Journal of Integer Sequences, Vol. 18 (2015),
Article 15.7.6

On a Sequence Involving Prime Numbers

Christian Axler
Departament of Mathematics
Heinrich-Heine-Universität
40225 Düsseldorf
Germany
Christian.Axler@hhu.de

Abstract

We study a particular sequence $C_n = np_n - \sum_{k \leq n} p_k$, $n \in \mathbb{N}$, involving prime numbers by deriving two asymptotic formulae, and we find a new lower bound for C_n that improves the currently known estimates. Furthermore, for the first time we determine an upper bound for C_n.

1 Introduction

In this paper, we study the sequence $(C_n)_{n \in \mathbb{N}}$ with

$$C_n = np_n - \sum_{k \leq n} p_k,$$

where p_n is the nth prime number. The motivation for considering this special sequence is an inequality conjectured by Mandl [7, p. 1] that asserts that

$$\frac{np_n}{2} - \sum_{k \leq n} p_k \geq 0 \tag{1}$$

for every $n \geq 9$. This inequality originally appeared without proof. In his 1998 thesis [4], Dusart used the equality

$$C_n = \int_2^{p_n} \pi(x)\, dx,$$

where $\pi(x)$ denotes the number of primes $\leq x$, and explicit estimates for the prime counting function $\pi(x)$ to prove that

$$C_n \geq \frac{np_n}{2},$$

which is equivalent to Mandl's inequality (1), for every $n \geq 9$. At the same time, Dusart [4] showed that

$$C_n \geq c + \frac{p_n^2}{2\log p_n} + \frac{3p_n^2}{4\log^2 p_n} \tag{2}$$

for every $n \geq 109$, where $c \doteq -47.1$. The first goal of this article is to study the asymptotic behaviour of the sequence C_n. This is done in the following two theorems.

Theorem 1 (Corollary 7)**.** *For each $s \in \mathbb{N}$ there is a unique monic polynomial U_s of degree s with rational coefficients, so that for every $m \in \mathbb{N}$*

$$C_n = \frac{n^2}{2}\left(\log n + \log\log n - \frac{1}{2} + \sum_{s=1}^{m} \frac{(-1)^{s+1}U_s(\log\log n)}{s\log^s n}\right) + O\left(\frac{n^2(\log\log n)^{m+1}}{\log^{m+1} n}\right).$$

Theorem 2 (Theorem 10)**.** *For each $m \in \mathbb{N}$ we have*

$$C_n = \sum_{k=1}^{m-1}(k-1)!\left(1 - \frac{1}{2^k}\right)\frac{p_n^2}{\log^k p_n} + O\left(\frac{p_n^2}{\log^m p_n}\right). \tag{3}$$

By setting $m = 9$ in (3), we get

$$C_n = \frac{p_n^2}{2\log p_n} + \frac{3p_n^2}{4\log^2 p_n} + \frac{7p_n^2}{4\log^3 p_n} + \chi(n) + O\left(\frac{p_n^2}{\log^9 p_n}\right), \tag{4}$$

where

$$\chi(n) = \frac{45p_n^2}{8\log^4 p_n} + \frac{93p_n^2}{4\log^5 p_n} + \frac{945p_n^2}{8\log^6 p_n} + \frac{5715p_n^2}{8\log^7 p_n} + \frac{80325p_n^2}{16\log^8 p_n}.$$

In view of (4), we improve the inequality (2) by finding the following lower bound for C_n.

Theorem 3 (Proposition 18)**.** *If $n \geq 52703656$, then*

$$C_n \geq \frac{p_n^2}{2\log p_n} + \frac{3p_n^2}{4\log^2 p_n} + \frac{7p_n^2}{4\log^3 p_n} + \Theta(n),$$

where

$$\Theta(n) = \frac{43.6p_n^2}{8\log^4 p_n} + \frac{90.9p_n^2}{4\log^5 p_n} + \frac{927.5p_n^2}{8\log^6 p_n} + \frac{5620.5p_n^2}{8\log^7 p_n} + \frac{79075.5p_n^2}{16\log^8 p_n}.$$

Finally, for the first time we give an upper bound for C_n, by proving the following theorem.

Theorem 4 (Proposition 21)**.** *For every* $n \in \mathbb{N}$,

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

where

$$\Omega(n) = \frac{46.4p_n^2}{8\log^4 p_n} + \frac{95.1p_n^2}{4\log^5 p_n} + \frac{962.5p_n^2}{8\log^6 p_n} + \frac{5809.5p_n^2}{8\log^7 p_n} + \frac{118848p_n^2}{16\log^8 p_n}.$$

2 Two asymptotic formulae for C_n

From here on, we use the following notation. Cipolla [3] showed that for each $s \in \mathbb{N}$ and each $0 \leq i \leq s$ there exist unique rational numbers a_{is}, where $a_{ss} = 1$, such that for every $m \in \mathbb{N}$

$$p_n = n\left(\log n + \log\log n - 1 + \sum_{s=1}^{m} \frac{(-1)^{s+1}}{s\log^s n} \sum_{i=0}^{s} a_{is}(\log\log n)^i\right) + O(c_m(n)), \tag{5}$$

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 5. 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
$$b_{s,i,0,0} = 1. \tag{6}$$
- If $j \geq 1$, then
$$b_{s,i,j,j} = b_{s,i,j-1,j-1} \cdot (-i + j - 1). \tag{7}$$
- If $j \geq 1$, then
$$b_{s,i,j,0} = b_{s,i,j-1,0} \cdot (s + j - 1). \tag{8}$$
- If $j > r \geq 1$, then
$$b_{s,i,j,r} = b_{s,i,j-1,r} \cdot (s + j - 1) + b_{s,i,j-1,r-1} \cdot (-i + r - 1). \tag{9}$$

Using (5) and [2, Thm. 2.5], we obtain the first asymptotic formula for C_n.

Theorem 6. *For each $m \in \mathbb{N}$ we have*

$$C_n = \frac{n^2}{2}\left(\log n + \log\log n - \frac{1}{2} + h_m(n)\right)$$
$$+ \frac{n^2}{2}\sum_{s=1}^{m}\frac{(-1)^{s+1}}{s\log^s n}\sum_{i=0}^{s} a_{is}\left(2(\log\log n)^i - \sum_{j=0}^{m-s}\sum_{r=0}^{\min\{i,j\}}\frac{b_{s,i,j,r}(\log\log n)^{i-r}}{2^j\log^j n}\right)$$
$$+ O(nc_m(n)).$$

Proof. From [2, Thm. 2.5] we know that

$$\sum_{k\le n} p_k = \frac{n^2}{2}\left(g(n) - h_m(n) + \sum_{s=1}^{m}\frac{(-1)^{s+1}}{s\log^s n}\sum_{i=0}^{s} a_{is}\sum_{j=0}^{m-s}\sum_{r=0}^{\min\{i,j\}}\frac{b_{s,i,j,r}(\log\log n)^{i-r}}{2^j\log^j n}\right)$$
$$+ O(nc_m(n)), \tag{10}$$

where $g(n) = \log n + \log\log n - 3/2$. Now we multiply (5) by n and substract (10) to get the result. □

Corollary 7. *For each $s \in \mathbb{N}$ there is a unique monic polynomial U_s of degree s with rational coefficients, so that for every $m \in \mathbb{N}$*

$$C_n = \frac{n^2}{2}\left(\log n + \log\log n - \frac{1}{2} + \sum_{s=1}^{m}\frac{(-1)^{s+1}U_s(\log\log n)}{s\log^s n}\right) + O(nc_m(n)). \tag{11}$$

In particular, $U_1(x) = x - 3/2$ and $U_2(x) = x^2 - 5x + 15/2$.

Proof. Since $a_{ss} = 1$ and $b_{s,s,0,0} = 1$, the formula (11) follows from Theorem 6. Now let $m = 2$. Cipolla [3] showed that $a_{01} = -2$, $a_{11} = 1$, $a_{02} = 11$, $a_{12} = -6$ and $a_{22} = 1$. Further, we use formulae (6)–(9) to compute the integers $b_{s,i,j,r}$. Then, using Theorem 6, we find the polynomials U_1 and U_2. □

To find another asymptotic formula for C_n, we use the identity (see Dusart [4, p. 50] or Hassani [5, p. 3])

$$C_n = \int_2^{p_n} \pi(x)\,dx, \tag{12}$$

which allows us to estimate C_n by using explicit bounds for $\pi(x)$. Further, we use the following integration rules (see Lemma 8), where the *logarithmic integral* $\mathrm{li}(x)$ is defined for every real $x \ge 2$ by

$$\mathrm{li}(x) = \int_0^x \frac{dt}{\log t} = \lim_{\varepsilon\to 0}\left\{\int_0^{1-\varepsilon}\frac{dt}{\log t} + \int_{1+\varepsilon}^{x}\frac{dt}{\log t}\right\} \approx \int_2^x \frac{dt}{\log t} + 1.04516\ldots.$$

Lemma 8. *Let $r, s \in \mathbb{R}$ with $s \ge r > 1$.*

(i) $\displaystyle\int_r^s \frac{x\,dx}{\log x} = \mathrm{li}(s^2) - \mathrm{li}(r^2)$.

(ii) $\displaystyle\int_r^s \frac{x\,dx}{\log^2 x} = 2\,\mathrm{li}(s^2) - 2\,\mathrm{li}(r^2) - \frac{s^2}{\log s} + \frac{r^2}{\log r}$.

(iii) *If* $n \in \mathbb{N}$, *then*

$$\int_r^s \frac{x\,dx}{\log^{n+1} x} = \frac{r^2}{n \log^n r} - \frac{s^2}{n \log^n s} + \frac{2}{n}\int_r^s \frac{x}{\log^n x}\,dx.$$

(iv) *For every* $m \in \mathbb{N}$ *with* $m \geq 2$ *we have*

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

Proof. The rules (i) and (ii) are from Dusart [4, Lemma 1.6]. Now, (iii) follows by integration by parts and (iv) can be shown by induction on m. □

The next proposition plays an important role for the proof of the second asymptotic formula (Theorem 2, see Introduction) for C_n.

Proposition 9. *Let* $m \in \mathbb{N}$ *with* $m \geq 2$. *Let* $a_2, \ldots, a_m \in \mathbb{R}$ *and let* $r, s \in \mathbb{R}$ *with* $s \geq r > 1$. *Then*

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

where

$$t_{i,j} = (j-1)!\sum_{l=j}^{i} \frac{2^{l-j}a_{l+1}}{l!}. \tag{13}$$

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

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

By Lemma 8(iii), we get

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

Now we use Lemma 8(iv) and the equality $t_{m-1,1} + 2^{m-1}a_{m+1}/m! = t_{m,1}$ to obtain

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

Since we have

$$\frac{2^{m-k}a_{m+1}(k-1)!}{m!} + t_{m-1,k} = t_{m,k}$$

and $t_{m,m} = a_{m+1}(m-1)!/(m!)$, the proposition is proved. □

Now we are able to prove Theorem 2.

Theorem 10. *For each $m \in \mathbb{N}$ we have*

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

Proof. A well-known asymptotic formula for the prime counting function $\pi(x)$ is given by

$$\pi(x) = \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x} + \frac{6x}{\log^4 x} + \ldots + \frac{(m-1)!x}{\log^m x} + O\left(\frac{x}{\log^{m+1} x}\right). \tag{14}$$

Using (14) and (12), 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

$$C_n = \sum_{k=1}^{m}(k-1)!\int_2^{p_n}\frac{x\,dx}{\log^k x} + O\left(\frac{p_n^2}{\log^m p_n}\right).$$

We now apply Proposition 9 to get

$$C_n = \int_2^{p_n}\frac{x\,dx}{\log x} + (2^{m-1}-1)\int_2^{p_n}\frac{x\,dx}{\log^2 x} - \sum_{k=2}^{m-1}\left(\frac{(k-1)!(2^{m-k}-1)p_n^2}{\log^k p_n}\right) + O\left(\frac{p_n^2}{\log^m p_n}\right).$$

It follows from Lemma 8(i) and Lemma 8(ii) that

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

Now we use the well-known asymptotic formula

$$\mathrm{li}(x) = \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x} + \frac{6x}{\log^4 x} + \ldots + \frac{(m-1)!x}{\log^m x} + O\left(\frac{x}{\log^{m+1} x}\right) \tag{15}$$

to obtain

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

and the theorem is proved. □

Using (14), we get the following corollary.

Corollary 11. *For each $m \in \mathbb{N}$ we have*

$$\sum_{k \le n} p_k = \pi(p_n^2) + O\left(\frac{p_n^2}{\log^m p_n}\right).$$

Proof. From Theorem 10 and the definition of C_n it follows that

$$\sum_{k \le n} p_k = np_n - \sum_{k=1}^{m-1} \frac{(k-1)!\, p_n^2}{\log^k p_n} + \sum_{k=1}^{m-1} \frac{(k-1)!\, p_n^2}{2^k \log^k p_n} + O\left(\frac{p_n^2}{\log^m p_n}\right).$$

Since $n = \pi(p_n)$, we obtain

$$\sum_{k \le n} p_k = \pi(p_n)p_n - \sum_{k=1}^{m-1} \frac{(k-1)!\, p_n^2}{\log^k p_n} + \sum_{k=1}^{m-1} \frac{(k-1)!\, p_n^2}{2^k \log^k p_n} + O\left(\frac{p_n^2}{\log^m p_n}\right).$$

Using (14), we get the asymptotic formula

$$\sum_{k \le n} p_k = \sum_{k=1}^{m-1} \frac{(k-1)!\, p_n^2}{2^k \log^k p_n} + O\left(\frac{p_n^2}{\log^m p_n}\right) = \pi(p_n^2) + O\left(\frac{p_n^2}{\log^m p_n}\right)$$

and the corollary is proved. □

Using (14), (15) and Corollary 11, we obtain the following result concerning the sum of the first n prime numbers.

Corollary 12. *For each $m \in \mathbb{N}$ we have*

$$\sum_{k \le n} p_k = \mathrm{li}(p_n^2) + O\left(\frac{p_n^2}{\log^m p_n}\right).$$

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

$$\pi(x) \geq \frac{x}{\log x} + \sum_{k=2}^{m} \frac{a_k x}{\log^k x} \tag{16}$$

for every $x \geq x_0$ and

$$\mathrm{li}(x) \geq \sum_{j=1}^{m-1} \frac{(j-1)!x}{\log^j x} \tag{17}$$

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

Theorem 13. *If $n \geq \max\{\pi(x_0)+1, \pi(\sqrt{y_0})+1\}$, then*

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

where $t_{i,j}$ is defined as in (13) and d_0 is given by

$$d_0 = d_0(m, a_2, \ldots, a_m, x_0) = \int_2^{x_0} \pi(x)\,dx - (1+2t_{m-1,1})\,\mathrm{li}(x_0^2) + \sum_{k=1}^{m-1} t_{m-1,k} \frac{x_0^2}{\log^k x_0}.$$

Proof. Since $p_n \geq x_0$, we use (12) and (16) to obtain

$$C_n \geq \int_2^{x_0} \pi(x)\,dx + \int_{x_0}^{p_n} \frac{x\,dx}{\log x} + \sum_{k=2}^{m} a_k \int_{x_0}^{p_n} \frac{x\,dx}{\log^k x}.$$

Now we apply Lemma 8(i) and Proposition 9 to get

$$C_n \geq \int_2^{x_0} \pi(x)\,dx - \mathrm{li}(x_0^2) + \mathrm{li}(p_n^2) + t_{m-1,1} \int_{x_0}^{p_n} \frac{x\,dx}{\log^2 x} - \sum_{k=2}^{m-1} t_{m-1,k} \left(\frac{p_n^2}{\log^k p_n} - \frac{x_0^2}{\log^k x_0} \right).$$

Using Lemma 8(ii), we obtain

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

Since $p_n^2 \geq y_0$, we use (17) to conclude

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

and it remains to apply the definition of t_{ij}. $\square$

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

$$\pi(x) \leq \frac{x}{\log x} + \sum_{k=2}^{m} \frac{a_k x}{\log^k x} \tag{18}$$

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

$$\mathrm{li}(x) \leq \sum_{j=1}^{m-2} \frac{(j-1)!x}{\log^j x} + \frac{\lambda x}{\log^{m-1} x} \tag{19}$$

for every $x \geq y_1$. Setting

$$d_1 = d_1(m, a_2, \ldots, a_m, x_1) = \int_2^{x_1} \pi(x)\,dx - (1 + 2t_{m-1,1})\,\mathrm{li}(x_1^2) + \sum_{k=1}^{m-1} t_{m-1,k} \frac{x_1^2}{\log^k x_1},$$

where $t_{m-1,k}$ is defined by (13), we obtain the following

Theorem 14. *If* $n \geq \max\{\pi(x_1) + 1, \pi(\sqrt{y_1}) + 1\}$, *then*

$$C_n \leq d_1 + \sum_{k=1}^{m-2} \left(\frac{(k-1)!}{2^k}(1 + 2t_{k-1,1}) \right) \frac{p_n^2}{\log^k p_n} + \left(\frac{(1 + 2t_{m-1,1})\lambda}{2^{m-1}} - \frac{a_m}{m-1} \right) \frac{p_n^2}{\log^{m-1} p_n}.$$

Proof. Since $p_n \geq x_1$, we use (12) and (18) to get

$$C_n \leq \int_2^{x_1} \pi(x)\,dx + \int_{x_1}^{p_n} \frac{x\,dx}{\log x} + \sum_{k=2}^{m} a_k \int_{x_1}^{p_n} \frac{x\,dx}{\log^k x}.$$

We apply Lemma 8(i) and Proposition 9 to obtain

$$C_n \leq \int_2^{x_1} \pi(x)\,dx - \mathrm{li}(x_1^2) + \mathrm{li}(p_n^2) + t_{m-1,1} \int_{x_1}^{p_n} \frac{x\,dx}{\log^2 x} - \sum_{k=2}^{m-1} t_{m-1,k} \left(\frac{p_n^2}{\log^k p_n} - \frac{x_1^2}{\log^k x_1} \right).$$

Using Lemma 8(ii), we get

$$C_n \leq d_1 + (1 + 2t_{m-1,1})\,\mathrm{li}(p_n^2) - \sum_{k=1}^{m-1} t_{m-1,k} \frac{p_n^2}{\log^k p_n}.$$

Now we use the inequality (19) to obtain

$$\begin{aligned} C_n \leq d_1 &+ \sum_{k=1}^{m-2} \left(\frac{(k-1)!}{2^k} + \frac{t_{m-1,1}(k-1)!}{2^{k-1}} - t_{m-1,k} \right) \frac{p_n^2}{\log^k p_n} \\ &+ \left(\frac{(1 + 2t_{m-1,1})\lambda}{2^{m-1}} - t_{m-1,m-1} \right) \frac{p_n^2}{\log^{m-1} p_n} \end{aligned}$$

and it remains to apply the definition of t_{ij}. □

5 Numerical results

5.1 An explicit lower bound for C_n

The goal of this subsection is to improve the inequality (2) in view of (4). In order to do this, we first give two lemmata concerning explicit estimates for $\mathrm{li}(x)$ and $\pi(x)$, respectively.

Lemma 15. *If $x \geq 4171$, then*

$$\mathrm{li}(x) \geq \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x} + \frac{6x}{\log^4 x} + \frac{24x}{\log^5 x} + \frac{120x}{\log^6 x} + \frac{720x}{\log^7 x} + \frac{5040x}{\log^8 x}.$$

Proof. We denote the right hand side of the above inequality by $\alpha(x)$ and let $f(x) = \mathrm{li}(x) - \alpha(x)$. Then, $f(4171) \geq 0.00019$ and $f'(x) = 40320/\log^9 x$, and the lemma is proved. □

Lemma 16. *If $x \geq 10^{16}$, then*

$$\mathrm{li}(x) \leq \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x} + \frac{6x}{\log^4 x} + \frac{24x}{\log^5 x} + \frac{120x}{\log^6 x} + \frac{900x}{\log^7 x}.$$

Proof. Similarly to the proof of Lemma 15. □

Lemma 17. *If $x \geq 1332450001$, then*

$$\pi(x) > \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x} + \frac{5.65x}{\log^4 x} + \frac{23.65x}{\log^5 x} + \frac{118.25x}{\log^6 x} + \frac{709.5x}{\log^7 x} + \frac{4966.5x}{\log^8 x}.$$

Proof. See [1, Thm. 1.2]. □

Setting

$$\Theta(n) = \frac{43.6p_n^2}{8\log^4 p_n} + \frac{90.9p_n^2}{4\log^5 p_n} + \frac{927.5p_n^2}{8\log^6 p_n} + \frac{5620.5p_n^2}{8\log^7 p_n} + \frac{79075.5p_n^2}{16\log^8 p_n}.$$

we get the following improvement of (2).

Proposition 18. *If $n \geq 52703656$, then*

$$C_n \geq \frac{p_n^2}{2\log p_n} + \frac{3p_n^2}{4\log^2 p_n} + \frac{7p_n^2}{4\log^3 p_n} + \Theta(n).$$

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 Lemma 17, we obtain the inequality (16) for every $x \geq x_0$ and (17) holds for every $x \geq y_0$ by Lemma 15. Substituting these values in Theorem 13, 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

$$d_0 = \int_2^{x_0} \pi(x)\,dx - \frac{753.1}{3}\,\mathrm{li}(x_0^2) + \frac{375.05x_0^2}{3\log x_0} + \frac{186.025x_0^2}{3\log^2 x_0} + \frac{183.025x_0^2}{3\log^3 x_0} + \frac{88.6875x_0^2}{\log^4 x_0} + \frac{165.55x_0^2}{\log^5 x_0} + \frac{354.75x_0^2}{\log^6 x_0} + \frac{709.5x_0^2}{\log^7 x_0}.$$

Since $x_0^2 \geq 10^{16}$, it follows from Lemma 16 that

$$d_0 \geq \int_2^{x_0} \pi(x)\,dx - \frac{x_0^2}{2\log x_0} - \frac{3x_0^2}{4\log^2 x_0} - \frac{7x_0^2}{4\log^3 x_0} - \frac{5.45x_0^2}{\log^4 x_0} - \frac{22.725x_0^2}{\log^5 x_0} - \frac{115.9375x_0^2}{\log^6 x_0} - \frac{1055.578125x_0^2}{\log^7 x_0}.$$

Using $\log x_0 \geq 21.01027$, we get

$$\begin{aligned} d_0 &\geq \int_2^{x_0} \pi(x)\,dx - 4.22512933 \cdot 10^{16} - 0.30164729 \cdot 10^{16} - 0.03349997 \cdot 10^{16} \\ &\quad - 0.0049656 \cdot 10^{16} - 0.00098548 \cdot 10^{16} - 0.0002393 \cdot 10^{16} - 0.0001037 \cdot 10^{16} \\ &= \int_2^{x_0} \pi(x)\,dx - 4.56657067 \cdot 10^{16}. \end{aligned} \tag{20}$$

Since $x_0 = p_{66773604}$, we use (12) a computer to obtain

$$\int_2^{x_0} \pi(x)\,dx = C_{66773604} = 45665745738169817.$$

Hence, by (20), 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. □

5.2 An explicit upper bound for C_n

We begin with the following two lemmata.

Lemma 19. *If $x \geq 10^{18}$, then*

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

Proof. Similarly to the proof of Lemma 15. □

Lemma 20. *If $x > 1$, then*

$$\pi(x) < \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x} + \frac{6.35x}{\log^4 x} + \frac{24.35x}{\log^5 x} + \frac{121.75x}{\log^6 x} + \frac{730.5x}{\log^7 x} + \frac{6801.4x}{\log^8 x}.$$

Proof. See [1, Thm. 1.1]. □

Using these upper bounds, we obtain the following explicit upper bound for C_n, where

$$\Omega(n) = \frac{46.4p_n^2}{8\log^4 p_n} + \frac{95.1p_n^2}{4\log^5 p_n} + \frac{962.5p_n^2}{8\log^6 p_n} + \frac{5809.5p_n^2}{8\log^7 p_n} + \frac{118848p_n^2}{16\log^8 p_n}.$$

Proposition 21. *For every $n \in \mathbb{N}$,*

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

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 Lemma 20, we get that the inequality (18) holds for every $x \geq x_1$ and by Lemma 19, that (19) holds for all $y \geq y_1$. By substituting these values in Theorem 14, we get

$$C_n \leq d_1 + \frac{p_n^2}{2\log p_n} + \frac{3p_n^2}{4\log^2 p_n} + \frac{7p_n^2}{4\log^3 p_n} + \Omega(n) - \frac{0.875p_n^2}{16\log^8 p_n} \tag{21}$$

for every $n \geq 50847535$, where $d_1 = d_1(9, 1, 2, 6.35, 24.35, 121.75, 730.5, 6801.4, 0, x_1)$ is given by

$$d_1 = \int_2^{x_1} \pi(x)\,dx - \frac{950777}{3150}\,\mathrm{li}(x_0^2) + \frac{947627x_0^2}{6300\log x_0} + \frac{941327x_0^2}{12600\log^2 x_0} + \frac{928727x_0^2}{12600\log^3 x_0} + \frac{902057x_0^2}{8400\log^4 x_0} + \frac{425461x_0^2}{2100\log^5 x_0} + \frac{187163x_0^2}{420\log^6 x_0} + \frac{34007x_0^2}{35\log^7 x_0}.$$

Since $\mathrm{li}(x_1^2) \geq 34.59$ and $\log x_1 \geq 2.39$, we obtain $d_1 \leq 450$. We define

$$f(x) = \frac{0.875x^2}{16\log^8 x} - 450.$$

Since $f(6 \cdot 10^6) \geq 109$ 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 (21) to obtain the desired inequality for every $n \geq 50847535$. For every $1 \leq n \leq 50847534$ a computer makes the rest of work. □

References

[1] C. Axler, New bounds for the prime counting function $\pi(x)$, preprint, 2015. Available at http://arxiv.org/abs/1409.1780.

[2] C. Axler, On the sum of the first n prime numbers, preprint, 2014. Available at http://arxiv.org/abs/1409.1777.

[3] M. Cipolla, La determinazione assintotica dell' n^{imo} numero primo, *Rend. Accad. Sci. Fis. Mat. Napoli* **8** (1902), 132–166.

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

[5] M. Hassani, A remark on the Mandl's inequality, preprint, 2006. Available at `http://arxiv.org/abs/math/0606765`.

[6] N. J. A. Sloane, The On-Line Encyclopedia of Integer Sequences, `http://oeis.org`.

[7] J. B. Rosser and L. Schoenfeld, Sharper bounds for the Chebyshev functions $\theta(x)$ and $\psi(x)$, *Math. Comp.* **29** (1975), 243–269.

2010 *Mathematics Subject Classification*: Primary 11A41; Secondary 11B83, 11N05.
Keywords: asymptotic formula, Mandl's inequality, prime number.

(Concerned with sequences A000040, A007504, A124478, and A152535.)

Received March 19 2015; revised versions received March 23 2015; June 19 2015; July 10 2015; July 13 2015. Published in *Journal of Integer Sequences*, July 16 2015.

Return to Journal of Integer Sequences home page.