IMPROVING THE ESTIMATES FOR A SEQUENCE INVOLVING PRIME NUMBERS

CHRISTIAN AXLER

Abstract. Based on new explicit estimates for the prime counting function, we improve the currently known estimates for the particular sequence $C_n = np_n - \sum_{k \leq n} p_k$, $n \geq 1$, involving the prime numbers.

1. Introduction

Let p_n denotes the nth prime number. In this paper, we establish new explicit estimates for the sequence $(C_n)_{n \geq 1}$ with

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

(see [6]). In [1, Theorem 10], the present author used the identity

$$C_n = \int_2^{p_n} \pi(x)\, dx, \tag{1}$$

where $\pi(x)$ denotes the number of primes not exceeding x, to derive that the asymptotic formula

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

holds for each positive integer m. By setting $m = 9$ in (2), we get

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

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 the direction of (3), the present author [1, Theorem 3 and Theorem 4] showed that

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

for every $n \geq 52703656$, where

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

and that the upper bound

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

holds for every positive integer n, where

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

Using new explicit estimates for the prime counting function $\pi(x)$, which are found in [2, Proposition 3.6 and Proposition 3.12], we improve the inequalities (4) and (5) by showing the following both results.

Date: June 14, 2017.
2010 *Mathematics Subject Classification.* Primary 11N05; Secondary 11A41.
Key words and phrases. prime counting function, prime numbers, sum of primes.

Theorem 1.1. *For every positive integer $n \geq 440200309$, we have*

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

where

$$L(n) = \frac{44.4p_n^2}{8\log^4 p_n} + \frac{92.1p_n^2}{4\log^5 p_n} + \frac{937.5p_n^2}{8\log^6 p_n} + \frac{5674.5p_n^2}{8\log^7 p_n} + \frac{79789.5p_n^2}{16\log^8 p_n}.$$

Theorem 1.2. *For every positive integer n, we have*

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

where

$$U(n) = \frac{45.6p_n^2}{8\log^4 p_n} + \frac{93.9p_n^2}{4\log^5 p_n} + \frac{952.5p_n^2}{8\log^6 p_n} + \frac{5755.5p_n^2}{8\log^7 p_n} + \frac{116371p_n^2}{16\log^8 p_n}.$$

2. Preliminaries

In 1793, Gauß [4] stated a conjecture concerning an asymptotic magnitude of $\pi(x)$, namely

$$\pi(x) \sim \mathrm{li}(x) \qquad (x \to \infty), \tag{6}$$

where the *logarithmic integral* $\mathrm{li}(x)$ defined for every real $x \geq 0$ as

$$\mathrm{li}(x) = \int_0^x \frac{dt}{\log t} = \lim_{\varepsilon \to 0} \left\{ \int_0^{1-\varepsilon} \frac{dt}{\log t} + \int_{1+\varepsilon}^x \frac{dt}{\log t} \right\} \approx \int_2^x \frac{dt}{\log t} + 1.04516\ldots. \tag{7}$$

Using the method of integration of parts, (7) implies that

$$\mathrm{li}(x) = \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x} + \frac{6x}{\log^4 x} + \frac{24x}{\log^5 x} + \ldots + \frac{(m-1)!x}{\log^m x} + O\left(\frac{x}{\log^{m+1} x}\right) \tag{8}$$

for every positive integer m. The asymptotic formula (6) was proved independently by Hadamard [5] and by de la Vallée-Poussin [7] in 1896, and is known as the *Prime Number Theorem*. By proving the existence of a zero-free region for the Riemann zeta-function $\zeta(s)$ to the left of the line $\mathrm{Re}(s) = 1$, de la Vallée-Poussin [8] was able to estimate the error term in the Prime Number Theorem by

$$\pi(x) = \mathrm{li}(x) + O(x\exp(-a\sqrt{\log x})),$$

where a is a positive absolute constant. Together with (8), we obtain that the asymptotic formula

$$\pi(x) = \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x} + \frac{6x}{\log^4 x} + \frac{24x}{\log^5 x} + \ldots + \frac{(m-1)!x}{\log^m x} + O\left(\frac{x}{\log^{m+1} x}\right). \tag{9}$$

holds for every positive integer m.

3. A proof of Theorem 1.1

Now, we use some recent obtained lower bound for the prime counting function $\pi(x)$ to give a proof of Theorem 1.1.

Proof of Theorem 1.1. First, let m be a positive integer with $m \geq 2$, and let $a_2, \ldots, a_m$, x_0, and y_0 be real numbers, so that

$$\pi(x) \geq \frac{x}{\log x} + \sum_{k=2}^{m} \frac{a_k x}{\log^k x} \tag{10}$$

for every $x \geq x_0$ and

$$\mathrm{li}(x) \geq \sum_{j=1}^{m-1} \frac{(j-1)!x}{\log^j x} \tag{11}$$

for every $x \geq y_0$. The asymptotic formulae (9) and (8) guarantee the existence of such parameters. In [1, Theorem 13], the present author showed that

$$C_n \geq d_0 + \sum_{k=1}^{m-1} \left(\frac{(k-1)!}{2^k}(1 + 2t_{k-1,1}) \right) \frac{p_n^2}{\log^k p_n} \tag{12}$$

for every $n \geq \max\{\pi(x_0) + 1, \pi(\sqrt{y_0}) + 1\}$, where $t_{i,j}$ is defined by

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

and d_0 is given by

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

Now, we choose $m = 9$, $a_2 = 1$, $a_3 = 2$, $a_4 = 5.85$, $a_5 = 23.85$, $a_6 = 119.25$, $a_7 = 715.5$, $a_8 = 5008.5$, $a_9 = 0$, $x_0 = 19027490297$ and $y_0 = 4171$. By [2, Proposition 3.12], we obtain that the inequality (10) holds for every $x \geq x_0$ and (11) holds for every $x \geq y_0$ by [1, Lemma 15]. Substituting these values in (12), we get

$$C_n \geq d_0 + \frac{p_n^2}{2 \log p_n} + \frac{3p_n^2}{4 \log^2 p_n} + \frac{7p_n^2}{4 \log^3 p_n} + L(n)$$

for every $n \geq 841160647 = \pi(x_0)$, where $d_0 = d_0(9, 1, 2, 5.85, 23.85, 119.25, 715.5, 5008.5, 0, x_0)$ is given by

$$\begin{aligned} d_0 = \int_2^{x_0} \pi(x)\, dx - 253.3 \operatorname{li}(x_0^2) + \frac{126.15x_0^2}{\log x_0} + \frac{62.575x_0^2}{\log^2 x_0} + \frac{61.575x_0^2}{\log^3 x_0} \\ + \frac{89.4375x_0^2}{\log^4 x_0} + \frac{165.95x_0^2}{\log^5 x_0} + \frac{357.75x_0^2}{\log^6 x_0} + \frac{715.5x_0^2}{\log^7 x_0}. \end{aligned} \tag{14}$$

The present author [1, Lemma 16] found that

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

for every $x \geq 10^{16}$. Applying this inequality to (14), we get

$$\begin{aligned} d_0 \geq \int_2^{x_0} \pi(x)\, dx - \frac{x_0^2}{2 \log x_0} - \frac{3x_0^2}{4 \log^2 x_0} - \frac{7x_0^2}{4 \log^3 x_0} - \frac{5.55x_0^2}{\log^4 x_0} - \frac{23.025x_0^2}{\log^5 x_0} \\ - \frac{117.1875x_0^2}{\log^6 x_0} - \frac{1065.515625x_0^2}{\log^7 x_0}. \end{aligned}$$

Computing the right-hand side of the last inequality, we get

$$d_0 \geq \int_2^{x_0} \pi(x)\, dx - 8.188366 \cdot 10^{18}. \tag{15}$$

Since $x_0 = p_{841160647}$, we use (1) and a computer to obtain

$$\int_2^{x_0} \pi(x)\, dx = C_{841160647} = 8188378036394419009.$$

Hence, by (15), we get $d_0 \geq 1.12 \cdot 10^{13} > 0$. So we obtain the desired inequality for every $n \geq 841160647$. For every $440200309 \leq n \leq 841160646$ we check the inequality with a computer. □

4. A proof of Theorem 1.2

Next, we use a recent result concerning an upper bound for the prime counting function $\pi(x)$ to establish the required inequality stated in Theorem 1.2.

Proof of Theorem 1.2. Let m be a positive integer with $m \geq 2$, let $a_2, \ldots, a_m, x_1$ be real numbers so that

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

for every $x \geq x_1$ and let λ, y_1 be real numbers so that

$$\operatorname{li}(x) \leq \sum_{j=1}^{m-2} \frac{(j-1)!x}{\log^j x} + \frac{\lambda x}{\log^{m-1} x} \tag{17}$$

for every $x \geq y_1$. Again, the asymptotic formulae (9) and (8) guarantee the existence of such parameters. The present author [1, Theorem 14] found that the inequality

$$C_n \leq d_1 + \sum_{k=1}^{m-2} \left(\frac{(k-1)!}{2^k}(1 + 2t_{k-1,1}) \right) \frac{p_n^2}{\log^k p_n} + \left(\frac{(1 + 2t_{m-1,1})\lambda}{2^{m-1}} - \frac{a_m}{m-1} \right) \frac{p_n^2}{\log^{m-1} p_n} \tag{18}$$

holds for every $n \geq \max\{\pi(x_1) + 1, \pi(\sqrt{y_1}) + 1\}$, where $t_{i,j}$ is defined by (13), and

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

Next, we choose $m = 9$, $a_2 = 1$, $a_3 = 2$, $a_4 = 6.15$, $a_5 = 24.15$, $a_6 = 120.75$, $a_7 = 724.5$,$a_8 = 6601$, $a_9 = 0$, $\lambda = 6300$, $x_1 = 13$ and $y_1 = 10^{18}$. By [2, Proposition 3.6], we get that the inequality (16) holds for every $x \geq x_1$ and by [1, Lemma 19], that (17) holds for every $y \geq y_1$. By substituting these values (18), we get

$$C_n \leq d_1 + \frac{p_n^2}{2 \log p_n} + \frac{3p_n^2}{4 \log^2 p_n} + \frac{7p_n^2}{4 \log^3 p_n} + U(n) - \frac{0.375p_n^2}{16 \log^8 p_n} \tag{19}$$

for every $n \geq 50847535$, where $d_1 = d_1(9, 1, 2, 6.15, 24.15, 120.75, 724.5, 6601, 0, x_1)$ is given by

$$d_1 = \int_2^{x_1} \pi(x)\, dx - \frac{26599}{90} \operatorname{li}(x_1^2) + \frac{26509x_1^2}{180 \log x_1} + \frac{26329x_1^2}{360 \log^2 x_1} + \frac{25969x_1^2}{360 \log^3 x_1} + \frac{25231x_1^2}{240 \log^4 x_1} + \frac{11891x_1^2}{60 \log^5 x_1} + \frac{5221x_1^2}{12 \log^6 x_1} + \frac{943x_1^2}{\log^7 x_1}.$$

A computation shows that $d_1 \leq 453$. We define

$$f(x) = \frac{0.375x^2}{16 \log^8 x} - 453.$$

Since $f(9187322) > 0$ and $f'(x) \geq 0$ for every $x \geq e^4$, we get $f(p_n) \geq 0$ for every $n \geq \pi(9187322) + 1 = 614124$. Now we can use (19) to obtain the desired inequality for every positive integer $n \geq 50847535$. Finally, we check the remaining cases with a computer. □

References

[1] C. Axler, On a sequence involving prime numbers, *J. Integer Seq.* **18** (2015), no. 7, Article 15.7.6, 13 pp.
[2] C. Axler, New estimates for some prime functions, preprint, 2017. Available at `arxiv.org/1703.08032`.
[3] P. Dusart, *Autour de la fonction qui compte le nombre de nombres premiers*, Dissertation, Université de Limoges, 1998.
[4] C. F. Gauß, *Werke*, 2 ed., Königlichen Gesellschaft der Wissenschaften, Göttingen, 1876.
[5] J. Hadamard, Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques, *Bull. Soc. Math. France* **24** (1896), 199–220.
[6] N. J. A. Sloane, Sequence A152535, The Online Encyclopedia of Integer Sequences, `oeis.org/A152535`
[7] C.-J. de la Vallée Poussin, Recherches analytiques la théorie des nombres premiers, *Ann. Soc. scient. Bruxelles* **20** (1896), 183–256.
[8] C.-J. de la Vallée Poussin, Sur la fonction $\zeta(s)$ de Riemann et le nombre des nombres premiers inférieurs à une limite donnée, *Mem. Couronnés de l'Acad. Roy. Sci. Bruxelles* **59** (1899), 1–74.

Institute of Mathematics, Heinrich-Heine University Düsseldorf, 40225 Düsseldorf, Germany
E-mail address: `christian.axler@hhu.de`