On a sequence involving the prime numbers

Christian Axler
April 20, 2015

Abstract

In this paper we study a sequence involving the prime numbers by deriving two asymptotic formulas and finding new upper and lower bounds, which improve the currently known estimates.

1 Introduction

In this paper, we study the difference

\[ C_n = n p_n - \sum_{k \leq n} p_k \]

(see also [3]), where \( p_n \) is the nth prime number, by proving two asymptotic formulas and finding lower and upper bounds for \( C_n \).

2 Two asymptotic formulas for \( C_n \)

Let \( m \in \mathbb{N} \). By [3], there exist unique \( a_{is} \in \mathbb{Q} \), where \( a_{ss} = 1 \) for all \( 1 \leq s \leq m \), such that

\[ p_n = n \left( \log n + \log \log n - 1 + \sum_{s=1}^{m} \frac{(-1)^s + 1}{s \log^s n} \sum_{i=0}^{s} a_{is} (\log \log n)^i \right) + O(c_m(n)), \]

where

\[ c_m(n) = \frac{n(\log \log n)^{m+1}}{\log^{m+1} n}. \]

We set

\[ h_m(n) = \sum_{j=1}^{m} \frac{(j-1)!}{2^j \log^j n}. \]

Further, we recall the following definition from [2].

Definition. Let \( s, i, j, r \in \mathbb{N}_0 \) with \( j \geq r \). We define the integers \( b_{s,i,j,r} \in \mathbb{Z} \) as follows:

- If \( j = r = 0 \), then
  \[ b_{s,i,0,0} = 1. \] (2)

- If \( j \geq 1 \), then
  \[ b_{s,i,j,j} = b_{s,i,j-1,j-1} \cdot (-i + j - 1). \] (3)

- If \( j \geq 1 \), then
  \[ b_{s,i,j,0} = b_{s,i,j-1,0} \cdot (s + j - 1). \] (4)

- If \( j > r \geq 1 \), then
  \[ b_{s,i,j,r} = b_{s,i,j-1,r} \cdot (s + j - 1) + b_{s,i,j-1,r-1} \cdot (-i + r - 1). \] (5)

Using (1) and Theorem 2.5 of [2], we obtain the first asymptotic formula for \( C_n \).
Theorem 2.1. Let \( m \in \mathbb{N} \). Then,
\[
C_n = \frac{n^2}{2} \left( \log n + \log \log n - \frac{1}{2} + h_m(n) \right) \\
+ \frac{n^2}{2} \sum_{s=1}^{\infty} \frac{(-1)^{s+1}}{s \log^s n} \sum_{i=0}^{s} a_{is} \left( 2(\log \log n)^i - \sum_{j=0}^{m-2} \sum_{r=0}^{\min(i,j)} b_{s,i,j,r} (\log \log n)^{i-r} \right) + O(nc_m(n)).
\]

Proof. First, we multiply the asymptotic formula (1) with \( n \). Then, we subtract the asymptotic formula for \( \sum_{k \leq n} p_k \) from [2, Theorem 2.5] to obtain our proposition.

Corollary 2.2. Let \( m \in \mathbb{N} \). Then there are unique monic polynomials \( U_s \in \mathbb{Q}[x] \), where \( 1 \leq s \leq m \) and \( \deg(U_s) = s \), such that
\[
C_n = \frac{n^2}{2} \left( \log n + \log \log n - \frac{1}{2} + \sum_{s=1}^{m} \frac{(-1)^{s+1} U_s(\log \log n)}{s \log^s n} \right) + O(nc_m(n)).
\]
In particular, we have \( U_1(x) = x - 3/2 \) and \( U_2(x) = x^2 - 5x + 15/2 \).

Proof. Since \( a_{ss} = 1 \) and \( b_{s,s,0,0} = 1 \), the first claim follows from Theorem 2.1. Now let \( m = 2 \). By [3], we have \( a_{01} = -2, a_{11} = 1, a_{02} = 11, a_{12} = -6 \) and \( a_{22} = 1 \). Further, we use the formulas (2)–(5) to compute the integers \( b_{s,i,j,r} \). Then, using Theorem 2.1, we obtain the polynomials \( U_1 \) and \( U_2 \).

To find another asymptotic formula for \( C_n \), we obtain the following identity, which leads to a possibility to estimate \( C_n \) by using estimates for \( \pi(x) \).

Lemma 2.3. For all \( n \in \mathbb{N} \),
\[
C_n = \int_{2}^{p_n} \pi(x) \, dx.
\]

Proof. See [4].

Now we give certain rules of integration.

Lemma 2.4. Let \( x, a \in \mathbb{R} \) with \( x \geq a > 1 \). Then,
\[
\int_{a}^{x} \frac{t \, dt}{\log t} = \text{li}(x^2) - \text{li}(a^2).
\]

Proof. See [4, Lemme 1.6].

Lemma 2.5. Let \( x, a \in \mathbb{R} \) with \( x \geq a > 1 \). Then,
\[
\int_{a}^{x} \frac{t \, dt}{\log^2 t} = 2 \text{li}(x^2) - 2 \text{li}(a^2) - \frac{x^2}{\log x} + \frac{a^2}{\log a}.
\]

Proof. See [4, Lemme 1.6].

Lemma 2.6. Let \( r, s \in \mathbb{R} \) with \( s \geq r > 1 \) and \( n \in \mathbb{N} \). Then,
\[
\int_{r}^{s} \frac{x \, dx}{\log^{s+1} x} = \frac{r^2}{n \log^n r} - \frac{s^2}{n \log^n s} + \frac{2}{n} \int_{r}^{s} \frac{x \, dx}{\log^n x}.
\]

Proof. Integration by parts.

Lemma 2.7. Let \( r, s \in \mathbb{R} \) with \( s \geq r > 1 \). Then, for all \( m \in \mathbb{N} \) with \( m \geq 2 \) we have
\[
\int_{r}^{s} \frac{x \, dx}{\log^n x} = \frac{2^{m-2} - (m-1)!}{(m-1)!} \int_{r}^{s} \frac{x \, dx}{\log^2 x} - \frac{2^{m-1-k} (k-1)!}{(m-1)! (m-k)!} \left( \frac{s^2}{\log^k s} - \frac{r^2}{\log^k r} \right).
\]
Proof. By induction on $m$. □

The next proposition plays an important role for the proof of the second asymptotic formula for $C_n$.

**Proposition 2.8.** Let $m \in \mathbb{N}$ with $m \geq 2$. Let $a_2, \ldots, a_m \in \mathbb{R}$ and $r, s \in \mathbb{R}$ with $s \geq r > 1$. Then,

$$
\sum_{k=2}^{m} a_k \int_{r}^{s} \frac{x \ dx}{\log^{k} x} = t_{m-1,1} \int_{r}^{s} \frac{x \ dx}{\log^{2} x} - \sum_{k=2}^{m-1} \sum_{k=2}^{m-1, k} \left( \frac{s^2}{\log^2 s} - \frac{r^2}{\log^2 r} \right),
$$

where

$$
t_{i,j} := (j-1)! \sum_{l=j}^{i} \frac{2l-j+1}{l!}.
$$

Proof. If $m = 2$, the claim is obviously true. By induction hypothesis, we have

$$
\sum_{k=2}^{m+1} a_k \int_{r}^{s} \frac{x \ dx}{\log^{k} x} = t_{m-1,1} \int_{r}^{s} \frac{x \ dx}{\log^{2} x} - \sum_{k=2}^{m} \sum_{k=2}^{m-1, k} \left( \frac{s^2}{\log^2 s} - \frac{r^2}{\log^2 r} \right) + a_{m+1} \int_{r}^{s} \frac{x \ dx}{\log^{m+1} x}.
$$

By Lemma 2.6 we get

$$
\sum_{k=2}^{m+1} a_k \int_{r}^{s} \frac{x \ dx}{\log^{k} x} = t_{m-1,1} \int_{r}^{s} \frac{x \ dx}{\log^{2} x} - \sum_{k=2}^{m} \sum_{k=2}^{m-1, k} \left( \frac{s^2}{\log^2 s} - \frac{r^2}{\log^2 r} \right) + \frac{2a_{m+1}}{m} \int_{r}^{s} \frac{x \ dx}{\log^{m} x}.
$$

Now we can use Lemma 2.6 and the equality $t_{m-1,1} + 2m^{-1}a_{m+1}/m = t_{m-1,1}$ to obtain

$$
\sum_{k=2}^{m+1} a_k \int_{r}^{s} \frac{x \ dx}{\log^{k} x} = t_{m,1} \int_{r}^{s} \frac{x \ dx}{\log^{2} x} - \sum_{k=2}^{m} \left( \frac{2^{m-k}a_{m+1}(k-1)!}{m!} t_{m-1,1} + \sum_{k=2}^{m-1, k} \left( \frac{s^2}{\log^2 s} - \frac{r^2}{\log^2 r} \right) \right) - \frac{a_{m+1}(m-1)!}{m!} \left( \frac{s^2}{\log^2 s} - \frac{r^2}{\log^2 r} \right).
$$

Since we have

$$
\frac{2^{m-k}a_{m+1}(k-1)!}{m!} t_{m-1,1} + \sum_{k=2}^{m-1, k} \left( \frac{s^2}{\log^2 s} - \frac{r^2}{\log^2 r} \right) = t_{m,m} = a_{m+1}(m-1)!/(m!),
$$

our proposition is proved. □

Now we give another asymptotic formula for $C_n$.

**Theorem 2.9.** Let $m \in \mathbb{N}$. Then,

$$
C_n = \sum_{k=1}^{m-1} (k-1)! \left( 1 - \frac{1}{2^k} \right) \frac{p_n}{\log^{k} p_n} + O \left( \frac{p_n}{\log^{m+1} p_n} \right).
$$

Proof. First we recall a well-known asymptotic formula for the prime counting function $\pi(x)$; i.e.

$$
\pi(x) = \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x} + \frac{6x}{\log^4 x} + \ldots + (m-1)! x \log^{m-1} x + O \left( \frac{x}{\log^{m+1} x} \right).
$$

Using (8) and Lemma 2.6 we get

$$
C_n = \sum_{k=1}^{m} (k-1)! \int_{2}^{p_n} \frac{x \ dx}{\log^{k} x} + O \left( \int_{2}^{p_n} \frac{x \ dx}{\log^{m+1} x} \right).
$$

Integration by parts gives

$$
C_n = \sum_{k=1}^{m} (k-1)! \int_{2}^{p_n} \frac{x \ dx}{\log^{k} x} + O \left( \frac{p_n}{\log^{m+1} p_n} \right).\]
We can apply Proposition 2.8 to get
\[ C_n = \int_2^{p_n} \frac{x \, dx}{\log x} + (2^m - 1) \int_2^{p_n} \frac{x \, dx}{\log^2 x} - \sum_{k=2}^{m-1} \frac{(k-1)! (2^{m-k} - 1) p_n^2}{\log^k p_n} + O \left( \frac{p_n^2}{\log^m p_n} \right). \]

Using Lemma 2.4 and Lemma 2.5, we get
\[ C_n = (2^m - 1) \log(p_n^2) - \sum_{k=1}^{m-1} \frac{(k-1)! (2^{m-k} - 1) p_n^2}{\log^k p_n} + O \left( \frac{p_n^2}{\log^m p_n} \right). \]

Now we use the asymptotic formula
\[ \log(x) = x + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x} + \frac{6x}{\log^4 x} + \ldots + \frac{(m-1)! x}{\log^m x} + O \left( \frac{x}{\log^{m+1} x} \right), \tag{9} \]
which can be showed by integration by parts, to obtain the equality
\[ C_n = (2^m - 1) \sum_{k=1}^{m-1} \frac{(k-1)! p_n^2}{2^k \log^k p_n} - \sum_{k=1}^{m-1} \frac{(k-1)! (2^{m-k} - 1) p_n^2}{\log^k p_n} + O \left( \frac{p_n^2}{\log^m p_n} \right). \]
and our theorem is proved. \( \square \)

Using (8), we get the following corollary.

**Corollary 2.10.** Let \( m \in \mathbb{N} \). Then,
\[ \sum_{k \leq n} p_k = \pi(p_n^2) + O \left( \frac{p_n^2}{\log^m p_n} \right). \]

**Proof.** From Theorem 2.9 and the definition of \( C_n \) it follows that
\[ \sum_{k \leq n} p_k = n p_n - \sum_{k=1}^{m-1} \frac{(k-1)! p_n^2}{\log^k p_n} + \sum_{k=1}^{m-1} \frac{(k-1)! p_n^2}{2^k \log^k p_n} + O \left( \frac{p_n^2}{\log^m p_n} \right). \]
Since \( n = \pi(p_n) \), we obtain
\[ \sum_{k \leq n} p_k = \pi(p_n) p_n - \sum_{k=1}^{m-1} \frac{(k-1)! p_n^2}{\log^k p_n} + \sum_{k=1}^{m-1} \frac{(k-1)! p_n^2}{2^k \log^k p_n} + O \left( \frac{p_n^2}{\log^m p_n} \right). \]
Using (8), we get the equality
\[ \sum_{k \leq n} p_k = \sum_{k=1}^{m-1} \frac{(k-1)! p_n^2}{2^k \log^k p_n} + O \left( \frac{p_n^2}{\log^m p_n} \right) = \pi(p_n^2) + O \left( \frac{p_n^2}{\log^m p_n} \right) \]
and the corollary is proved. \( \square \)

Comparing (8) and (9), we see that \( \pi(x) \) and \( \log(x) \) have the same asymptotic formula. Hence, using Corollary 2.10 we also get the following result on the sum of the first \( n \) prime numbers.

**Corollary 2.11.** Let \( m \in \mathbb{N} \). Then,
\[ \sum_{k \leq n} p_k = \log(p_n^2) + O \left( \frac{p_n^2}{\log^m p_n} \right). \]
3 A lower bound for $C_n$

Let $m \in \mathbb{N}$ with $m \geq 2$ and let $a_2, \ldots, a_m, x_0, y_0 \in \mathbb{R}$, so that

\[
\pi(x) \geq \frac{x}{\log x} + \sum_{k=2}^{m} \frac{a_k x}{\log^k x}\]

for every $x \geq x_0$ and

\[
\text{li}(x) \geq \sum_{j=1}^{m-1} \frac{(j-1)! x}{\log^j x}\]

for every $x \geq y_0$. Then, we obtain the following lower bound for $C_n$.

**Theorem 3.1.** If $n \geq \max\{\pi(x_0) + 1, \pi(\sqrt{y_0}) + 1\}$, then

\[
C_n \geq d_0 + \sum_{k=1}^{m-1} \left( \frac{(k-1)!}{2^k} (1 + 2t_{k-1,1}) \right) p_n^2 \log^k p_n,
\]

where $t_{i,j}$ is defined as in (6) and $d_0$ is given by

\[
d_0 = d_0(m, a_2, \ldots, a_m, x_0) = \int_{x_0}^{x_0} \pi(x) \, dx - (1 + 2t_{m-1,1}) \text{li}(x_0^2) + \sum_{k=1}^{m-1} t_{m-1,k} \frac{x_0^2}{\log^k x_0}.
\]

**Proof.** Since $p_n \geq x_0$, we use Lemma 2.3 and (10) to obtain

\[
C_n \geq \int_{x_0}^{x_0} \pi(x) \, dx + \int_{x_0}^{p_n} x \, dx + \sum_{k=2}^{m} a_k \int_{x_0}^{p_n} \frac{x \, dx}{\log^k x}.
\]

Now, we apply Lemma 2.4 and Proposition 2.8 to get

\[
C_n \geq \int_{x_0}^{x_0} \pi(x) \, dx - \text{li}(x_0^2) + \text{li}(p_n^2) + t_{m-1,1} \int_{x_0}^{p_n} \frac{x \, dx}{\log^2 x} - \sum_{k=1}^{m-1} t_{m-1,k} \left( \frac{p_n^2}{\log^k p_n} - \frac{x_0^2}{\log^k x_0} \right).
\]

Using Lemma 2.5 we obtain

\[
C_n \geq d_0 + (1 + 2t_{m-1,1}) \text{li}(p_n^2) - \sum_{k=1}^{m-1} t_{m-1,k} \frac{p_n^2}{\log^k p_n}.
\]

Since $p_n^2 \geq y_0$, we use (11) to conclude

\[
C_n \geq d_0 + \sum_{k=1}^{m-1} \left( \frac{(k-1)!}{2^k} (1 + 2t_{k-1,1}) \right) \frac{p_n^2}{\log^k p_n}
\]

and it remains to use the definition of $t_{i,j}$.

4 An upper bound for $C_n$

Next, we derive for the first time an upper bound for $C_n$. Let $m \in \mathbb{N}$ with $m \geq 2$ and let $a_2, \ldots, a_m, x_1 \in \mathbb{R}$ so that

\[
\pi(x) \leq \frac{x}{\log x} + \sum_{k=2}^{m} \frac{a_k x}{\log^k x}\]

for every $x \geq x_1$ and let $\lambda, y_1 \in \mathbb{R}$ so that

\[
\text{li}(x) \leq \sum_{j=1}^{m-2} \frac{(j-1)! x}{\log^j x} + \frac{\lambda x}{\log^{m-1} x}.
\]
for every $x \geq y_1$. Setting
\[ d_1 := d_1(m, a_2, \ldots, a_m, x_1) = \int_2^{x_1} \pi(x) \, dx - (1 + 2t_{m-1,1}) \log(x_1^2) + \sum_{k=1}^{m-1} t_{m-1, k} \frac{x_k^2}{\log^k x_1}, \]
where $t_{m-1, k}$ is defined by (6), we obtain the following

**Theorem 4.1.** If $n \geq \max\{\pi(x_1) + 1, \pi(\sqrt{y_1}) + 1\}$, then
\[ C_n \leq d_1 + \sum_{k=1}^{m-2} \left( \frac{(k-1)!}{2^k} (1 + 2t_{k-1,1}) \right) \frac{p_k^2}{\log^k p_n} + \left( \frac{(1 + 2t_{m-1,1}) \lambda}{2^{m-1}} - \frac{a_m}{m-1} \right) \frac{p_n^2}{\log^{m-1} p_n}. \]

**Proof.** Since $p_n \geq x_1$, we use Lemma 2.3 and (12) to get
\[ C_n \leq \int_2^{x_1} \pi(x) \, dx + \int_{x_1}^{p_n} x \frac{dx}{\log x} + \sum_{k=2}^{m} a_k \int_{x_1}^{p_n} x \frac{dx}{\log^k x}. \]
We apply Lemma 2.4 and Proposition 2.8 to obtain
\[ C_n \leq \int_2^{x_1} \pi(x) \, dx - \log(x_1^2) + \log(p_n^2) + t_{m-1, 1} \int_{x_1}^{p_n} x \frac{dx}{\log x} - \sum_{k=2}^{m-1} t_{m-1, k} \left( \frac{p_k^2}{\log^k p_n} - \frac{x_1^2}{\log^k x_1} \right). \]
Using Lemma 2.5 we get
\[ C_n \leq d_1 + (1 + 2t_{m-1, 1}) \log(p_n^2) - \sum_{k=1}^{m-1} t_{m-1, k} \frac{p_k^2}{\log^k p_n}. \]
Now we can use the inequality (13) to obtain
\[ C_n \leq d_1 + \sum_{k=1}^{m-2} \left( \frac{(k-1)!}{2^k} + \frac{t_{m-1,1}(k-1)!}{2^{k-1}} - t_{m-1, k} \right) \frac{p_k^2}{\log^k p_n} + \left( \frac{(1 + 2t_{m-1,1}) \lambda}{2^{m-1}} - t_{m-1, m-1} \right) \frac{p_n^2}{\log^{m-1} p_n} \]
and it remains to use the definition of $t_{ij}$. \qed

## 5 Numerical results

By setting $m = 8$ in Theorem 2.9 we obtain
\[ C_n = \frac{p_n^2}{2 \log p_n} + \frac{3p_n^2}{4 \log^2 p_n} + \frac{7p_n^2}{4 \log^3 p_n} + \chi(n) + O\left( \frac{p_n^2}{\log^4 p_n} \right), \]
where $\chi(n)$ is defined by
\[ \chi(n) = \frac{45p_n^2}{8 \log^4 p_n} + \frac{93p_n^2}{4 \log^4 p_n} + \frac{945p_n^2}{8 \log^5 p_n} + \frac{5715p_n^2}{8 \log^6 p_n}. \]

### 5.1 An explicit lower bound for $C_n$

Dusart [4] proved, that
\[ C_n \geq c + \frac{p_n^2}{2 \log p_n} + \frac{3p_n^2}{4 \log^2 p_n} \]
for every $n \geq 109$, where $c \approx -47.1$. The goal of this subsection is to improve inequality (14). In order to do this, we first give two lemmata concerning explicit estimates for $\log(x)$.  

6
Lemma 5.1. If \( x \geq 4171 \), then
\[
\operatorname{li}(x) \geq \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x} + \frac{6x}{\log^4 x} + \frac{24x}{\log^5 x} + \frac{120x}{\log^6 x} + \frac{720x}{\log^7 x} + \frac{5040x}{\log^8 x}.
\]

Proof. We denote the right hand side by \( \alpha(x) \). Let \( f(x) = \operatorname{li}(x) - \alpha(x) \). Then, \( f(4171) \geq 0.00019 \) and \( f'(x) = 40320/\log^5 x \), and our lemma is proved. \( \square \)

Lemma 5.2. If \( x \geq 10^{16} \), then
\[
\operatorname{li}(x) \leq \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x} + \frac{6x}{\log^4 x} + \frac{24x}{\log^5 x} + \frac{120x}{\log^6 x} + \frac{900x}{\log^8 x}.
\]

Proof. Similarly to the proof of Lemma 5.1. \( \square \)

Setting
\[
\Theta(n) = \frac{43.6p_n^2}{8 \log^4 p_n} + \frac{90.9p_n^2}{4 \log^4 p_n} + \frac{927.5p_n^2}{8 \log^6 p_n} + \frac{10.5p_n^2}{3 \log^7 p_n} + \frac{4942.21875p_n^2}{\log^8 p_n},
\]
we get the following improvement of (14).

Proposition 5.3. If \( n \geq 52703656 \), then
\[
C_n \geq d_0 + \frac{p_n^2}{2 \log p_n} + \frac{3p_n^2}{4 \log^2 p_n} + \frac{7p_n^2}{4 \log^3 p_n} + \Theta(n).
\]

Proof. We choose \( m = 9 \), \( a_2 = 1 \), \( a_3 = 2 \), \( a_4 = 5.65 \), \( a_5 = 23.65 \), \( a_6 = 118.25 \), \( a_7 = 709.5 \), \( a_8 = 4966.5 \), \( a_9 = 0 \), \( x_0 = 1332450001 \) and \( y_0 = 4171 \). By (11), we obtain the inequality (10) for every \( x \geq x_0 \) and (11) holds for every \( x \geq y_0 \) by Lemma 5.1. Substituting these values in Theorem 3.1, we get
\[
C_n \geq d_0 + \frac{p_n^2}{2 \log p_n} + \frac{3p_n^2}{4 \log^2 p_n} + \frac{7p_n^2}{4 \log^3 p_n} + \Theta(n)
\]
for every \( n \geq 66773605 \), where \( d_0 = d_0(9, 1, 2, 5.65, 23.65, 118.25, 709.5, 4966.5, 0, x_0) \) is given by
\[
d_0 = \int_2^{x_0} \pi(x) \, dx - \frac{753.1}{3} \ln(x_0^2) + \frac{375.05x_0^2}{3 \log x_0} + \frac{186.025x_0^2}{3 \log^2 x_0} + \frac{183.025x_0^2}{3 \log^3 x_0} + \frac{88.6875x_0^2}{\log^4 x_0}
\]
\[
+ \frac{165.55x_0^2}{\log^5 x_0} + \frac{354.75x_0^2}{\log^6 x_0} + \frac{709.5x_0^2}{\log^7 x_0} + \frac{709.5x_0^2}{\log^8 x_0}
\]
Since \( x_0^2 \geq 10^{16} \), we obtain using Lemma 5.2,
\[
d_0 \geq \int_2^{x_0} \pi(x) \, dx - \frac{x_0^2}{2 \log x_0} - \frac{3x_0^2}{4 \log^2 x_0} - \frac{7x_0^2}{4 \log^3 x_0} - \frac{5.45x_0^2}{\log^4 x_0} - \frac{22.725x_0^2}{\log^5 x_0} - \frac{115.9375x_0^2}{\log^6 x_0} - \frac{1055.578125x_0^2}{\log^7 x_0}.
\]
Using \( \log x_0 \geq 21.01027 \), we get
\[
d_0 \geq \int_2^{x_0} \pi(x) \, dx - 4.22512933 \cdot 10^{16} - 0.30164729 \cdot 10^{16} - 0.03349997 \cdot 10^{16} - 0.0049656 \cdot 10^{16}
\]
\[
- 0.00098548 \cdot 10^{16} - 0.0002393 \cdot 10^{16} - 0.0001037 \cdot 10^{16}
\]
\[
= \int_2^{x_0} \pi(x) \, dx - 4.56657067 \cdot 10^{16}. \quad (16)
\]
Since \( x_0 = p_{66773604} \), we obtain using Lemma 2.3 and a computer,
\[
\int_2^{x_0} \pi(x) \, dx = 45665745738169817.
\]
Hence, by (16), we get \( d_0 \geq 3.9 \cdot 10^{10} > 0 \). So we obtain the asserted inequality for every \( n \geq 66773605 \). For every \( 52703656 \leq n \leq 66773604 \) we check the inequality with a computer. \( \square \)
5.2 An explicit upper bound for $C_n$

We begin with the following lemma.

**Lemma 5.4.** If $x \geq 10^{18}$, then

$$\ln(x) \leq \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x} + \frac{6x}{\log^4 x} + \frac{24x}{\log^5 x} + \frac{120x}{\log^6 x} + \frac{720x}{\log^7 x} + \frac{6300x}{\log^8 x}.$$  

*Proof.* Similarly to the proof of Lemma 5.1.

Using an upper bound for $\pi(x)$ from [1], we obtain the following explicit upper bound for $C_n$, where

$$\Omega(n) = \frac{46.4p_n^2}{8 \log^4 p_n} + \frac{95.1p_n^2}{4 \log^5 p_n} + \frac{962.5p_n^2}{8 \log^6 p_n} + \frac{5809.5p_n^2}{8 \log^7 p_n} + \frac{59424p_n^2}{8 \log^8 p_n}.$$ (17)

**Proposition 5.5.** For every $n \in \mathbb{N}$,

$$C_n \leq \frac{p_n^2}{2 \log p_n} + \frac{3p_n^2}{4 \log^2 p_n} + \frac{7p_n^2}{4 \log^3 p_n} + \Omega(n).$$ (18)

*Proof.* We choose $a_2 = 1$, $a_3 = 2$, $a_4 = 6.35$, $a_5 = 24.35$, $a_6 = 121.75$, $a_7 = 730.5$, $a_8 = 6801.4$, $\lambda = 6300$, $x_1 = 11$ and $y_1 = 10^{18}$. By [1], we get that the inequality (12) holds for every $x \geq x_1$ and by Lemma 5.4 that (13) holds for all $y \geq y_1$. By substituting these values in Theorem 4.1, we get

$$C_n \leq d_1 + \frac{p_n^2}{2 \log p_n} + \frac{3p_n^2}{4 \log^2 p_n} + \frac{7p_n^2}{4 \log^3 p_n} + \Omega(n) - \frac{0.4375p_n^2}{8 \log^5 p_n}$$

for every $n \geq 50847553$, where $d_1 = d_1(9, 1, 2, 6.35, 24.35, 121.75, 730.5, 6801.4, 0, x_1)$ is given by

$$d_1 = \int_{2}^{x_1} \pi(x) \, dx - \frac{950777}{3150} \ln(x_0^2) + \frac{947627x_0^2}{6300 \log x_0} + \frac{941327x_0^2}{12600 \log^2 x_0} + \frac{928727x_0^2}{12600 \log^3 x_0} + \frac{902057x_0^2}{8400 \log^4 x_0} + \frac{425461x_0^2}{2100 \log^5 x_0} + \frac{187163x_0^2}{420 \log^6 x_0} + \frac{34007x_0^2}{35 \log^7 x_0}.$$  

Since $\ln(x_1^2) \geq 34.59$ and $\log x_1 \geq 2.39$, we obtain $d_1 \leq 450$. We define $f(x) = 0.4375x^2/(8 \log^5 x) - 450$. Since $f(6 \cdot 10^6) \geq 100$ and $f'(x) \geq 0$ for every $x \geq e^4$, we get $f(p_n) \geq 0$ for every $n \geq \pi(6 \cdot 10^6) + 1 = 412850$. Now we can use (18) to obtain the claim for every $n \geq 50847553$. For every $1 \leq n \leq 50847534$ we check the asserted inequality with a computer.

References

[1] Axler, C., *New bounds for the prime counting function $\pi(x)$*, arXiv:1409.1780v3 (2015).

[2] ——, *On the sum of the first $n$ prime numbers*, arXiv:1409.1777 (2014).

[3] Cipolla, M., *La determinazione assintotica dell’ $n$° primo numero*, Rend. Accad. Sci. Fis-Mat. Napoli (3) 8 (1902), 132-166.

[4] Dusart, P., *Autour de la fonction qui compte le nombre de nombres premiers*, Dissertation, Université de Limoges, 1998.

[5] Pol, O. E., *Sequence A152535*, The On-Line Encyclopedia of Integer Sequences, http://oeis.org/A152535

Mathematisches Institut, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany

E-mail address: axler@math.uni-duesseldorf.de