AN AVERAGE ESTIMATE FOR A CERTAIN INTEGRAL OVER INTEGERS WITH SPECIFIED NUMBER OF PRIME FACTORS

THEOPHILUS AGAMA

Abstract. Using some properties of the prime, we establish an asymptotic for the sum
\[ \sum_{k \geq 2} \left( \frac{1}{2} + o(1) \right) \int_{2}^{x} \frac{\pi_k(t)}{t} dt = \frac{x}{2} + O\left( \frac{x}{\log x} \right) . \]

1. Introduction and statement

A slight generalization of the prime number theorem is the estimate
\[ \pi_k(x) = (1 + o(1)) \frac{x \log \log k^{-1} x}{(k - 1)! \log x} \]
for a fixed \( k \), where \( \pi_k(x) = \sum_{\Omega(n) = k, n \leq x} 1 \) [1]. Since for the case \( k = 1 \), we obtain the well-known weaker estimate
\[ \pi(x) = (1 + o(1)) \frac{x}{\log x} , \]
the prime number theorem [2]. It is generally not known to hold uniformly in \( k \). However, It is known to hold for all \( k \) such that \( k \leq \log \log x \). This makes an attempt to estimate the sum directly
\[ \sum_{k \geq 2} \left( \frac{1}{2} + o(1) \right) \int_{2}^{x} \frac{\pi_k(t)}{t} dt \]
a non-trivial task coupled with some convergence issues. However, we can get around this problem by using an earlier result of the author in a careful manner, without having to exploit the estimate for \( \pi_k(x) \) for a uniform \( k \).

2. Notation

Throughout this paper a prime number will either be denoted by \( p \) or the subscripts of \( p \). Any other letter will be clarified. The function \( \Omega(n) := \sum_{p \mid n} 1 \) 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: December 25, 2018.
2000 Mathematics Subject Classification. Primary 54C40, 14E20; Secondary 46E25, 20C20.
Key words and phrases. prime, theta.
3. Preliminary results

**Lemma 3.1.** For all $x \geq 2$

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

**Proof.** For a proof see for instance [2].

**Theorem 3.2.** For every positive integer $x$

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

**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 $2, 2^2, \ldots, 2^b$ such that $2^b < x$; clearly there are $\left\lfloor \frac{\log x}{\log 2} \right\rfloor$ such number of terms in the sequence. Again consider those of the form $3 \cdot 2, \ldots, 3 \cdot 2^b$ such that $3 \cdot 2^b < x$; clearly there are $\left\lfloor \frac{\log (x/3)}{\log 2} \right\rfloor$ such terms in this sequence. We terminate the process by considering those of the form $2 \cdot j, \ldots, 2^b \cdot j$ such that $(2,j) = 1$ and $2^b \cdot j < x$; Clearly there are $\left\lfloor \frac{\log (x/j)}{\log 2} \right\rfloor$ 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 $(x-1)/2 + \frac{1}{4} = \sum_{\substack{j \leq x \\ (2,j)=1}} \left\lfloor \frac{\log (x/j)}{\log 2} \right\rfloor$. By combining both cases, the result follows immediately.

**Remark 3.3.** As a consequence of Stirlings formula, we obtain the following useful estimate.

**Corollary 3.1.**

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

\[(3.1) \quad \sum_{n \leq x} \log n = x \log x - x + \frac{1}{2} \log x + \log(\sqrt{2\pi}) + O\left(\frac{1}{x}\right).\]

Also from Theorem 3.2, we obtain

\[(3.2) \quad \sum_{n \leq x} \log n = x \log x - x \log 2 - \log 2 \sum_{n \leq x} \left\lfloor \log \left(\frac{x}{n}\right) \log 2 \right\rfloor + O(1).\]

Comparing equation (3.1) and (3.2), we have

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

thereby establishing the estimate. \qed

As is crucial to our problem, we obtain an exact formula for the prime counting function \(\pi(x)\) in terms of the number of prime factor counting multiplicity function and some other elementary functions in the following sequel.

**Theorem 3.4.** 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:

\[H(x) := \sum_{p \leq x} \left\lfloor \frac{\log \left(\frac{x}{p}\right)}{\log 2} \right\rfloor, \quad G(x) := \frac{\log x}{\log 2} + \sum_{n \leq x} \left\lfloor \frac{\log \left(\frac{x}{n}\right)}{\log 2} \right\rfloor, \quad T(x) := \left\lfloor \frac{\log \left(\frac{x}{2}\right)}{\log 2} \right\rfloor\]

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

denotes the fractional part of any real number.

Proof. We use Theorem 3.2 to establish an explicit formula for the prime counting function. Theorem 3.2 gives \((x - 1)/2 = \sum_{n \leq x} \left\lfloor \frac{\log \left(\frac{x}{n}\right)}{\log 2} \right\rfloor - (1 + (-1)^x)1/4,

which can then be recast as

\[(x - 1)/2 = \sum_{p \leq x} \left\lfloor \frac{\log \left(\frac{x}{p}\right)}{\log 2} \right\rfloor - \left\lfloor \frac{\log \left(\frac{x}{2}\right)}{\log 2} \right\rfloor + \frac{\log x}{\log 2} + \sum_{n \leq x} \left\lfloor \frac{\log \left(\frac{x}{n}\right)}{\log 2} \right\rfloor - \left(1 + (-1)^x\right)/4,\]
where $p$ runs over the primes. It follows by further simplification that

$$(x - 1)/2 = \frac{1}{\log 2} \left( \log x \sum_{p \leq x} \frac{1}{\log p} - \sum_{p \leq x} \left\lfloor \frac{\log(x/p)}{\log 2} \right\rfloor + \frac{\log x}{\log 2} - \frac{\log(x/2)}{\log 2} \right)$$

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

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

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

and $\theta(x) := \sum_{p \leq x} \log p$.

and the result follows immediately. □

**Remark 3.5.** Using Theorem 3.4, we relate the prime counting function to the sum we seek to estimate in the following sequel.

### 4. Main result

**Theorem 4.1.** For all positive integers $x \geq 2$

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

where

$$\Theta(x) = \frac{\theta(x)}{\log x} + \frac{x}{2 \log x} - \frac{1}{4} \frac{1}{\log x} \sum_{k \geq 2} \left( \frac{1}{2} + o(1) \right) \int_{2}^{x} \frac{\pi_k(t)}{t} dt.$$

**Proof.** By Theorem 3.4, 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

$$H(x) := \sum_{p \leq x} \left\{ \frac{\log(x/p)}{\log 2} \right\}, \quad G(x) := \left\lfloor \frac{\log x}{\log 2} \right\rfloor + \sum_{n \leq x \atop \Omega(n) = k \atop k \geq 2} \left\lfloor \frac{\log(n/\pi)}{\log 2} \right\rfloor, \quad T(x) := \left\lfloor \frac{\log(x)}{\log 2} \right\rfloor.$$
Now, we estimate the term $H(x) - G(x) + T(x)$. Clearly we can write

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

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

$$= - \frac{\log x}{\log 2} \sum_{k \geq 2} \left( \frac{1}{2} + o(1) \right) \sum_{\substack{n \leq x \\Omega(n) = k \\gcd(2,n) = 1}} 1 + \frac{\log x}{\log 2} \sum_{k \geq 2} \left( \frac{1}{2} + o(1) \right) \sum_{\substack{n \leq x \\Omega(n) = k \\gcd(2,n) = 1}} 1$$

$$- \frac{1}{\log 2} \sum_{k \geq 2} \left( \frac{1}{2} + o(1) \right) \int_{\frac{1}{2}}^{x} \pi_k(t) \frac{dt}{t} + \sum_{\substack{n \leq x \\Omega(n) = k \\gcd(2,n) = 1}} \left\lfloor \frac{\log \left( \frac{x}{n} \right)}{\log 2} \right\rfloor + O(1)$$

$$= - \frac{1}{\log 2} \sum_{k \geq 2} \left( \frac{1}{2} + o(1) \right) \int_{\frac{1}{2}}^{x} \pi_k(t) \frac{dt}{t} + \sum_{\substack{n \leq x \\Omega(n) = k \\gcd(2,n) = 1}} \left\lfloor \frac{\log \left( \frac{x}{n} \right)}{\log 2} \right\rfloor + O(1).$$

It follows that

$$H(x) - G(x) + T(x) = \sum_{n \leq x} \left\lfloor \frac{\log \left( \frac{x}{n} \right)}{\log 2} \right\rfloor - \frac{1}{\log 2} \sum_{k \geq 2} \left( \frac{1}{2} + o(1) \right) \int_{\frac{1}{2}}^{x} \pi_k(t) \frac{dt}{t} + O(1)$$

$$= \frac{x}{\log 4} - \frac{x}{2} - \frac{\log x}{\log 16} - \frac{1}{\log 2} \sum_{k \geq 2} \left( \frac{1}{2} + o(1) \right) \int_{\frac{1}{2}}^{x} \pi_k(t) \frac{dt}{t} + O(1)$$

where we have used Corollary 3.1 and the proof is complete. □

**Corollary 4.1.** The estimate

$$\sum_{k \geq 2} \left( \frac{1}{2} + o(1) \right) \int_{\frac{1}{2}}^{x} \pi_k(t) \frac{dt}{t} = \frac{x}{2} + O \left( \frac{x}{\log x} \right)$$

holds.

**Proof.** By Theorem 4.1, we can write

$$\frac{1}{\log x} \sum_{k \geq 2} \left( \frac{1}{2} + o(1) \right) \int_{\frac{1}{2}}^{x} \pi_k(t) \frac{dt}{t} = \frac{\theta(x)}{\log x} - \pi(x) + \frac{x}{2 \log x} - \frac{1}{4} + O \left( \frac{1}{\log x} \right).$$
Using Lemma 3.1, we can write
\[
\frac{1}{\log x} \sum_{k \geq 2} \left( \frac{1}{2} + o(1) \right) \frac{\pi_k(t)}{t} dt = \frac{x}{2 \log x} + O\left( \frac{x}{\log^2 x} \right)
\]
and the result follows immediately. \qed

5. Final remarks

The estimate
\[
\sum_{k \geq 2} \left( \frac{1}{2} + o(1) \right) \int_2^{\pi} \frac{\pi_k(t)}{t} dt = \frac{x}{2} + O\left( \frac{x}{\log x} \right)
\]
established, we hope, might be useful in other related problems. We also note that standard heuristics reveals the error term to the above problem should be \( \ll \sqrt{x} \).
Thus we state as a conjecture

Conjecture 5.1.

\[
\sum_{k \geq 2} \left( \frac{1}{2} + o(1) \right) \int_2^{\pi} \frac{\pi_k(t)}{t} dt = \frac{x}{2} + O(\sqrt{x}).
\]

1.

References


Department of Mathematics, African Institute for Mathematical science, Ghana
E-mail address: theophilus@aims.edu.gh/emperordagama@yahoo.com