

IMPROVED ESTIMATE FOR THE PRIME COUNTING FUNCTION

THEOPHILUS AGAMA

ABSTRACT. Using some simple combinatorial arguments, we establish some new estimates for the prime counting function and its allied functions. In particular we show that

$$\pi(x) = \Theta(x) + O\left(\frac{1}{\log x}\right),$$

where

$$\Theta(x) = \frac{\theta(x)}{\log x} + \frac{x}{2 \log x} - \frac{1}{4} - \frac{\log 2}{\log x} \sum_{\substack{n \leq x \\ \Omega(n)=k \\ k \geq 2 \\ 2|n}} \frac{\log(\frac{x}{n})}{\log 2}.$$

This is an improvement to the estimate

$$\pi(x) = \frac{\theta(x)}{\log x} + O\left(\frac{x}{\log^2 x}\right)$$

found in the literature.

1. INTRODUCTION

Let us set $\pi(x) = \sum_{p \leq x} 1$, where $\pi(x)$ counts the number of primes no more than a fixed real number. The irregularity of the primes makes it difficult to obtain an exact practical formula for the prime counting function. It is one of the well studied functions in number theory and the whole of mathematics given its profound connection to the grand Riemann hypothesis, but not much is known concerning their distribution. In the following sequel, our mission is to improve a known estimate for the prime counting function involving the chebyshev theta function, given by

$$\pi(x) = \frac{\theta(x)}{\log x} + O\left(\frac{x}{\log^2 x}\right).$$

2. NOTATIONS

Through out this paper a prime number will either be denoted by p or the subscripts of p . Any other letter will be clarified. The function $\Omega(n) := \sum_{p|n} 1$ counts the number of prime factors of n with multiplicity. The inequality $|k(n)| \leq Mp(n)$ for sufficiently large values of n will be compactly written as $k(n) \ll p(n)$ or $k(n) = O(p(n))$.

Date: May 28, 2020.

2000 Mathematics Subject Classification. Primary 54C40, 14E20; Secondary 46E25, 20C20.

Key words and phrases. prime, theta.

3. PRELIMINARY RESULTS

Lemma 3.1. *There exist some constant $c > 0$ such that*

$$\theta(x) = x + O\left(\frac{x}{e^{c\sqrt{\log x}}}\right).$$

Proof. For a proof, see for instance [3]. □

Lemma 3.2. *For all $x \geq 2$*

$$\pi(x) = \frac{\theta(x)}{\log x} + \int_2^x \frac{\theta(t)}{t \log^2 t} dt.$$

Proof. For a proof see for instance [1]. □

Lemma 3.3. *For all $x \geq 2$*

$$\theta(x) = \pi(x) \log x - \int_2^x \frac{\pi(t)}{t} dt.$$

Proof. For a proof see for instance [1]. □

4. MAIN RESULT

Theorem 4.1. *For every positive integer x*

$$\sum_{\substack{n \leq x \\ (2,n)=1}} \left\lfloor \frac{\log(\frac{x}{n})}{\log 2} \right\rfloor = \frac{x-1}{2} + \left(1 + (-1)^x\right) \frac{1}{4}.$$

Proof. The plan of attack is to examine the distribution of odd natural numbers and even numbers. We break the proof of this result into two cases; The case x is odd and the case x is even. For the case x is odd, we argue as follows: We first observe that there are as many even numbers as odd numbers less than any given odd number x . That is, for $1 \leq m < x$, there are $(x-1)/2$ such possibilities. On the other hand, consider the sequence of even numbers less than x , given as $2, 2^2, \dots, 2^b$ such that $2^b < x$; clearly there are $\lfloor \frac{\log x}{\log 2} \rfloor$ such number of terms in the sequence. Again consider those of the form $3 \cdot 2, \dots, 3 \cdot 2^b$ such that $3 \cdot 2^b < x$; Clearly there are $\lfloor \frac{\log(x/3)}{\log 2} \rfloor$ such terms in this sequence. We terminate the process by considering those of the form $2 \cdot j, \dots, 2^b \cdot j$ such that $(2, j) = 1$ and $2^b \cdot j < x$; Clearly there are $\lfloor \frac{\log(x/j)}{\log 2} \rfloor$ such number of terms in this sequence. The upshot is that $(x-1)/2 = \sum_{\substack{j \leq x \\ (2,j)=1}} \left\lfloor \frac{\log(x/j)}{\log 2} \right\rfloor$. We now turn to the case x is even. For

the case x is even, we argue as follows: First we observe that there are $x/2$ even numbers less than or equal to x . On the other hand, there are $\sum_{\substack{j \leq x \\ (2,j)=1}} \left\lfloor \frac{\log(x/j)}{\log 2} \right\rfloor$

even numbers less than or equal to x . This culminates into the assertion that $(x-1)/2 + \frac{1}{2} = \sum_{\substack{j \leq x \\ (2,j)=1}} \left\lfloor \frac{\log(x/j)}{\log 2} \right\rfloor$. By combining both cases, the result follows immediately. □

Remark 4.2. Understanding the fractional parts of real numbers is an important problem in number theory and gains much paramountcy. The next result gives a little insight into the problem.

Corollary 4.1.

$$\sum_{n \leq x} \left\{ \frac{\log \left(\frac{x}{n} \right)}{\log 2} \right\} = \frac{x}{\log 2} - x - \frac{\log x}{\log 4} + O(1).$$

Proof. Stirling's formula [4] gives

$$(4.1) \quad \sum_{n \leq x} \log n = x \log x - x + \frac{1}{2} \log x + \log(\sqrt{2\pi}) + O\left(\frac{1}{x}\right).$$

Also from Theorem 4.1, we obtain

$$(4.2) \quad \sum_{n \leq x} \log n = x \log x - x \log 2 - \log 2 \sum_{n \leq x} \left\{ \frac{\log \left(\frac{x}{n} \right)}{\log 2} \right\} + O(1).$$

Comparing equation (4.1) and (4.2), we have

$$\sum_{n \leq x} \left\{ \frac{\log \left(\frac{x}{n} \right)}{\log 2} \right\} = \frac{x}{\log 2} - x - \frac{\log x}{\log 4} + O(1),$$

thereby establishing the estimate. \square

The prime counting function is one of the most important functions in number theory, given it's connection with the famous Riemann hypothesis. The prime counting function has been studied by many authors in the past decade. For instance in 2010, Dusart [2] showed that

$$\pi(x) \leq \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2.334x}{\log^3 x}$$

holds for every $x \geq 2953652287$ and that

$$\pi(x) \geq \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x}$$

for every $x \geq 88783$. We now obtain, as an immediate consequence of Theorem 4.1, an explicit formula for the prime counting function $\pi(x)$. It relates the prime counting function to the Chebychev theta function $\theta(x)$.

Theorem 4.3. *For all positive integers x*

$$\pi(x) = \frac{(x-1) \log(\sqrt{2}) + \theta(x) + (\log 2) \left(H(x) - G(x) + T(x) \right) + \left(1 + (-1)^x \right) \frac{\log 2}{4}}{\log x},$$

where

$$H(x) := \sum_{p \leq x} \left\{ \frac{\log(\frac{x}{p})}{\log 2} \right\}, \quad G(x) := \left\lfloor \frac{\log x}{\log 2} \right\rfloor + \sum_{\substack{n \leq x \\ \Omega(n)=k \\ k \geq 2 \\ 2 \nmid k}} \left\lfloor \frac{\log(\frac{x}{n})}{\log 2} \right\rfloor, \quad T(x) := \left\lfloor \frac{\log(\frac{x}{2})}{\log 2} \right\rfloor$$

$$\theta(x) := \sum_{p \leq x} \log p, \quad \text{and} \quad \{\cdot\}$$

denotes the fractional part of any real number.

Proof. We use Theorem 4.1 to establish an explicit formula for the prime counting function. Theorem 4.1 gives $(x-1)/2 = \sum_{\substack{n \leq x \\ (2,n)=1}} \left\lfloor \frac{\log(x/n)}{\log 2} \right\rfloor - (1 + (-1)^x)/4$,

which can then be recast as

$$(x-1)/2 = \sum_{p \leq x} \left\lfloor \frac{\log(x/p)}{\log 2} \right\rfloor - \left\lfloor \frac{\log(x/2)}{\log 2} \right\rfloor + \left\lfloor \frac{\log x}{\log 2} \right\rfloor + \sum_{\substack{n \leq x \\ \Omega(n)=k \\ k \geq 2 \\ 2 \nmid k}} \left\lfloor \frac{\log(x/n)}{\log 2} \right\rfloor - \left(1 + (-1)^x\right)/4,$$

where p runs over the primes. It follows by further simplification that

$$(x-1)/2 = \frac{1}{\log 2} \left(\log x \sum_{p \leq x} 1 - \sum_{p \leq x} \log p \right) - \sum_{p \leq x} \left\{ \frac{\log(x/p)}{\log 2} \right\} + \left\lfloor \frac{\log x}{\log 2} \right\rfloor - \left\lfloor \frac{\log(x/2)}{\log 2} \right\rfloor \\ + \sum_{\substack{n \leq x \\ \Omega(n)=k \\ k \geq 2 \\ 2 \nmid k}} \left\lfloor \frac{\log(x/n)}{\log 2} \right\rfloor - \left(1 + (-1)^x\right)/4.$$

It follows that

$$(x-1) \log \sqrt{2} = \pi(x) \log x - \theta(x) - (\log 2) \left(H(x) - G(x) + T(x) \right) - \left(1 + (-1)^x\right)/4,$$

where

$$H(x) := \sum_{p \leq x} \left\{ \frac{\log(\frac{x}{p})}{\log 2} \right\}, \quad G(x) := \left\lfloor \frac{\log x}{\log 2} \right\rfloor + \sum_{\substack{n \leq x \\ \Omega(n)=k \\ k \geq 2 \\ 2 \nmid k}} \left\lfloor \frac{\log(\frac{x}{n})}{\log 2} \right\rfloor, \quad T(x) := \left\lfloor \frac{\log(\frac{x}{2})}{\log 2} \right\rfloor,$$

$$\text{and} \quad \theta(x) := \sum_{p \leq x} \log p,$$

and the result follows immediately. \square

Remark 4.4. Using Theorem 4.3, we obtain an estimate for the prime counting and its allied functions, with very small error term in the following sequel.

Theorem 4.5. For all positive integers $x \geq 2$

$$\pi(x) = \Theta(x) + O\left(\frac{1}{\log x}\right),$$

where

$$\Theta(x) = \frac{\theta(x)}{\log x} + \frac{x}{2 \log x} - \frac{1}{4} - \frac{\log 2}{\log x} \sum_{\substack{n \leq x \\ \Omega(n)=k \\ k \geq 2 \\ 2 \nmid k}} \left\lfloor \frac{\log(\frac{x}{n})}{\log 2} \right\rfloor.$$

Proof. By Theorem 4.3, we can write

$$\pi(x) = \left(\frac{\log 2}{2}\right) \frac{x}{\log x} + \frac{\theta(x)}{\log x} + \frac{\log 2}{\log x} \left(H(x) - G(x) + T(x)\right) + O\left(\frac{1}{\log x}\right),$$

where

$$H(x) := \sum_{p \leq x} \left\lfloor \frac{\log(\frac{x}{p})}{\log 2} \right\rfloor, \quad G(x) := \left\lfloor \frac{\log x}{\log 2} \right\rfloor + \sum_{\substack{n \leq x \\ \Omega(n)=k \\ k \geq 2 \\ 2 \nmid k}} \left\lfloor \frac{\log(\frac{x}{n})}{\log 2} \right\rfloor, \quad T(x) := \left\lfloor \frac{\log(\frac{x}{2})}{\log 2} \right\rfloor.$$

Now, we estimate the term $H(x) - G(x) + T(x)$. Clearly we can write

$$-G(x) + T(x) = - \sum_{\substack{n \leq x \\ \Omega(n)=k \\ k \geq 2 \\ 2 \nmid k}} \left\lfloor \frac{\log(\frac{x}{n})}{\log 2} \right\rfloor + O(1)$$

It follows that

$$\begin{aligned} H(x) - G(x) + T(x) &= \sum_{\substack{n \leq x \\ (2,n)=1}} \left\lfloor \frac{\log\left(\frac{x}{n}\right)}{\log 2} \right\rfloor - \sum_{\substack{n \leq x \\ \Omega(n)=k \\ k \geq 2 \\ 2 \nmid k}} \left\lfloor \frac{\log(\frac{x}{n})}{\log 2} \right\rfloor + O(1) \\ &= \frac{x}{\log 4} - \frac{x}{2} - \frac{\log x}{\log 16} - \sum_{\substack{n \leq x \\ \Omega(n)=k \\ k \geq 2 \\ 2 \nmid k}} \left\lfloor \frac{\log(\frac{x}{n})}{\log 2} \right\rfloor + O(1), \end{aligned}$$

where we have used Corollary 4.1 and the proof is complete. \square

Corollary 4.2. There exist a constant $c > 0$ such that

$$\pi(x) = \nu(x) + O\left(\frac{x}{e^{c\sqrt{\log x}} \log x}\right),$$

where

$$\nu(x) := \frac{3x}{2 \log x} - \frac{1}{4} - \frac{\log 2}{\log x} \sum_{\substack{n \leq x \\ \Omega(n)=k \\ k \geq 2 \\ 2 \nmid k}} \frac{\log(\frac{x}{n})}{\log 2}.$$

Proof. The result follows by plugging the estimate in Lemma 3.1 into Theorem 4.5. \square

It is known that

$$\pi(x) = \frac{\theta(x)}{\log x} + O\left(\frac{x}{\log^2 x}\right)$$

[3], where $\theta(x) := \sum_{p \leq x} \log p$. The above estimate derived indicates that we can do better than this with

$$\pi(x) = \Theta(x) + O\left(\frac{1}{\log x}\right),$$

where $\Theta(x)$ is some explicit function.

Remark 4.6. It needs to be said that Theorem 4.5 relates the prime counting function to the Chebyshev theta function. Thus, very good and sharp estimates for $\theta(x)$ will eventually yield a sharp estimate for the prime counting function. For the purpose of applications, we stick to the current estimate involving the Chebyshev theta function.

Theorem 4.7. *For all positive integers $x \geq 2$, we have*

$$\int_2^x \frac{\theta(t)}{t \log^2 t} dt = R(x) + O\left(\frac{1}{\log x}\right),$$

where

$$R(x) = \frac{x}{2 \log x} - \frac{1}{4} - \frac{\log 2}{\log x} \sum_{\substack{n \leq x \\ \Omega(n)=k \\ k \geq 2 \\ 2|n}} \frac{\log(\frac{x}{n})}{\log 2}.$$

Proof. It is known that

$$\pi(x) = \frac{\theta(x)}{\log x} + \int_2^x \frac{\theta(t)}{t \log^2 t} dt$$

for all $x \geq 2$ [1]. Comparing this result with Theorem 4.5, the result follows immediately. \square

Theorem 4.8. *For all positive integers $x \geq 2$, we have*

$$\int_2^x \frac{\pi(t)}{t} dt = \eta(x) + O(1),$$

where

$$\eta(x) = \frac{x}{2} - \frac{\log x}{4} - \log 2 \sum_{\substack{n \leq x \\ \Omega(n)=k \\ k \geq 2 \\ 2|n}} \frac{\log(\frac{x}{n})}{\log 2}.$$

Proof. Recall the well-known identity [1]

$$\theta(x) = \pi(x) \log x - \int_2^x \frac{\pi(t)}{t} dt.$$

By comparing this identity with Theorem 4.5, the result follows immediately. \square

Remark 4.9. Theorem 4.7 measures the discrepancies of the prime counting function and $\theta(x)/\log x$.

5. FINAL REMARKS

In this paper we have obtained a new estimate for the prime counting function. We have shown that indeed the estimate

$$\pi(x) = \Theta(x) + O\left(\frac{1}{\log x}\right),$$

where

$$\Theta(x) = \frac{\theta(x)}{\log x} + \frac{x}{2 \log x} - \frac{1}{4} - \frac{\log 2}{\log x} \sum_{\substack{n \leq x \\ \Omega(n)=k \\ k \geq 2 \\ 2 \nmid k}} \frac{\log(\frac{x}{n})}{\log 2}$$

is better than

$$\pi(x) = \frac{\theta(x)}{\log x} + O\left(\frac{x}{\log^2 x}\right)$$

found in the literature.

1.

REFERENCES

1. Rosser, J. Barkley and Schoenfeld, Lowell and others, *Approximate formulas for some functions of prime numbers.*, Illinois Journal of Mathematics, vol. 6.1, University of Illinois at Urbana-Champaign, Department of Mathematics, 2015, pp.64–94, 1962.
2. P. Dusart, *Estimates of some functions over primes without RH*, arXiv preprint arXiv:1002.0442, 2010.
3. Montgomery, H.L, and Vaughan, R.C, *Multiplicative number theory 1:Classical theory.* vol.97, Cambridge university press, 2006.
4. Robbins, Herbert, *A remark on Stirling's formula*, Amer. Math. Mon., vol. 62.1, 1955, pp.26–29

DEPARTMENT OF MATHEMATICS, AFRICAN INSTITUTE FOR MATHEMATICAL SCIENCE, GHANA
Email address: theophilus@aims.edu.gh/emperordagama@yahoo.com