**Notes on Number Theory and Discrete Mathematics**
**Print ISSN 1310–5132, Online ISSN 2367–8275**
**2024, Volume 30, Number 3, 580–586**
**DOI: 10.7546/nntdm.2024.30.3.580-586**

# New estimations for numerical analysis approach to twin primes conjecture

**Gabriele Di Pietro**

Via delle Ville, 18, 64026, Roseto degli Abruzzi (TE), Italy
e-mail: petreius@libero.it

**Received:** 27 February 2024 **Revised:** 3 October 2024
**Accepted:** 9 October 2024 **Online First:** 11 October 2024

**Abstract:** This paper provides a better approximation of the functions presented in the article "Numerical Analysis Approach to Twin Primes Conjecture" (see [3]). The new estimates highlight the approximations used in the previous article and the validity of Theorems 1 and 2 through the use of the false hypothesis based on the distribution of primes punctually following the Logarithmic Integral $\mathrm{Li}(x)$ (see [4] and [7], pp. 174–176) will be re-evaluated.
**Keywords:** Numerical analysis, Number theory, Sieves.
**2020 Mathematics Subject Classification:** 11N35.

## 1 Introduction

This paper represents a follow-up of the previous article "Numerical Analysis Approach to Twin Primes Conjecture" (see [3]) in which the sieve of Eratosthenes was modified to obtain the twin primes counting function called $\Upsilon(x)$. In order to avoid burdening the discussion, a summary of the main notations and reasonings will be carried out, focusing on the new estimations and the validity of the theorems on the basis of the latest modifications.

First of all, the sieve of Eratosthenes should be considered until a certain number $x$. For the sake of simplicity, consider a table divided in two specific parts, namely the eliminators zone from 1 to $\sqrt{x}$ and the would-be twin primes zone from $\sqrt{x}$ to $x$. The numbers of the $6k+1$ and $6k-1$

Copyright © 2024 by the Author. This is an Open Access paper distributed under the terms and conditions of the Creative Commons Attribution 4.0 International License (CC BY 4.0). https://creativecommons.org/licenses/by/4.0/

forms will also be called would-be twin primes, then the amount of couples $(6k+1, 6k-1)$ in the would-be twin prime zone will be counted. Through this calculation, the obtained result is

$$\left\lfloor \frac{x}{6} \right\rfloor - \left\lfloor \frac{\sqrt{x}}{6} \right\rfloor, \tag{1}$$

from which the amount of the twin prime numbers deleted from the would-be twin prime zone will be deducted.

At this point, consider

$$K = \left\lfloor \frac{x}{3y} \right\rfloor - \left\lfloor \frac{\sqrt{x}}{3y} \right\rfloor, \tag{2}$$

where $y$ represents a number in the eliminators zone and $K$ is the amount of the would-be twin prime numbers deleted by $y$. If we assign to $y$ the prime numbers $p$ in the eliminators zone $(5 \leq p \leq \sqrt{x})$, we shall obtain

$$\sum_{p \in [5, \sqrt{x}]} K = \sum_{p \in [5, \sqrt{x}]} \left\lfloor \frac{x}{3p} \right\rfloor - \left\lfloor \frac{\sqrt{x}}{3p} \right\rfloor, \tag{3}$$

which represents the total number of deletions within the would-be twin prime zone. The set of these deletions contains the repetitions due to the fact that a would-be prime number may be deleted by more eliminators, so there is a need for a function which shall take into account the various factors of a number $x$ in order to remove the repetitions from (3).

In literature it is called $\omega(x)$ (Hardy and Wright in [6], p. 354) and is defined as the function counting the distinct prime factors of a $x$ number. Here we can define the amount for which the deletions of (3) should be divided

$$L = \frac{1}{\#Z} \sum_{z \in [1, x]} \omega(z), \tag{4}$$

where $z$ is an integer number of the form $6k+1$ or $6k-1$ in the $[1, x]$ Sieve of Eratosthenes table; $Z$ is the set of integer numbers $z$ and $\#Z$ is its cardinality.

The function (4) encloses the parity problem (see [9]) introduced by Selberg in 1949 and in order to bypass it, we should rely on the result

$$\frac{1}{L} \sum_{p \in [5, \sqrt{x}]} K = \left\lfloor \frac{x}{3} \right\rfloor - \left\lfloor \frac{\sqrt{x}}{3} \right\rfloor - \left[\pi(x) - \pi(\sqrt{x})\right], \tag{5}$$

which does not need any approximation of $\omega(x)$, but rises the problem of re-adding within the zone of deletions all that ones that do not distinguish the would-be prime of the forms $6k+1$ and $6k-1$ that is to say the ones that will delete both the would-be prime and its twin. It can be defined as

$$T = \sum_{p \in [\sqrt{x}, x]} \left\lfloor \frac{\operatorname{dist}(p) - 3}{6} \right\rfloor, \tag{6}$$

where $\operatorname{dist}(p)$ represents the distance function between $p$ (prime) and the following prime number.

However, using the false hypothesis that the average distance between two consecutive prime numbers may be $\ln(x)$,

$$\tilde{T} = \sum_{p\in[\sqrt{x},x]} \left\lfloor \frac{\ln(p)-3}{6} \right\rfloor \tag{7}$$

can be defined as an approximation of (6). The choice of such hypothesis derives from the observation of the logarithmic integral function $\mathrm{Li}(x)$ (see [2], pp. 116–117) introduced by Gauss:

$$\mathrm{Li}(x) = \frac{x}{\ln x} \sum_{k=0}^{\infty} \frac{k!}{(\ln x)^k}, \tag{8}$$

in which the development in its first order is $\frac{x}{\ln(x)}$. The set of these three functions allows us to define the twin prime couples counting function

$$\Upsilon(x) = \left\lfloor \frac{x}{6} \right\rfloor - \left\lfloor \frac{\sqrt{x}}{6} \right\rfloor - \frac{1}{L} \sum_{p\in[5,\sqrt{x}]} K + T \tag{9}$$

and its approximation

$$\tilde{\Upsilon}(x) = \left\lfloor \frac{x}{6} \right\rfloor - \left\lfloor \frac{\sqrt{x}}{6} \right\rfloor - \frac{1}{\tilde{L}} \sum_{p\in[5,\sqrt{x}]} \tilde{K} + \tilde{T}, \tag{10}$$

where

$$\frac{1}{\tilde{L}} \sum_{p\in[5,\sqrt{x}]} \tilde{K} = \left\lfloor \frac{x}{3} \right\rfloor - \left\lfloor \frac{\sqrt{x}}{3} \right\rfloor - \left( \frac{x}{\ln x} - \frac{\sqrt{x}}{\ln \sqrt{x}} \right) \tag{11}$$

by approximating (5) with the prime number theorem (see [1,5]).

In 1962, Rosser and Schoenfeld published “Approximate formulas for some functions of prime numbers” (see [8]). This paper will rely on their results in order to recalculate (7) and it is structured as follows: Section 2 is used to compare approximations of (7) and its new version; in Section 3 a new version of (10) may be obtained and its monotonicity may be studied; Section 4 is characterized by an extension of the results in Section 3 to any distance of primes $k\ln(x)$, where $0 < k < 1$.

## 2 A better estimate for $\tilde{T}$

Relying on Rosser and Schoenfeld’s introduction in [8], $f(p)$ can be defined as the following function:

$$f : P \subset \mathbb{N} \to \mathbb{R}, \tag{12}$$

where $P$ represents the set of prime numbers. At this point, we can consider the logarithm product of primes which are less than a natural number $x$

$$\theta(x) = \ln\left(\prod_{p\leq x} p\right) = \sum_{p\leq x} \ln(p), \tag{13}$$

or the function counting the amount of primes which are less than $x$

$$\pi(x) = \sum_{p \le x} 1. \tag{14}$$

Bearing in mind the classic definition of Gaussian $\mathrm{Li}(x)$ and its property of representing the average distribution of primes, it is natural to approximate

$$\sum_{p \le x} f(p) \approx \int_2^x \frac{f(y)}{\ln y} dy, \tag{15}$$

which, if applied to (13), is

$$\theta(x) = \sum_{p \le x} \ln(p) \approx \int_2^x dy \approx x, \tag{16}$$

while if applied to (14), is

$$\pi(x) = \sum_{p \le x} 1 \approx \int_2^x \frac{1}{\ln(y)} dy, \tag{17}$$

as clearly demonstrated by De la Vallée Poussin in [1].

The Theorem 4 in [8], p. 70, provides a lower bound for (13):

$$x\left(1 - \frac{1}{2\ln(x)}\right) < \theta(x) = \sum_{p \le x} \ln(p), \quad x \ge 563 \tag{18}$$

that will be used for the following calculations. In fact, by developing (7) through the prime number theorem, one has

$$\tilde{T} = \sum_{p \in [\sqrt{x}, x]} \left\lfloor \frac{\ln(p)}{6} \right\rfloor - \left\lfloor \frac{3}{6}\left(\frac{x}{\ln(x)} - \frac{\sqrt{x}}{\ln\sqrt{x}}\right) \right\rfloor, \tag{19}$$

from which, by using the inequality in (18), the result is

$$\sum_{p \in [\sqrt{x}, x]} \frac{\ln(p)}{6} \sim \frac{x}{6}\left(1 - \frac{1}{2\ln(x)}\right) - \frac{\sqrt{x}}{6}\left(1 - \frac{1}{2\ln(\sqrt{x})}\right). \tag{20}$$

So the function in (19) may be approximated by the quantity that will be called $\tilde{S}$ in order to differentiate it from $\tilde{T}$, avoiding any duplication of notation:

$$\tilde{T} \approx \tilde{S} = \frac{2x\ln(x) - x - 2\sqrt{x}\ln(x) + 2\sqrt{x}}{12\ln x} - \frac{3}{6}\left(\frac{x}{\ln(x)} - \frac{\sqrt{x}}{\ln\sqrt{x}}\right), \tag{21}$$

that is to say

$$\tilde{S} = \frac{2x\ln(x) - 2\sqrt{x}\ln(x) - 7x + 14\sqrt{x}}{12\ln x}. \tag{22}$$

The latter represents a better approximation of the function (19) so it is logical to suppose that any other chosen approximation may have $\tilde{S}$ as an upper bound. It can be verified, for example, for the $\tilde{T}$ approximation in [3]:

$$\tilde{T} = \frac{\ln(x)}{6}\left(\frac{x}{\ln(x)} - \frac{\sqrt{x}}{\ln\sqrt{x}}\right) - \frac{5.5}{6}\left(\frac{x}{\ln(x)} - \frac{\sqrt{x}}{\ln\sqrt{x}}\right) \tag{23}$$

due to the inequality

$$\frac{3}{6} < \frac{5.5}{6}. \tag{24}$$

In fact, one has

$$\frac{\ln(x) - 5.5}{6}\left(\frac{x}{\ln(x)} - \frac{\sqrt{x}}{\ln\sqrt{x}}\right) = \frac{2x\ln(x) - 4\sqrt{x}\ln(x) - 11x + 22\sqrt{x}}{12\ln x} \tag{25}$$

and the difference between (22) and (25) is

$$\frac{2x\ln(x) - 2\sqrt{x}\ln(x) - 7x + 14\sqrt{x} - 2x\ln(x) + 4\sqrt{x}\ln(x) + 11x - 22\sqrt{x}}{12\ln x}, \tag{26}$$

hence

$$\frac{2\sqrt{x}\ln(x) + 4x - 8\sqrt{x}}{12\ln x} > 0 \tag{27}$$

for any $x \geq 563$ as stated in (18), proving

$$\tilde{S} > \frac{\ln(x) - 5.5}{6}\left(\frac{x}{\ln(x)} - \frac{\sqrt{x}}{\ln\sqrt{x}}\right) \tag{28}$$

for any $x \geq 563$.

# 3 A new approximation of twin prime couples function

On the basis of the latest considerations, it may be interesting to test the twin prime couples counting function defined as (10) through the use of the new approximations. Instead of considering the natural $\tilde{S}$, assign the following limit value:

$$\tilde{T}_{min} = \frac{2x\ln(x) - x - 2\sqrt{x}\ln(x) + 2\sqrt{x}}{12\ln x} - \frac{5.5}{6}\left(\frac{x}{\ln(x)} - \frac{\sqrt{x}}{\ln\sqrt{x}}\right), \tag{29}$$

which represents the minimum value of approximation of (19) in view of the above. At this point, the approximation of $\Upsilon(x)$ may be re-evaluated and the twin prime couples counting function is the following:

$$\tilde{\Upsilon}(x) = \left\lfloor\frac{x}{6}\right\rfloor - \left\lfloor\frac{\sqrt{x}}{6}\right\rfloor - \frac{1}{\tilde{L}}\sum_{\tilde{y}\in\pi(\sqrt{x})}\tilde{K} + \tilde{T}_{\min}, \tag{30}$$

where

$$\frac{1}{\tilde{L}}\sum_{p\in[5,\sqrt{x}]}\tilde{K} = \left\lfloor\frac{x}{3}\right\rfloor - \left\lfloor\frac{\sqrt{x}}{3}\right\rfloor - \left(\frac{x}{\ln x} - \frac{\sqrt{x}}{\ln\sqrt{x}}\right), \tag{31}$$

**Theorem 1.** *Supposing a punctually* $\mathrm{Li}(x)$ *function of first order, the* $\tilde{\Upsilon}(x)$ *function, defined by* (29)–(31)*, is identically equal to zero. This function is an approximation of* $\Upsilon(x)$*, so the numerical twin prime couples counting function is* $\Upsilon(x) > \tilde{\Upsilon}(x) = 0$ *for any* $x \geq 563$.

*Proof.*

$$\frac{\sqrt{x} - x}{6} + \frac{2x\ln(x) - x - 2\sqrt{x}\ln(x) + 2\sqrt{x}}{12\ln x} + \frac{0.5}{6}\left(\frac{x}{\ln(x)} - \frac{\sqrt{x}}{\ln\sqrt{x}}\right) = 0,$$

$$\frac{2\sqrt{x}\ln(x) - 2x\ln(x) + 2x\ln(x) - x - 2\sqrt{x}\ln(x) + 2\sqrt{x}}{12\ln x} + \frac{0.5}{6}\left(\frac{x}{\ln(x)} - \frac{\sqrt{x}}{\ln\sqrt{x}}\right) = 0,$$

using the properties of logarithms

$$\frac{2\sqrt{x}-x}{12\ln x}+\frac{0.5}{6}\left(\frac{x}{\ln x}-\frac{2\sqrt{x}}{\ln x}\right)=0,$$

$$\frac{2\sqrt{x}-x+x-2\sqrt{x}}{12\ln x}=0,$$

simplifying

$$\frac{0}{12\ln x}=0.$$

The assertion $\Upsilon(x)>\tilde{\Upsilon}(x)=0$ for any $x\geq 563$ is a consequence of the previous approximations in Section 2, proving the theorem. $\square$

Obviously, choosing values between $\frac{3}{6}$ and $\frac{5.5}{6}$ in (24), the function $\tilde{\Upsilon}(x)$ will be greater than zero further justifying Theorem 1 proof.

# 4 Extension to upper orders of $\mathbf{Li}(x)$

By calling the Logarithmic Integral in (8), it may be observed that the distribution of primes $\frac{x}{\ln(x)}$ is simply its first order expansion. In addition, as in Section 6 of [3], for any $0<k<1$, we may consider $\pi(x)\sim\frac{1}{k}\frac{x}{\ln x}\sim\mathrm{Li}(x)$ and assume as a false hypothesis that $k\ln(x)$ may be the average distance between twin primes.

Consequently, we wonder if the inequality in (18) is still valid through the introduction of the new average but in order to establish that we need (12) and (17) to develop (15) by using the Stieltjes integral:

$$\sum_{p\leq x}f(p)=\int_{2^-}^{x}f(y)d(\pi(y)). \tag{32}$$

Integration by parts provides

$$\sum_{p\leq x}f(p)=f(x)\pi(x)-\int_{2}^{x}f'(y)\pi(y)dy, \tag{33}$$

so the approximations can be applied as follows:

$$\sum_{p\leq x}k\ln(p)=\frac{kx\ln(x)}{k\ln(x)}-\int_{2}^{x}\frac{k}{y}\frac{y}{k\ln(y)}dy=x-\int_{2}^{x}\frac{1}{\ln(y)}dy\sim x\left(1-\frac{1}{\ln(x)}\right) \tag{34}$$

as the approximation in (17) has been used in the last stage. Now, following the proof of Theorem 4 in [8],

$$x\left(1-\frac{1}{2\ln(x)}\right)<\sum_{p\leq x}k\ln(p) \tag{35}$$

is valid for each average distance $k\ln(x)$.

**Theorem 2.** *Using Theorem 1 with $\frac{1}{k}\cdot\frac{x}{\ln(x)}$ as a distribution of primes and $k\ln(x)$ as an average distance between twin primes, then the approximating function $\tilde{\Upsilon}(x)$ is positive where $x>4$ and $0<k<1$.*

*Proof.* On the basis of the previous demonstration we obtain

$$\frac{\sqrt{x}-x}{6}+\frac{2x\ln(x)-x-2\sqrt{x}\ln(x)+2\sqrt{x}}{12\ln x}+\frac{0.5}{6}\frac{1}{k}\left(\frac{x}{\ln x}-\frac{2\sqrt{x}}{\ln x}\right)>0,$$

from which

$$\frac{2\sqrt{x}\ln(x)-2x\ln(x)+2x\ln(x)-x-2\sqrt{x}\ln(x)+2\sqrt{x}+\frac{x}{k}-\frac{2\sqrt{x}}{k}}{12\ln x}>0,$$

and so

$$\frac{x\left(\frac{1}{k}-1\right)-2\sqrt{x}\left(\frac{1}{k}-1\right)}{12\ln x}>0,$$

where the numerator is greater than zero for any $x>4$ and the denominator when $x>1$ ensuring that $\tilde{\Upsilon}(x)$ is increasingly monotonous for any $x>4$ where $0<k<1$. $\square$

Obviously, for previous approximations, $\Upsilon(x)>\tilde{\Upsilon}(x)\geq 563$ and, in order to prove that there are infinite twin prime couples supposing a punctually $\mathrm{Li}(x)$ function distribution of primes, it is sufficient to assume for any $0<k<1$ that

$$\frac{1}{k}=\sum_{t=0}^{\infty}\frac{t!}{(\ln x)^t}.$$

# References


[1] De la Vallée Poussin, C. J. (1896). Recherches analytiques la théorie des nombres premiers. *Annales de la Société scientifique de Bruxelles*, 20, 183–256.

[2] Derbyshire, J. (2004). *Prime Obsession: Bernhard Riemann and the Greatest Unsolved Problem in Mathematics*. New York: Penguin.

[3] Di Pietro, G. (2021). Numerical analysis approach to twin primes conjecture. *Notes on Number Theory and Discrete Mathematics*, 27(3), 175–183.

[4] Gauss, C. F. (1863). *Werke, Band 10*, Teil I. p. 10.

[5] Hadamard, J. (1896). Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques. *Bulletin de la Societe Mathematique de France*, 24, 199–220.

[6] Hardy, G. H., & Wright, E. M. (1979). *An Introduction to the Theory of Numbers* (5th ed.). Oxford Science Publications, Oxford, England: Clarendon Press, pp. 354–358.

[7] Havil, J., & Dyson, F. (2003). Problems with Primes. *In Gamma: Exploring Euler's Constant*. Princeton; Oxford: Princeton University Press, pp. 163–188.

[8] Rosser, J. B., & Schoenfeld, L. (1962). Approximate formulas for some functions of prime numbers. *Illinois Journal of Mathematics*, 6(1), 64–94.

[9] Selberg, A. (1949). An elementary proof of the Prime-Number Theorem. *Annals of Mathematics*, 50(2), 305–313.