

# ON THE ASYMPTOTIC EXPANSION OF THE SUM OF THE FIRST $n$ PRIMES

*Nilotpal Kanti Sinha*

November 9, 2010

*Dedicated to Dr. Madan Mohan Singh, my first mathematics teacher..*

## Abstract

An asymptotic formula for the sum of the first  $n$  primes is derived. This result improves the previous results of P. Dusart. Using this asymptotic expansion, we prove the conjectures of R. Mandl and G. Robin on the upper and the lower bound of the sum of the first  $n$  primes respectively.

## 1 Introduction

Let  $p_n$  denote the  $n^{\text{th}}$  prime <sup>1 2</sup>. Robert Mandl conjectured that

$$\sum_{r \leq n} p_r < \frac{np_n}{2}.$$

This conjecture was proven by Rosser and Schoenfeld in [1] and is now referred to as Mandl's inequality. An alternate version of the proof was given by Dusart in [2]. In the same paper, Dusart also showed that

$$\sum_{r \leq n} p_r = \frac{n^2}{2}(\ln n + \ln \ln n - 3/2 + o(1)). \quad (1)$$

Currently, the best upper bound on Mandl's inequality is due to Hassani who showed that (See [3]) for for  $n \geq 10$ ,

$$\sum_{r \leq n} p_r < \frac{np_n}{2} - \frac{n^2}{14}. \quad (2)$$

With regards to the lower bound, G. Robin conjectured that

$$np_{[n/2]} < \sum_{r < n} p_r. \quad (3)$$

This conjecture was also proved by Dusart in [2]. However, neither 2 nor 3 give exact growth rate of  $\sum_{r \leq n} p_r$ . In this paper, we shall derive the asymptotic formula for  $\sum_{r \leq n} p_r$ . Both Hassani's improvement of Mandl's inequality and Robin's conjecture follow as corollaries of our asymptotic formula.

<sup>1</sup> 2000 *Mathematics Subject Classification*. 11A41.

<sup>2</sup> *Key words and phrases*. Primes, Inequalities.

## 2 Asymptotic expansion of $\sum_{r \leq n} p_r$

**Theorem 2.1.** (*M. Cipolla*) *There exists a sequence  $(P_m)_{m \geq 1}$  of polynomials with rational coefficients such that, for any integer  $m$ ,*

$$p_n = n \left[ \ln n + \ln \ln n - 1 + \sum_{r=1}^m \frac{(-1)^r P_r(\ln \ln n)}{\ln^r n} + o\left(\frac{1}{\ln^m n}\right) \right].$$

This was proved by M. Cipolla in a beautiful paper (See [5]) in 1902. In the same paper, Cipolla gives recurrence formula for  $P_m$  and shows that every  $P_m$  has degree  $m$  and leading coefficient  $\frac{1}{m}$ . In particular,

$$P_1(x) = x - 2, P_2(x) = \frac{1}{2}(x^2 - 6x + 11). \quad (4)$$

**Lemma 2.2.** *If  $f$  is monotonic and continuous and defined in  $[1, n]$  and then,*

$$\sum_{r \leq n} f(r) = \int_1^n f(x) dx + O(|f(n)| + |f(1)|).$$

*Proof.* Well known. (See [4], 1.62-1.67, Page 19-20) □

**Theorem 2.3.** *There exists a sequence  $(S_m)_{m \geq 1}$  of polynomials with rational coefficients such that, for any integer  $m$ ,*

$$\sum_{r \leq n} p_r = \frac{n^2}{2} \left[ \ln n + \ln \ln n - \frac{3}{2} + \sum_{r=1}^m \frac{(-1)^r S_r(\ln \ln n)}{\ln^r n} + o\left(\frac{1}{\ln^m n}\right) \right].$$

*Further, every  $S_m$  has degree  $m$  and leading coefficient  $1/m$ . In particular*

$$S_1(x) = x - 3, S_2(x) = \frac{x^2}{2} - \frac{7x}{2} + \frac{27}{4}.$$

*Proof.* We define  $p(x)$  as

$$p(x) = x \ln x + x \ln \ln x - x + \sum_{r=1}^m \frac{(-1)^r x P_r(\ln \ln n)}{\ln^r n}. \quad (5)$$

where  $P_r(x)$  is the same sequence of polynomials as in Theorem 2.1. It follows from Lemma 2.2 that

$$\sum_{r \leq n} p_r = 2 + 3 + \int_3^n p(x) dx + O(p_n) + o\left(\int_3^n \frac{x}{\ln^m x}\right).$$

Now  $p_n \sim n \ln n$  where as

$$\int_3^n \frac{x}{\ln^m x} \sim \frac{n^2}{2 \ln^m n}$$

which grows much faster than  $n \ln n$ . Hence

$$\sum_{r \leq n} p_r = \int_3^n p(x) dx + o\left(\frac{n^2}{\ln^2 n}\right). \quad (6)$$

All we need to do is the integrate each term of 5. Except for a couple of simple terms, integration of the terms of 5 will result in an infinite series and due to 6, we can stop the series when the growth rate of a new term is equal to or slower than the error term in 6. Since  $P_m(\ln \ln x)$  is a polynomial of degree  $m$  and has rational coefficients with leading coefficient  $1/m$ , the integration of each terms of  $p(x)$  will result in an infinite series of the type

$$\int_3^n x P_m(\ln \ln x) dx = \frac{(-1)^m n^2}{2} \sum_{i=1}^{\infty} Q_{m,i}(\ln \ln n) + O(1)$$

where  $Q_{m,i}(x)$  is a polynomial of degree  $m$  with rational coefficients and leading coefficient  $1/m$ . Thus the polynomial  $S_m(x)$  is of degree  $m$  and has rational coefficients with leading coefficient  $1/m$ .

To find the first two terms of the polynomial  $S_m(x)$  we integrate the first four terms of  $p(x)$ . The first four terms of  $p(x)$  are

$$x \ln x + x \ln \ln x - x + \frac{x \ln \ln x - 2x}{\ln x} - \frac{x \ln^2 \ln x - 6x \ln \ln x + 11x}{2 \ln^2 x}. \quad (7)$$

Integrating each term separately, we have

$$\int_3^n x \ln x dx = \frac{n^2 \ln n}{2} - \frac{n^2}{4} + O(1) \quad (8)$$

$$\int_3^n x \ln \ln x dx = \frac{n^2 \ln \ln n}{2} - \frac{n^2}{4 \ln n} - \frac{n^2}{8 \ln^2 n} + O\left(\frac{n^2}{\ln^3 n}\right) \quad (9)$$

$$- \int_3^n x dx = -\frac{n^2}{2} + O(1) \quad (10)$$

$$\int_3^n \frac{x \ln \ln x}{\ln x} dx = \frac{n^2 \ln \ln n}{2 \ln n} + \frac{n^2 \ln \ln n}{4 \ln^2 n} - \frac{n^2}{4 \ln n} + O\left(\frac{n^2 \ln \ln n}{\ln^3 n}\right) \quad (11)$$

$$- 2 \int_3^n \frac{x}{\ln x} dx = -\frac{n^2}{\ln n} - \frac{n^2}{2 \ln^2 n} + O\left(\frac{n^2}{\ln^3 n}\right) \quad (12)$$

$$- \frac{1}{2} \int_3^n \frac{x \ln^2 \ln x}{\ln^2 x} dx = -\frac{n^2 \ln^2 \ln n}{4 \ln^2 n} + O\left(\frac{n^2 \ln^2 \ln n}{\ln^3 n}\right) \quad (13)$$

$$3 \int_3^n \frac{x \ln \ln x}{\ln^2 x} dx = \frac{3n^2 \ln \ln n}{2 \ln^2 n} + O\left(\frac{n^2 \ln^2 \ln n}{\ln^3 n}\right) \quad (14)$$

$$- \frac{11}{2} \int_3^n \frac{x}{\ln^2 x} dx = -\frac{11n^2}{4 \ln^2 n} + O\left(\frac{n^2 \ln n}{\ln^3 n}\right) \quad (15)$$

Adding 8 - 15 we obtain

$$\sum_{r \leq n} p_r = \frac{n^2}{2} \left[ \ln n + \ln \ln n - \frac{3}{2} + \frac{\ln \ln n}{\ln n} - \frac{3}{\ln n} - \frac{\ln^2 \ln n}{2 \ln^2 n} + \frac{7 \ln \ln n}{2 \ln^2 n} - \frac{27}{4 \ln^2 n} + o\left(\frac{1}{\ln^2 n}\right) \right]. \quad (16)$$

Notice that taking the first four terms of 16 we obtain Dusart's result in 1. This proves the theorem.  $\square$

### 3 The inequalities of Mandl and Robin

From the asymptotic expansion of  $\sum_{r \leq n} p_r$  we can not only prove the inequalities of Mandl 1 and Robin 3 but also refine them.

**Lemma 3.1.**

$$\sum_{r \leq n} p_r = \frac{np_n}{2} - \frac{n^2}{4} - \frac{n^2}{2 \ln n} + \frac{n^2 \ln \ln n}{4 \ln^2 n} - \frac{49n^2}{8 \ln^2 n} + o\left(\frac{n^2}{\ln^2 n}\right).$$

*Proof.* Follows trivial by using the asymptotic expansion of  $p_n$  from Theorem 2.1 in 16  $\square$

Since the second and the third term in Lemma 3.1 are negative, it implies that for all sufficiently large  $n$ ,  $\sum_{r \leq n} p_r < \frac{np_n}{2}$ . This proves Mandl's conjecture given in 1. But using Lemma 3.1, we can do much better than Mandl's conjecture and Hassani's bound given in 2. Based on actual computation, we suggest the following refinement of Mandl's inequality. For  $n \geq 835$

$$\sum_{r \leq n} p_r < \frac{np_n}{2} - \frac{n^2}{4} - \frac{n^2}{2 \ln n} + \frac{n^2 \ln \ln n}{4 \ln^2 n}. \quad (17)$$

**Lemma 3.2.**

$$\sum_{r < n} p_r = np_{[n/2]} + \frac{2 \ln 2 - 1}{4} n^2 + O\left(\frac{n^2 \ln \ln n}{\ln n}\right).$$

*Proof.* Taking  $[n/2]$  in place of  $n$  in the asymptotic expansion of  $n^{\text{th}}$  prime we obtain

$$\begin{aligned} np_{[n/2]} &= \frac{n^2}{2} (\ln n + \ln \ln n - 1 - \ln 2) + O\left(\frac{n^2 \ln \ln n}{\ln n}\right) \\ &= \frac{np_n}{2} - \frac{n^2 \ln 2}{2} + O\left(\frac{n^2 \ln \ln n}{\ln n}\right) \end{aligned}$$

Using Lemma 3.1 we can reduce this to

$$= \sum_{r < n} p_r + \frac{n^2}{4} - \frac{n^2 \ln 2}{2} + O\left(\frac{n^2 \ln \ln n}{\ln n}\right)$$

This proves the lemma.  $\square$

Since the second term of Lemma 3.2 is positive, it follows that for all sufficiently large  $n$ , Robin's conjecture is true.

## References

- [1] J. Barkley Rosser and L. Schoenfeld: Sharper Bounds for the Chebyshev Functions  $\theta(x)$  and  $\psi(x)$ , Math. Of Computation, Vol. 29, Number 129 (January 1975) pp. 243-269.
- [2] Pierre Dusart: Sharper bounds for  $\psi, \theta, \pi, p_k$ , Rapport de recherche n 1998-06, Universit de Limoges.
- [3] Mehdi Hassani: A remark on the Mandl's inequality, Octogon Mathematical Magazine, vol. 15 (2007), no.2, pp. 567-572.
- [4] Henryk Iwaniec, Emmanuel Kowalski: Analytic Number Theory; Vol. 53: (2004).
- [5] M. Cipolla: La determinazione assintotica dell  $n^{imo}$  numero primo, Rend. Acad. Sci. Fis. Mat. Napoli 8(3) (1902) 132166.

e-mail: [nilotpalsinha@gmail.com](mailto:nilotpalsinha@gmail.com)