows that the inequality (17) also holds for every \( n \) satisfying \( 396 \leq n \leq 399 \).

Next, we give a proof of (18). It is easy to see that
\[
x^2 + \log x (\log x - 1) > \frac{x}{2} \log x (\log x - 1)
\]
for every \( x > 0 \). Now, for \( x \geq 1 \), the last inequality is seen to be equivalent to
\[
\left(1 - \frac{1}{x} + \frac{\log x}{x^2}\right) \frac{\log x - 1}{x + 0.5(\log x - 1)} \geq \frac{\log x - 2}{x}.
\]
Since \( \frac{\log^2 x}{x^2} \geq 0 \) for every \( x > 0 \), we get
\[
\frac{\log^2 x}{x^2} + \left(1 - \frac{1}{x} + \frac{\log x}{x^2}\right) \frac{\log x - 1}{x + 0.5(\log x - 1)} \geq \frac{\log x - 2}{x}
\]
(19) for every \( x \geq 1 \). Substituting \( x = \log n \) in (19), we obtain the inequality (18) for every integer \( n \geq 3 \). We can directly check that (18) holds for \( n = 2 \) as well.

Finally, we give a proof of Theorem 1.1.

Proof of Theorem 1.1. We use (4), (17) and (11) respectively to see that for every \( n \geq 396 \),
\[
\vartheta(p_n)/n \geq G(n, 2.1454) \geq F(n, 0.25) U(n) > F(n, 0.25) \log p_{n+1}.
\]
A direct computation shows that the left-hand side inequality of (9) also holds for every integer \( n \) with \( 2 \leq n \leq 395 \).

In order to prove the right-hand side inequality of (9), we combine (5), (18) and (12), respectively, to get
\[
\vartheta(p_n)/n \leq G(n, 2) \leq F(n, 1) V(n) \leq F(n, 1) \log p_{n+1}
\]
for every \( n \geq 198 \). For smaller values of \( n \), we use a computer.
4 Proof of Theorem 1.2

The right-hand side of (10) has been established already. To show the left-hand side, we start with the following lemma.

Lemma 4.1. For any \( 0 < \varepsilon < 1 \), there exists \( m_{\varepsilon} \in \mathbb{N} \) such that

\[
G(n, 2.1454) \geq \mathcal{F}(n, 1 - \varepsilon) \cdot U(n)
\]

holds for every \( n \geq m_{\varepsilon} \). Here \( U(n) \) is defined as in Lemma 3.1.

Proof. Fix any \( 0 < \varepsilon < 1 \). We denote \( a = 2.1454 \), \( b = 0.8 - 0.018 \) and set \( x = \log n \) to transform the inequality (20) into

\[
x + \log x - 1 + \frac{\log x - a}{x} \geq \left( 1 - \frac{1}{x} + (1 - \varepsilon) \frac{\log x}{4x^2} \right) \left( x + \log x + \frac{\log x - b}{x} \right).
\]

This is equivalent to

\[
\left( \varepsilon \log x + \frac{\log x}{x} \right) + \left( -a - (1 - \varepsilon) \left( \frac{\log^2 x}{x} + \frac{\log^2 x}{x^2} \right) \right) + b \left( 1 - \frac{1}{x} + (1 - \varepsilon) \frac{\log x}{x} \right) \geq 0.
\]

Now, the left-hand side is a sum of three functions, each of which is strictly increasing for all sufficiently large \( x \), and the limit of the left-hand side, as \( x \to \infty \), is \( +\infty \). Therefore we conclude that the last inequality holds for all sufficiently large \( x \).

\( \Box \)

Proof of Theorem 1.2. For any \( 0 < \varepsilon < 1 \), we have \( m_{\varepsilon} \in \mathbb{N} \) such that (20) holds for every \( n \geq m_{\varepsilon} \). We combine this with (4) and (11) to obtain that for every \( n \geq n_{\varepsilon} := \max\{m_{\varepsilon}, 140\} \)

\[
\vartheta(p_n)/n \geq G(n, 2.1454) \geq \mathcal{F}(n, 1 - \varepsilon) U(n) \geq \mathcal{F}(n, 1 - \varepsilon) \log p_{n+1}.
\]

This completes the proof.

\( \Box \)

5 Remarks

1. For every \( n \geq 599 \), we have

\[
\frac{\pi(n)}{n} \geq \frac{1}{\log n} + \frac{1}{\log^2 n},
\]

which was found by Dusart [3]. Using this and a computer, we get

\[
\frac{\pi(n)}{n} \geq \frac{1}{\log n - 1} \left( 1 - \frac{\log \log n}{4 \log n} \right)
\]

for every integer \( n \geq 83 \). Hence, (2) is an improvement of (1).

2. The bounds given in (2) are particularly useful for comparing \( \vartheta(p_n)/n \) with \( \log p_{n+1} \). To see a numerical example, we use a computer to find that for \( n \geq 23 \) the relative error in approximating \( \vartheta(p_n)/n \) with \( \mathcal{F}(n, 0.25) \) is less than 5% and for \( n \geq 114 \) it is less than 2%.

An important feature of (2) is that it holds even for very small values of \( n \).
Acknowledgements

I am thankful to Mridul Nandi (Indian Statistical Institute, Kolkata, India) and Mehdi Hassani (University of Zanjan, Iran) for their valuable suggestions.

References


