IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 21, Issue 3, Ser. 2 (May. – June. 2025), PP 39-43
www.iosrjournals.org

Refinements of Inequalities in Prime Number Sequences

Qianxi Zhang[1], Xinhua Xiong[2]
[1](*College of Mathematics and Physics, China Three Gorges University, Yichang, China*)
[2](*College of Mathematics and Physics, China Three Gorges University, Yichang, China*)

Abstract: *This study conducts an in-depth analysis and improvement of the asymptotic properties of the prime sequence $C_n = np_n - \sum_{k\le n} p_k$ (for $n \ge 1$) and the upper bounds of the prime summation function $S(x)$, based on the upper and lower bound estimates of the prime counting function $\pi(x)$. By developing novel estimation methods, we obtain more precise asymptotic estimates than existing results in the literature. Specifically, this work not only optimizes the upper and lower bound estimates of Cn but also significantly improves the upper bound estimation of $S(x)$. These refinements deepen our understanding of the structural characteristics of prime sequences and hold potential applications in computational number theory and related fields.*
Background*: The study of prime numbers and their asymptotic behavior has been a central topic in number theory since the pioneering work of Gauss and Legendre on the Prime Number Theorem. The prime counting function $\pi(x)$, which enumerates primes not exceeding x, provides fundamental insights into the distribution of primes. While the Riemann Hypothesis offers the most precise conjectural bounds for $\pi(x)$, practical applications often rely on computationally verifiable estimates. The summation of primes $S(x)$ and related sequences like $C_n = np_n - \sum_{k\le n} p_k$ naturally arise in various contexts, including prime gap analysis and verification algorithms for prime certificates. Previous work by Rosser and Schoenfeld established rigorous bounds for $\pi(x)$, and subsequent refinements by Dusart and others have improved these estimates.*
Key Word*: Prime counting function; Prime sequence; Sum of primes; Prime numbers.*

Date of Submission: 09-06-2025 Date of Acceptance: 20-06-2025

I. Introduction

In 1998, Dusart proved[5] that for every integer $n \ge 109$,

$$C_n = np_n - \sum_{k\le n} p_k \ge c + \frac{p_n^2}{2\log p_n}.$$

Where $c \approx -47.1$. In 2015[1], Axler established that for every positive integer m,

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

By setting $m = 9$ in equation (1), we obtain

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

where

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

Building on this result, Axler[2] derived in 2019 the following estimates for the prime sequence $(C_n)_{n\ge 1}$. For every integer $n \ge 440200309$,

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

where

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

And for every integer n,

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

where

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

Regarding the sum of primes $S(x)$ (denoting the sum of all primes not exceeding x), Szalay[9] proved

$$S(x) = \mathrm{li}(x^2) + O(x^2 e^{-a\sqrt{\log x}}).$$

Building upon his previous results, Axler[3] established that

$$S(x) = \frac{x^2}{2\log x} + \frac{x^2}{4\log^2 x} + \frac{x^2}{4\log^3 x} + \frac{3x^2}{8\log^4 x} + O\left(\frac{x^2}{\log^5 x}\right).$$

Massias and Robin[8] proved that for all $x \ge 24281$,

$$S(x) \le \frac{x^2}{2\log x} + \frac{3x^2}{10\log^2 x}.$$

Subsequently, Axler improved upon these results by establishing that for all $x \ge 110118925$,

$$S(x) < \frac{x^2}{2\log x} + \frac{x^2}{4\log^2 x} + \frac{x^2}{4\log^3 x} + \frac{5.3x^2}{8\log^4 x}.$$

II. Preliminaries

This section introduces the fundamental concepts and lemmas necessary for the subsequent developments in this paper.

Definition 1:

$$\pi(x) = \sum_{p \le x} 1,$$

here, $\pi(x)$ represents the prime-counting function, which gives the number of prime numbers less than or equal to x.

Lemma 1[5]:

$$C_n = \int_2^{p_n} \pi(x)dx, \tag{2}$$

Definition 2: Gauss[7] observed that the logarithmic integral $\mathrm{li}(x) = \int_0^x \frac{dt}{\log t}$ provides an exceptionally accurate approximation to $\pi(x)$,

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

Lemma 2[1]: For all x satisfying $x \ge 4171$, we have

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

Lemma 3[1]: If $x \ge 10^{16}$, then

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

Lemma 4[1]: If $x \ge 10^{18}$, then

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

Lemma 5[1]: For $n \ge max\{\pi(x_0) + 1, \pi(\sqrt{y_0}) + 1\}$, we have

$$C_n \ge d_0 + \sum_{k=1}^{m-1}\left(\frac{(k-1)!}{2^k}(1 + 2t_{k-1,1})\right)\frac{p_n^2}{\log^k p_n}, \tag{4}$$

where

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

$$d_0 = d_0(m, a_2, \cdots, a_m, x_0) = \int_2^{x_0} \pi(x)dx - (1 + 2t_{m-1,1})\mathrm{li}(x_0^2) + \sum_{k=1}^{m-1} t_{m-1,k}\frac{x_0^2}{\log^k x_0}. \tag{5}$$

Lemma 6[1]: For $n \geq max\{\pi(x_1)+1, \pi(\sqrt{y_1})+1\}$, we have

$$C_n \leq d_1 + \sum_{k=1}^{m-2}\left(\frac{(k-1)!}{2^k}(1+2t_{k-1,1})\right)\frac{p_n^2}{\log^k p_n} + \left(\frac{(1+2t_{m-1,1})\lambda}{2^{m-1}} - \frac{a_m}{m-1}\right)\frac{p_n^2}{\log^{m-1} p_n}, \tag{6}$$

where

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

$$d_1 = d_1(m, a_2, \cdots, a_m, x_1) = \int_2^{x_1}\pi(x)dx - (1+2t_{m-1,1})\mathrm{li}(x_1^2) + \sum_{k=1}^{m-1} t_{m-1,k}\frac{x_1^2}{\log^k x_1}. \tag{7}$$

Lemma 7[4]: Provided that $x > 1$, it follows that

$$\pi(x) > \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x} + \frac{5.975666x}{\log^4 x} + \frac{23.975666x}{\log^5 x} + \frac{119.87833x}{\log^6 x} + \frac{719.26998x}{\log^7 x} + \frac{5034.88986x}{\log^8 x}.$$

Lemma 8[4]: Provided that $x \leq 1681111802141$, it follows that

$$\pi(x) < \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{2x}{\log^3 x} + \frac{6.024334x}{\log^4 x} + \frac{4.024334x}{\log^5 x} + \frac{120.12167x}{\log^6 x} + \frac{720.73002x}{\log^7 x} + \frac{6098x}{\log^8 x}.$$

III. Theorems and Proofs

We now present the main theoretical contributions of this paper, accompanied by their complete proofs.

Theorem 1: If $n \geq 62009690659$, then

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

where

$$\Delta(n) = \frac{44.902664p_n^2}{8\log^4 p_n} + \frac{92.853996p_n^2}{4\log^5 p_n} + \frac{943.7833p_n^2}{8\log^6 p_n} + \frac{5708.42982p_n^2}{8\log^7 p_n} + \frac{80238.12762p_n^2}{16\log^8 p_n}.$$

Proof: Throughout the proof, we set the parameters as follows: $m = 9, a_2 = 1, a_3 = 2, a_4 = 5.975666, a_5 = 23.975666, a_6 = 119.87833, a_7 = 719.26998, a_8 = 5034.88986, a_9 = 0, x_0 = 1681111802141$ and $y_0 = 4171$. According to Lemma 7, we obtain that when $x \geq x_0$,

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

Then, by Lemma 2, for all $x \geq y_0$,

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

Substituting the chosen parameters into (4), we derive that for all $n \geq 62009690659 = \pi(x_0) + 1$,

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

where

$$d_0 = d_0(9, 1, 2, 5.975666, 23.975666, 119.87833, 719.26998, 5034.88986, 0, x_0).$$

From (5), we obtain

$$d_0 = \int_2^{x_0}\pi(x)\mathrm{d}x - \frac{764.172644}{3}\mathrm{li}(x_0^2) + \frac{380.586322x_0^2}{3\log x_0} + \frac{188.793161x_0^2}{3\log^2 x_0} + \frac{185.793161x_0^2}{3\log^3 x_0} + \frac{89.9087475x_0^2}{\log^4 x_0} + \frac{167.829662x_0^2}{\log^5 x_0} + \frac{359.63499x_0^2}{\log^6 x_0} + \frac{719.26998x_0^2}{\log^7 x_0}.$$

From $x_0^2 \geq 10^{16}$ and Lemma 3, we derive

$$d_0 \geq \int_2^{x_0}\pi(x)dx - \frac{x_0^2}{2\log x_0} - \frac{3x_0^2}{4\log^2 x_0} - \frac{7x_0^2}{4\log^3 x_0} - \frac{5.612833x_0^2}{\log^4 x_0} - \frac{23.213499x_0^2}{\log^5 x_0} - \frac{117.9729125x_0^2}{\log^6 x_0} - \frac{1071.759654375x_0^2}{\log^7 x_0}. \tag{8}$$

Since $x_0 = p_{62009690658}$ and by combining equation (2), the computation via Wolfram Mathematica yields

$$\begin{aligned}&\int_2^{x_0}\pi(x)dx\\&=C_{62009690658}\\&=np_n-\sum_{k\le n}p_k\\&=62009690658\times 1681111802141-51121966756475752753601\\&=53123256055800559345177.\end{aligned}$$

Therefore, according to inequality (8), $d_0 \geq 5.12 \times 10^{22}$. Hence, for all $n \geq 62009690659$, the inequality in Theorem 1 holds.

Theorem 2: If $n \geq 50847535$, then

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

where

$$\nabla(n)=\frac{45.097336p_n^2}{8\log^4 p_n}+\frac{93.146004p_n^2}{4\log^5 p_n}+\frac{946.2167p_n^2}{8\log^6 p_n}+\frac{5721.57018p_n^2}{8\log^7 p_n}+\frac{111044p_n^2}{16\log^8 p_n}.$$

Proof: Similar to the proof of Theorem 4.1, we select $m = 9, a_2 = 1, a_3 = 2, a_4 = 6.024334, a_5 = 24.024334, a_6 = 120.12167, a_7 = 720.73002, a_8 = 6098, a_9 = 0, \lambda = 6300, x_1 = 17$ and $y_1 = 10^{18}$. By Lemma 8, we conclude that for all $x \geq x_1$,

$$\pi(x) \le \frac{x}{\log x}+\sum_{k=2}^{m}\frac{a_k x}{\log^k x}.$$

By Lemma 4, when $x \geq y_1$,

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

Substituting the selected values into equation (6), we obtain: for all $n \geq 50847535 = \pi(y_1) + 1$,

$$C_n \le d_1+\frac{p_n^2}{2\log p_n}+\frac{3p_n^2}{4\log^2 p_n}+\frac{7p_n^2}{4\log^3 p_n}+\nabla(n)-\frac{0.260925p_n^2}{16\log^8 p_n}, \tag{9}$$

where

$$d_1 = d_1(9,1,2,6.024334,24.024334,120.12167,720.73002,6098,0,x_1).$$

Derived from Equation (7)

$$d_1=\int_2^{x_1}\pi(x)dx-\frac{4441749563}{15750000}\mathrm{li}(x_1^2)+\frac{4425999563x_1^2}{31500000\log x_1}+\frac{4394499563x_1^2}{63000000\log^2 x_1}+\frac{4331499563x_1^2}{63000000\log^3 x_1}$$
$$+\frac{4204988549x_1^2}{42000000\log^4 x_1}+\frac{1976366521x_1^2}{10500000\log^5 x_1}+\frac{862055507x_1^2}{2100000\log^6 x_1}+\frac{6098x_1^2}{7\log^7 x_1}.$$

Given that , it follows that $d_1 \leq 202$. Define the function

$$f(x)=\frac{0.260925x^2}{16\log^8 x}-202.$$

Since $f'(x) \geq 0$ for $x \geq e^4$ and $f(50847535) > 0$, it follows that $f(p_n) \geq 0$ for all $n \geq \pi(50847535) + 1 = 3048956$. Together with (9), this completes the proof.

Theorem 3: For all $x \ge 11681111802143$,

$$S(x) < \frac{x^2}{2\log x}+\frac{x^2}{4\log^2 x}+\frac{x^2}{4\log^3 x}+\frac{3.7x^2}{8\log^4 x}.$$

Proof: The proof of Theorem 1.1 in reference[3] yields the inequality

$$S(x)=\pi(p_n)p_n-np_n+\sum_{k\le n}p_k\ ,$$

combining Lemma 8 with Theorem 1 yields

$$S(x) < \frac{p_n^2}{2\log p_n} + \frac{p_n^2}{4\log^2 p_n} + \frac{p_n^2}{4\log^3 p_n} + \frac{3.292008 p_n^2}{8\log^4 p_n} + \frac{3.24334 p_n^2}{4\log^5 p_n} + \frac{17.19006 p_n^2}{8\log^6 p_n} + \frac{57.41034 p_n^2}{8\log^7 p_n} + \frac{17329.87238 p_n^2}{16\log^8 p_n}. \quad (10)$$

We denote the right-hand side of inequality (10) as $h(p_n) = h(x)$, and the right-hand side of the inequality in Theorem 3 as $f(p_n) = f(x)$. Since $h(x)$ is an increasing function, for $x \geq p_{62009690659}$, we have $S(x) < h(x)$. Moreover, because $h(x) < f(x)$, the theorem is proved.

References

[1]. Axler C. On a sequence involving prime numbers. Journal of Integer Sequences.2015;18:Article 15.7.6.

[2]. Axler C. Improving the estimates for a sequence involving prime numbers. Notes on Number Theory and Discrete Mathematics.2019; 25(1):8-13

[3]. Axler C. On the sum of the first *n* prime numbers, Journal de Théorie des Nombres de Bordeaux. 2019;31:293–311.

[4]. Axler C. Effective estimates for some functions defined over primes. Integers: The Electronic Journal of Combinatorial Number Theory. 2024;24:Paper No.A34.

[5]. Dusart P. Autour de la fonction qui compte le nombre de nombres premiers. Ph.D. Thesis, Limoges, 1998.

[6]. Dusart P. The k-th prime is grater than $k(\ln k + \ln\ln k - 1)$ for $k \geq 2$. Mathematics of Computation.1999;68(225): 411-415.

[7]. Gauss C. F. Königliche Gesellschaft der Wissenschaften zu Göttingen.Werke, Vol. II.1863;pp.444–447.

[8]. Massias J. P. and Robin G. Bornes effectives pour certaines fonctions concernant les nom-bres premiers. Journal de Théorie des Nombres de Bordeaux.1986: 213-238.

[9]. Szalay M. On the maximal order in S_n and S_n^*. Acta Arithmetica. 1980;37: 321-331.