Int. J. Contemp. Math. Sci., Vol. 3, 2008, no. 22, 1069 - 1086

An Approximate Formula for Prime Numbers

Rafael Jakimczuk

Division Matematica, Universidad Nacional de Lujan
Buenos Aires, Argentina
jakimczu@mail.unlu.edu.ar

Abstract

Let p_n be the n-th prime number. We prove if $n \geq 4$ the following formula holds

$$\begin{aligned} p_n &= n\log n + \log(n\log n)(n - Li(n\log n)) \\ &+ \sum_{k=2}^{\infty} \frac{(-1)^k Q_{k-1}(\log(n\log n))}{k! n^{k-1}\log^{k-1} n}(n - Li(n\log n))^k + O(h(n)) \end{aligned}$$

Where $Li(x)$ is the well known classical function $\int_2^x \frac{1}{\log t}\, dt$, $h(n) = \frac{n\log^2 n}{exp(d\sqrt{\log n})}$ and the $Q_{k-1}(x)$ are polynomials defined in this article. Note that the mistake term $O(h(n))$ in this formula is better than the mistake term in others previous known aproximate formulas for p_n. If the Riemann's hypothesis is true then $h(n) = \sqrt{n}\log^{5/2} n$.

Mathematics Subject Classification: 11A41, 11N05

Keywords: Prime numbers, approximate formula, the function $Li(x)$.

1 The function $y = Li(x)$

Let us consider the classical function

$$Li(x) = \int_2^x \frac{1}{\log t}\, dt \qquad (x > 1)$$

Note that $\lim_{x\to 1} Li(x) = -\infty$. Its first derivative is

$$Li^{(1)}(x) = \frac{1}{\log x} \tag{1}$$

The following derivatives satisfy the following theorem

Theorem 1.1 *We have*

$$Li^{(k+1)}(x) = (-1)^k \frac{P_{k-1}(\log x)}{x^k \log^{k+1} x} \qquad (k \geq 1) \tag{2}$$

Where $P_{k-1}(\log x)$ is a polynomial in $\log x$ of degree $k-1$ and positive integer coefficients.

The polynomials $P_k(X)$ can be obtained from the following recursive formulae

$$P_0(X) = 1$$

$$P_k(X) = kXP_{k-1}(X) + (k+1)P_{k-1}(X) - XP'_{k-1}(X) \qquad (k \geq 1) \tag{3}$$

Example.

$$P_1(X) = X+2 \qquad P_2(X) = 2X^2+6X+6 \qquad P_3(X) = 6X^3+22X^2+36X+24$$

Proof. Mathematical induction

Remark 1. The function $Li(x)$ is increasing and convex since its first derivative $\frac{1}{\log x}$ is positive and its second derivative $-\frac{1}{x\log^2 x}$ is negative.

Theorem 1.2 *Let $a > 1$, then*

$$Li(x) = Li(a) + \frac{1}{\log a}(x-a) + \sum_{k=1}^{\infty} \frac{(-1)^k P_{k-1}(\log a)}{(k+1)!a^k \log^{k+1} a}(x-a)^{k+1} \tag{4}$$

where $x \in (1, 2a-1)$.

Proof. The function

$$g(z) = \frac{1}{\log z}$$

where

$$\log z = \log \rho + i\theta \qquad z = \rho(\cos\theta + i\sin\theta) \qquad (-\pi < \theta \leq \pi)$$

is analitical in the semiplane $x > 1$. Hence (theorem 1.1) we find that

$$\frac{1}{\log x} = \frac{1}{\log a} + \sum_{k=1}^{\infty} \frac{(-1)^k P_{k-1}(\log a)}{k!a^k \log^{k+1} a}(x-a)^k$$

where $x \in (1, 2a-1)$. The theorem is proved.

2 The inverse function of $y = Li(x)$: $x = Li^{-1}(y)$.

The function $y = Li(x)$ is increasing in the interval $(1, \infty)$, consequently it has an inverse function $x = Li^{-1}(y)$ in the interval $(-\infty, \infty)$. This inverse function is positive, increasing and it has infinite derivatives. The first derivative is

$$\log\left(Li^{-1}(y)\right) \tag{5}$$

We have the following theorem.

Theorem 2.1 *The k-th derivative of $x = Li^{-1}(y)$ is*

$$(-1)^k \frac{Q_{k-1}(\log x)}{x^{k-1}} = (-1)^k \frac{Q_{k-1}(\log(Li^{-1}(y)))}{\left(Li^{-1}(y)\right)^{k-1}} \qquad (k \geq 2) \tag{6}$$

Where $Q_{k-1}(\log x)$ is a polynomial in $\log x$ of degree $k-1$ and integer coefficients alternately positive and negative, the leading coefficient is positive and the last coefficient is zero.

The polynomials $Q_k(X)$ can be obtained from the following recursive formulae

$$Q_1(X) = X$$
$$Q_k(X) = X\left((k-1)Q_{k-1}(X) - Q'_{k-1}(X)\right) \qquad (k \geq 2) \tag{7}$$

Example.

$$Q_2(X) = X^2 - X \qquad Q_3(X) = 2X^3 - 4X^2 + X$$
$$Q_4(X) = 6X^4 - 18X^3 + 11X^2 - X$$

Proof. Mathematical Induction.

Theorem 2.2 *Let us consider the function*

$$y = f(x) = a_1 x + a_2 x^2 + a_3 x^3 + \ldots \qquad (-l < x < l) \tag{8}$$

where $a_1 = 1$. Since its first derivative in $x = 0$ is $1 \neq 0$, this function has an inverse function $x = g(y)$ in a neighborhood of $y = 0$ which has infinite derivatives.

Let us consider the function

$$b_1 y + b_2 y^2 + b_3 y^3 + \ldots \tag{9}$$

where $b_n = \frac{g^{(n)}(0)}{n!}$. Being $g^{(n)}(0)$ the n-th derivative of $x = g(y)$ in $y = 0$.

Let ρ_1 be a positive number less than l, and let M_1 be a positive number such that $|a_n| \rho_1^n \leq M_1$. Then (9) converges in the interval $(-\mu_1, \mu_1)$ where

$$\mu_1 = 2M_1 + \rho_1 - 2\sqrt{M_1 (M_1 + \rho_1)}$$

Let ρ_2 be a positive number less than μ_1, and let M_2 be a positive number such that $|b_n|\,\rho_2^n \leq M_2$. Then

$$x = g(y) = b_1 y + b_2 y^2 + b_3 y^3 + \dots$$

in the interval $(-\mu_2, \mu_2)$, where

$$\mu_2 = \frac{\rho_2 l}{l + M_2}$$

Proof. This theorem is an inmediate consequence of the Art. 55 (page 156) and the Art. 36 (page 95) of the reference [1].

We shall use theorem 2.2 as a fundamental lemma in the proof of the following theorem.

Theorem 2.3 *From a certain value of a we have*

$$Li^{-1}(y) = a + \log a(y - Li(a)) + \sum_{k=2}^{\infty} \frac{(-1)^k Q_{k-1}(\log a)}{k! a^{k-1}} (y - Li(a))^k \qquad (10)$$

in the interval

$$\left(Li(a) - \frac{c\ a}{4\log^{3/2} a}, Li(a) + \frac{c\ a}{4\log^{3/2} a}\right) \qquad (c = 3 - 2\sqrt{2}) \qquad (11)$$

Proof. We shall suppose a sufficiently large.

Let us consider the function $Li(x + a) - Li(a)$, that is (theorem 1.2)

$$\frac{1}{\log a} x + \sum_{k=1}^{\infty} \frac{(-1)^k P_{k-1}(\log a)}{(k+1)! a^k \log^{k+1} a} x^{k+1} \qquad -(a-1) < x < (a-1) \qquad (12)$$

From the derivatives of the inverse function of (12) in $y = 0$ we obtain the function (see theorem 2.1)

$$\log a\ y + \sum_{k=2}^{\infty} \frac{(-1)^k Q_{k-1}(\log a)}{k! a^{k-1}} y^k \qquad (13)$$

If we multiply (12) by $\log a$ we obtain the function

$$y = x + \sum_{k=1}^{\infty} \frac{(-1)^k P_{k-1}(\log a)}{(k+1)! a^k \log^k a} x^{k+1} \qquad -(a-1) < x < (a-1) \qquad (14)$$

This function is the function (8) of theorem 2.2.

The function (9) of theorem 2.2 will be (see (13))

$$y+\sum_{k=2}^{\infty}\frac{(-1)^k Q_{k-1}(\log a)}{k!a^{k-1}\log^k a}y^k \tag{15}$$

Let choose us (see theorem 2.2)

$$0<\rho_1=f(a)=\frac{a}{\log^{1/2}a}<a-1=l$$

If we substituting $x=\rho_1=f(a)$ into (14) then we obtain a series where the quotient of the absolute value of a term with the absolute value of the former term will be $(k\geq 1)$

$$\begin{aligned}
&\frac{P_k(\log a)f(a)^{k+2}}{(k+2)!a^{k+1}\log^{k+1}a}\frac{(k+1)!a^k\log^k a}{P_{k-1}(\log a)f(a)^{k+1}}=\frac{f(a)}{(k+2)a\log a}\frac{P_k(\log a)}{P_{k-1}(\log a)}\\
=\;&\frac{f(a)}{(k+2)a\log a}\frac{k\log aP_{k-1}(\log a)+(k+1)P_{k-1}(\log a)-\log aP'_{k-1}(\log a)}{P_{k-1}(\log a)}\\
\leq\;&\frac{f(a)}{(k+2)a\log a}\frac{k\log aP_{k-1}(\log a)+(k+1)P_{k-1}(\log a)}{P_{k-1}(\log a)}\\
=\;&\frac{f(a)}{a\log a}\left(\frac{k}{k+2}\log a+\frac{k+1}{k+2}\right)<\frac{f(a)}{a\log a}(\log a+1)=\frac{1+\log a}{\log^{3/2}a}<1
\end{aligned}$$

That is, less than 1 from a certain value of a. Consequently, the term of greatest absolute value will be between the first term

$$\rho_1=f(a)=\frac{a}{\log^{1/2}a}$$

and the second term

$$\frac{f(a)^2}{2a\log a}=\frac{a}{2\log^2 a}$$

Now, from a certain value of a the first term is greater than the second term. Then we shall take

$$M_1=\rho_1=\frac{a}{\log^{1/2}a}$$

Therefore

$$\mu_1=2M_1+\rho_1-2\sqrt{M_1\left(M_1+\rho_1\right)}=(3-2\sqrt{2})\rho_1=c\frac{a}{\log^{1/2}a} \qquad (c=3-2\sqrt{2})$$

Now, let choose us

$$0<\rho_2=g(a)=\frac{c\ a}{2\log^{1/2}a}<\mu_1$$

Let us consider the polynomials $H_k(X)$ whose coefficients are the absolute values of the coefficients of the polynomials $Q_k(X)$. Clearly, these polynomials can be obtained from the following recursive formulae

$$H_1(X) = X$$

$$H_k(X) = (k-1)XH_{k-1}(X) + XH'_{k-1}(X) \qquad (k \geq 2)$$

If we substituting $y = \rho_2 = g(a)$ into (15) then we obtain a series where the absolute value of each term from the second satisfies the inequality

$$\left| \frac{(-1)^k Q_{k-1}(\log a)}{k! a^{k-1} \log^k a} g(a)^k \right| \leq \frac{H_{k-1}(\log a)}{k! a^{k-1} \log^k a} g(a)^k \qquad (k \geq 2)$$

Let us consider the series

$$g(a) + \sum_{k=2}^{\infty} \frac{H_{k-1}(\log a)}{k! a^{k-1} \log^k a} g(a)^k$$

The quotient of a term with the former term is $(k \geq 2)$

$$\begin{aligned}
& \frac{H_k(\log a) g(a)^{k+1}}{(k+1)! a^k \log^{k+1} a} \frac{k! a^{k-1} \log^k a}{H_{k-1}(\log a) g(a)^k} = \frac{g(a)}{(k+1) a \log a} \frac{H_k(\log a)}{H_{k-1}(\log a)} \\
= \; & \frac{g(a)}{(k+1) a \log a} \frac{(k-1) \log a H_{k-1}(\log a) + \log a H'_{k-1}(\log a)}{H_{k-1}(\log a)} \\
= \; & \frac{g(a)}{a} \left(\frac{k-1}{k+1} + \frac{\frac{H'_{k-1}(\log a)}{k-1}}{\frac{H_{k-1}(\log a)}{\log a}} \frac{k-1}{k+1} \frac{1}{\log a} \right) < \frac{g(a)}{a} \left(1 + \frac{1}{\log a} \right) \\
= \; & \frac{c(1 + \log a)}{2 \log^{3/2} a} < 1
\end{aligned}$$

That is, less than 1 from a certain value of a. Consequently, the greatest term will be between the first term

$$\rho_2 = g(a) = \frac{c\, a}{2 \log^{1/2} a}$$

and the second term

$$\frac{g(a)^2}{2a \log a} = \frac{c^2 a}{8 \log^2 a}$$

Now, from a certain value of a the first term is greater than the second term. Then we shall take

$$M_2 = \rho_2 = \frac{c\, a}{2 \log^{1/2} a}$$

Therefore (15) is the inverse function of (14) in the interval $(-\mu_2, \mu_2)$, where (see theorem 2.2)

$$\mu_2 = \frac{\rho_2 l}{l + M_2} \sim \frac{c\ a}{2 \log^{1/2} a}$$

On the other hand, from a certain value of a we have

$$\frac{c\ a}{4 \log^{1/2} a} < \mu_2$$

Therefore (15) is the inverse function of (14) in the interval

$$\left(-\frac{c\ a}{4 \log^{1/2} a}, \frac{c\ a}{4 \log^{1/2} a} \right)$$

Consequently (13) is the inverse function of (12) in the interval

$$\left(-\frac{1}{\log a} \frac{c\ a}{4 \log^{1/2} a}, \frac{1}{\log a} \frac{c\ a}{4 \log^{1/2} a} \right) = \left(-\frac{c\ a}{4 \log^{3/2} a}, \frac{c\ a}{4 \log^{3/2} a} \right)$$

The theorem is thus proved.

3 The sequence $Li^{-1}(n)$

It is well known that (prime number theorem)

$$\pi(x) = Li(x) + O(f(x)) \tag{16}$$

where for example (de la Vallee Poussin)

$$f(x) = \frac{x}{exp(d\sqrt{\log x})} \qquad (d > 0)$$

There exist in the literature better functions $f(x)$. For sake of simplicity we shall use the de la Vallee Poussin's function.

On the other hand, if the Riemann's hypothesis is true then (Von Koch)

$$f(x) = \sqrt{x} \log x$$

Theorem 3.1 *If $a = n \log n$, then from a certain value of n we have*

$$0 < n - Li(a) < \frac{c\ a}{4 \log^{3/2} a} \tag{17}$$

Proof. We shall suppose n sufficiently large.

Substituting $x = p_n$ into (16) we obtain

$$|n - Li(p_n)| < Kf(p_n) \qquad (K > 0)$$

That is

$$Li(p_n) - Kf(p_n) < n < Li(p_n) + Kf(p_n) \tag{18}$$

We shall chosse in (16) the de la Vallee Poussin's function

$$f(x) = \frac{x}{exp(d\sqrt{\log x})}$$

Therefore

$$f(p_n) = \frac{p_n}{exp(d\sqrt{\log p_n})} \sim \frac{n \log n}{exp(d\sqrt{\log n})} \tag{19}$$

It is well known the aproximate formula

$$p_n = n \log n + n \log \log n - n + o(n) \tag{20}$$

Hence, we have

$$n \log n < p_n \tag{21}$$

(19) and (20) give

$$\frac{1}{\log p_n}(p_n - n \log n) - Kf(p_n) \sim \frac{n \log \log n}{\log n}$$

Therefore we have

$$\frac{1}{\log p_n}(p_n - n \log n) - Kf(p_n) > 0 \tag{22}$$

Consequently from (18), the Lagrange's theorem and (22) we find that

$$n - Li(n \log n) \geq Li(p_n) - Li(n \log n) - Kf(p_n) = \frac{1}{\log b}(p_n - n \log n)$$

$$- \quad Kf(p_n) \geq \frac{1}{\log p_n}(p_n - n \log n) - Kf(p_n) > 0$$

Where $n \log n < b < p_n$. That is

$$n - Li(n \log n) > 0 \tag{23}$$

If $a = n \log n$ we obtain

$$\frac{c\ a}{4 \log^{3/2} a} = \frac{c\ (n \log n)}{4 \log^{3/2}(n \log n)} \sim \frac{c\ n}{4 \log^{1/2} n}$$

Therefore we have

$$\frac{c\ n}{8\log^{1/2} n} < \frac{c\ a}{4\log^{3/2} a} \tag{24}$$

From (23) and (18) we obtain

$$0 < n - Li(n\log n) \leq Li(p_n) - Li(n\log n) + Kf(p_n) \tag{25}$$

From the Lagrange's theorem and (21) we find that

$$Li(p_n) - Li(n\log n) = \frac{1}{\log b}(p_n - n\log n) \leq \frac{p_n - n\log n}{\log(n\log n)} \tag{26}$$

where $n\log n < b < p_n$. Equations (25) and (26) give

$$0 < n - Li(n\log n) \leq \frac{p_n - n\log n}{\log(n\log n)} + Kf(p_n) \tag{27}$$

(19) and (20) give

$$\frac{p_n - n\log n}{\log(n\log n)} + Kf(p_n) \sim \frac{n\log\log n}{\log n} \tag{28}$$

From (27) and (28) we find that

$$0 < n - Li(n\log n) < \frac{2\ n\log\log n}{\log n} \tag{29}$$

Now

$$\frac{2\ n\log\log n}{\log n} < \frac{c\ n}{8\log^{1/2} n}$$

Therefore

$$0 < n - Li(n\log n) < \frac{c\ n}{8\log^{1/2} n} \tag{30}$$

Finally, (24) and (30) give (17). The theorem is thus proved.

Theorem 3.2 *From a certain value of n we have*

$$\begin{aligned} Li^{-1}(n) &= n\log n + \log(n\log n)(n - Li(n\log n)) \\ &+ \sum_{k=2}^{\infty} \frac{(-1)^k Q_{k-1}(\log(n\log n))}{k! n^{k-1} \log^{k-1} n}(n - Li(n\log n))^k \end{aligned}$$

Proof. It is an inmediate consequence of theorem 3.1 and theorem 2.3.

4 The mistake term

Theorem 4.1 *The following formula holds*

$$p_n = Li^{-1}(n) + O(h(n)) \qquad (n \geq 2)$$

where

$$h(n) = \frac{n \log^2 n}{exp(d\sqrt{\log n})}$$

Proof. We shall suppose x and n sufficiently large.

We shall choose in (16) the de la Vallee Poussin's function

$$f(x) = \frac{x}{exp(d\sqrt{\log x})}$$

From (18) we obtain that n belongs to the following interval of the y-axis

$$(Li(p_n) - Kf(p_n), Li(p_n) + Kf(p_n))$$

Therefore, since $y = Li(x)$ is increasing, $Li^{-1}(n)$ will belong to the following interval of the x-axis

$$\left(Li^{-1}\left(Li(p_n) - Kf(p_n)\right), Li^{-1}\left(Li(p_n) + Kf(p_n)\right)\right)$$

Note that the point $x = p_n$ belongs to this interval since $Li^{-1}(Li(p_n)) = p_n$.

If $n \leq Li(p_n)$ then $Li^{-1}(n) \leq Li^{-1}(Li(p_n)) = p_n$. Hence

$$p_n - Li^{-1}(n) < p_n - Li^{-1}\left(Li(p_n) - Kf(p_n)\right) \tag{31}$$

From the Lagrange's theorem applied to $Li^{-1}(y)$ we have (see (5))

$$\begin{aligned} p_n &- Li^{-1}\left(Li(p_n) - Kf(p_n)\right) = Li^{-1}(Li(p_n)) - Li^{-1}\left(Li(p_n) - Kf(p_n)\right) \\ &= \log a(Kf(p_n)) \leq \log p_n Kf(p_n) \end{aligned} \tag{32}$$

where $x = a$ satisfies

$$Li^{-1}\left(Li(p_n) - Kf(p_n)\right) < a < p_n$$

(31) and (32) give

$$p_n - Li^{-1}(n) < K \log p_n f(p_n) \tag{33}$$

On the other hand, if

$$n > Li(p_n) \tag{34}$$

then $Li^{-1}(n) > Li^{-1}(Li(p_n)) = p_n$. Hence

$$Li^{-1}(n) - p_n < Li^{-1}\left(Li(p_n) + Kf(p_n)\right) - p_n \tag{35}$$

From the Lagrange's theorem applied to $Li^{-1}(y)$ we have

$$\begin{aligned} Li^{-1}\left(Li(p_n)+Kf(p_n)\right)-p_n &= Li^{-1}\left(Li(p_n)+Kf(p_n)\right)-Li^{-1}(Li(p_n)) \\ &= \log a(Kf(p_n)) \end{aligned} \tag{36}$$

where $x=a$ satisfies

$$p_n < a < Li^{-1}\left(Li(p_n)+Kf(p_n)\right) \tag{37}$$

Let us consider the function $g(x)=Li(x)-Kf(x)<Li(x)$. We have $g(x)\sim Li(x)\sim(x/\log x)$. Therefore $g(x)\to\infty$. Its derivative is $g'(x)\sim(1/\log x)>0$. This imply that $g(x)$ is increasing from a certain value of x. Consequently, there exists $x=b$ such that $g(b)=Li(p_n)+Kf(p_n)$.

Now, since

$$\begin{aligned} g\left(Li^{-1}\left(Li(p_n)+Kf(p_n)\right)\right) &= Li(p_n)+Kf(p_n) \\ &- Kf\left(Li^{-1}\left(Li(p_n)+Kf(p_n)\right)\right) \\ &< Li(p_n)+Kf(p_n)=g(b) \end{aligned}$$

We have

$$Li^{-1}\left(Li(p_n)+Kf(p_n)\right)<b \tag{38}$$

(37) and (38) give

$$p_n<a<b \tag{39}$$

Since $g(x)$ is increasing, in the interval (p_n,b) of the x-axis we have

$$g(x)-g(b)<0 \tag{40}$$

On the other hand in the interval (b,∞) we have

$$g(x)-g(b)>0 \tag{41}$$

Now, we have (see (19))

$$f(p_{2n})=\frac{p_{2n}}{exp(d\sqrt{\log p_{2n}})}\sim\frac{2n\log n}{exp(d\sqrt{\log n})}$$

$$f(p_n)=\frac{p_n}{exp(d\sqrt{\log p_n})}\sim\frac{n\log n}{exp(d\sqrt{\log n})}$$

Therefore

$$(2n-2Kf(p_{2n}))-(n+Kf(p_n))\sim n$$

and consequently

$$(2n-2Kf(p_{2n}))-(n+Kf(p_n))>0 \tag{42}$$

Now (see (34))

$$Li(p_n) + Kf(p_n) < n + Kf(p_n) \tag{43}$$

On the other hand (see (18))

$$Li(p_{2n}) + Kf(p_{2n}) > 2n$$

That is

$$g(p_{2n}) = Li(p_{2n}) - Kf(p_{2n}) > 2n - 2Kf(p_{2n}) \tag{44}$$

(42), (43) and (44) give

$$g(p_{2n}) - (Li(p_n) + Kf(p_n)) = g(p_{2n}) - g(b) > 0 \tag{45}$$

(39), (40), (41) and (45) give

$$a < b < p_{2n}$$

That is

$$\log a < \log p_{2n} \tag{46}$$

(35), (36) and (46) give

$$Li^{-1}(n) - p_n < K \log p_{2n} f(p_n) \tag{47}$$

Now (see (19))

$$\log p_{2n} f(p_n) \sim \log p_n f(p_n) \sim \frac{n \log^2 n}{exp(d\sqrt{\log n})} \tag{48}$$

(33), (47) and (48) give

$$\left| p_n - Li^{-1}(n) \right| = O\left(\frac{n \log^2 n}{exp(d\sqrt{\log n})} \right)$$

The theorem is thus proved.

Remark 2. If the Riemann's hypothesis is true we obtain in the same way

$$h(n) = \sqrt{n} \log^{5/2} n$$

Since in this case $f(x) = \sqrt{x} \log x$ and consequently we have

$$\log p_n f(p_n) \sim \sqrt{n} \log^{5/2} n$$

If we use a function $f(x)$ better than the de la Vallee Poussin's function, in the same way we shall obtain a better function $h(n)$.

5 The aproximate formula for p_n

Theorem 5.1 *From a certain value of n the following formula holds*

$$\begin{aligned} p_n &= n\log n + \log(n\log n)(n - Li(n\log n)) \\ &+ \sum_{k=2}^{\infty} \frac{(-1)^k Q_{k-1}(\log(n\log n))}{k! n^{k-1}\log^{k-1} n}(n - Li(n\log n))^k + O(h(n)) \end{aligned}$$

where

$$h(n) = \frac{n\log^2 n}{exp(d\sqrt{\log n})}$$

Proof. It is an inmediate consequence of theorem 3.2 and theorem 4.1.

Theorem 5.2 *The series*

$$\begin{aligned} & n\log n + \log(n\log n)(n - Li(n\log n)) \\ + & \sum_{k=2}^{\infty} \frac{(-1)^k Q_{k-1}(\log(n\log n))}{k! n^{k-1}\log^{k-1} n}(n - Li(n\log n))^k \end{aligned} \tag{49}$$

converges absolutely if $n \geq 4$.

Proof. We have (see theorem 2.3)

$$\left|\frac{(-1)^k Q_{k-1}(\log(n\log n))}{k! n^{k-1}\log^{k-1} n}\right| \leq \frac{H_{k-1}(\log(n\log n))}{k! n^{k-1}\log^{k-1} n} \qquad (k \geq 2)$$

Let us consider the series

$$\begin{aligned} & n\log n + \log(n\log n)\,|n - Li(n\log n)| \\ + & \sum_{k=2}^{\infty} \frac{H_{k-1}(\log(n\log n))}{k! n^{k-1}\log^{k-1} n}\,|n - Li(n\log n)|^k \end{aligned} \tag{50}$$

We shall prove that the series (50) converges if $n \geq 4$. Therefore the series (49) will be absolutely convergent if $n \geq 4$.

Let us consider the quotient of a term with the former term in the series (50). Then if $k \geq 2$ we have

$$\begin{aligned} & \frac{H_k(\log(n\log n))\,|n - Li(n\log n)|^{k+1}}{(k+1)! n^k \log^k n} \frac{k! n^{k-1}\log^{k-1} n}{H_{k-1}(\log(n\log n))\,|n - Li(n\log n)|^k} \\ = & \frac{1}{(k+1)n\log n}\frac{H_k(\log(n\log n))}{H_{k-1}(\log(n\log n))}\,|n - Li(n\log n)| = \frac{1}{(k+1)n\log n} \\ & \frac{(k-1)\log(n\log n)H_{k-1}(\log(n\log n)) + \log(n\log n)H'_{k-1}(\log(n\log n))}{H_{k-1}(\log(n\log n))} \end{aligned}$$

$$|n - Li(n\log n)| = \frac{\log(n\log n)}{n\log n}\left(\frac{k-1}{k+1} + \frac{k-1}{k+1}\frac{1}{\log(n\log n)}\right.$$
$$\left.\frac{\frac{H'_{k-1}(\log(n\log n))}{k-1}}{\frac{H_{k-1}(\log(n\log n))}{\log(n\log n)}}\right)|n - Li(n\log n)| \leq \frac{|n - Li(n\log n)|\log(n\log n)}{n\log n}$$
$$\left(1 + \frac{1}{\log(n\log n)}\right) = \frac{|n - Li(n\log n)|}{n}\left(1 + \frac{\log\log n}{\log n} + \frac{1}{\log n}\right)$$

From the ratio test the series (50) will be convergent for all n such that

$$\frac{|n - Li(n\log n)|}{n}\left(1 + \frac{\log\log n}{\log n} + \frac{1}{\log n}\right) < 1 \tag{51}$$

Note that

$$\lim_{n\to\infty}\frac{|n - Li(n\log n)|}{n}\left(1 + \frac{\log\log n}{\log n} + \frac{1}{\log n}\right) = 0$$

Since $Li(x) \sim \frac{x}{\log x}$. Therefore (51) is true from a certain value of n. Here we shall prove (51) is true if $n \geq 4$.

First case. Suppose that for a certain n we have

$$n > Li(n\log n)$$

If we apply the Lagrange's theorem to the function $Li(x)$ in the interval $[2, n\log n]$ then we find that

$$\frac{n\log n - 2}{\log(n\log n)} < Li(n\log n) \qquad (n \geq 3) \tag{52}$$

(51) and (52) give $(n \geq 3)$

$$\begin{aligned} & \frac{|n - Li(n\log n)|}{n}\left(1 + \frac{\log\log n}{\log n} + \frac{1}{\log n}\right) \\ \leq\ & \frac{n - \frac{n\log n - 2}{\log(n\log n)}}{n}\left(1 + \frac{\log\log n}{\log n} + \frac{1}{\log n}\right) \\ =\ & \frac{n\log n\log\log n + n(\log\log n)^2 + n\log\log n + 2\log n + 2\log\log n + 2}{n\log^2 n + n\log n\log\log n} \\ =\ & f(n) \end{aligned}$$

Let us consider the function $f(x)$. We have

$$f'(x) = \frac{A(x)}{B(x)}$$

where

$$B(x) = \left(x\log^2 x + x\log x\log\log x\right)^2$$

and

$$\begin{aligned} & A(x) = (1-\log\log x)x\log^2 x + (1-2(\log\log x)^2)x\log x \\ - & \ 6\log x\log\log x - 4\log^2 x\log\log x - 2\log^3 x - 4\log^2 x - 4\log x \\ - & \ 2\log x(\log\log x)^2 - 2\log\log x - x(\log\log x)^3 - 2(\log\log x)^2 - 2 \end{aligned}$$

Therefore if $x \geq 20$ we have $A(x) < 0$ and consequently $f'(x) < 0$.

On the other hand, if $4 \leq n \leq 20$ we find (using a computer and the software Microsoft Excell) that $f(n) < 1$. Consequently if $n \geq 4$ we have $f(n) < 1$.

If we apply the Lagrange's theorem to the function $Li(x)$ in the intervals $[2,3]$, $[3,4]$, $[4,5]$ and $[5,6]$ we find that

$$Li(4\log 4) = \frac{3-2}{\log 2} + \frac{4-3}{\log 3} + \frac{5-4}{\log 4} + \frac{6-5}{\log 5} < 4$$

That is, if $n = 4$ we are in this case. Therefore the series (49) converges if $n = 4$.

Second case. Suppose that for a certain n we have

$$n < Li(n\log n)$$

If we apply the Lagrange's theorem to the function $Li(x)$ in the intervals $[2, n/2]$, $[n/2, n]$ and $[n, n\log n]$, we find that

$$\frac{1}{\log 2}\left(\frac{n}{2}-2\right)+\frac{1}{\log\left(\frac{n}{2}\right)}\left(n-\frac{n}{2}\right)+\frac{1}{\log n}(n\log n - n) > Li(n\log n) \qquad (n \geq 5)$$

That is

$$\frac{n}{2\log 2} - \frac{2}{\log 2} + \frac{n}{2(\log n - \log 2)} + n - \frac{n}{\log n} > Li(n\log n) \qquad (n \geq 5) \quad (53)$$

(51) and (53) give $(n \geq 5)$

$$\begin{aligned} & \frac{|n - Li(n\log n)|}{n}\left(1 + \frac{\log\log n}{\log n} + \frac{1}{\log n}\right) \\ \leq & \left(\frac{1}{2\log 2} - \frac{2}{n\log 2} + \frac{1}{2(\log n - \log 2)} - \frac{1}{\log n}\right)\left(1 + \frac{\log\log n}{\log n} + \frac{1}{\log n}\right) \\ = & \ f(n) \end{aligned}$$

Let us consider the function $f(x)$. We have

$$f'(x) = \frac{A(x)}{B(x)}$$

where

$$B(x) = (2\log 2)x^2\log^3 x(\log x - \log 2)^2$$

and

$$\begin{aligned} A(x) &= -x\log^3 x\log\log x + (\log 2)x\log^3 x + (4\log 2)x\log^2 x\log\log x \\ &- \log 2(4\log 2 - 1)x\log^2 x + 2(\log 2)^3 x - 8(\log 2)^2 x\log x\log\log x \\ &+ 4(\log 2)^3 x\log\log x + 4\log^4 x\log\log x - 2(\log 2)^2(2-\log 2)x\log x \\ &+ 4\log^5 x - 4(2\log 2 - 1)\log^4 x + 4(\log 2)^2\log^2 x \\ &- 4(2\log 2 - 1)\log^3 x\log\log x + 4(\log 2)^2\log x\log\log x \\ &- 4\log 2(2-\log 2)\log^3 x - 4\log 2(2-\log 2)\log^2 x\log\log x \end{aligned}$$

We shall prove that $A(x) < 0$ if $x \geq 50.000$ and consequently that $f'(x) < 0$ if $x \geq 50.000$.

It is sufficient to prove that (see $A(x)$):

$$-x\log^3 x\log\log x + (\log 2)x\log^3 x + (4\log 2)x\log^2 x\log\log x < 0$$

$$-\log 2(4\log 2 - 1)x\log^2 x + 2(\log 2)^3 x < 0$$

$$-8(\log 2)^2 x\log x\log\log x + 4(\log 2)^3 x\log\log x + 4\log^4 x\log\log x < 0$$

$$-2(\log 2)^2(2-\log 2)x\log x + 4\log^5 x < 0$$

$$-4(2\log 2 - 1)\log^4 x + 4(\log 2)^2\log^2 x < 0$$

$$-4(2\log 2 - 1)\log^3 x\log\log x + 4(\log 2)^2\log x\log\log x < 0$$

if $x \geq 50.000$. That is

$$-\log x\log\log x + (\log 2)\log x + (4\log 2)\log\log x < 0 \tag{54}$$

$$-(4\log 2 - 1)\log^2 x + 2(\log 2)^2 < 0 \tag{55}$$

$$-2(\log 2)^2 x\log x + (\log 2)^3 x + \log^4 x < 0 \tag{56}$$

$$-(\log 2)^2(2-\log 2)x + 2\log^4 x < 0 \tag{57}$$

$$-(2\log 2 - 1)\log^2 x + (\log 2)^2 < 0 \tag{58}$$

if $x \geq 50.000$.

Clearly, the inequalities (55) and (58) are true if $x \geq 50.000$.

Consider the inequality (54) and let us write

$$j(x) = -\log x\log\log x + (\log 2)\log x + (4\log 2)\log\log x$$

Then

$$j'(x) = \frac{1}{x}\left(-\log\log x - (1-\log 2) + \frac{4\log 2}{\log x}\right)$$

Now, the inequality

$$-(1-\log 2)+\frac{4\log 2}{\log x}<0$$

is true if

$$x\geq 50.000>exp\left(\frac{4\log 2}{1-\log 2}\right)$$

Therefore $j'(x)<0$ if $x\geq 50.000$.

On the other hand $j(50.000)<0$, consequently $j(x)<0$ if $x\geq 50.000$. That is, the inequality (54) is true if $x\geq 50.000$.

Clearly, if $x\geq 50.000$ the following inequality holds

$$\sqrt[4]{x}>\log x \tag{59}$$

Consider the inequality (57) and let us write

$$h(x)=-(\log 2)^2(2-\log 2)x+2\log^4 x$$

Then

$$h'(x)=-(\log 2)^2(2-\log 2)+\frac{8\log^3 x}{x}$$

Let us consider the inequality

$$-(\log 2)^2(2-\log 2)+\frac{8\log^3 x}{x}<0 \tag{60}$$

Substituting $\log x$ by $\sqrt[4]{x}$ we obtain the inequality

$$-(\log 2)^2(2-\log 2)+\frac{8}{\sqrt[4]{x}}<0 \tag{61}$$

This inequality is true if

$$x\geq 50.000>\left(\frac{8}{\log 2)^2(2-\log 2)}\right)^4$$

Therefore (59) imply that the inequality (60) is also true if $x\geq 50.000$.

That is, $h'(x)<0$ if $x\geq 50.000$.

On the other hand $h(50.000)<0$, consequently $h(x)<0$ if $x\geq 50.000$.

That is, the inequality (57) is true if $x\geq 50.000$.

Consider the inequality (56) and let us write

$$g(x)=-2(\log 2)^2x\log x+(\log 2)^3x+\log^4 x$$

Then

$$g'(x)=-2(\log 2)^2\log x-(2-\log 2)(\log 2)^2+\frac{4\log^3 x}{x}$$

Let us consider the inequality

$$-(2-\log 2)(\log 2)^2+\frac{4\log^3 x}{x}<0 \tag{62}$$

Substituting $\log x$ by $\sqrt[4]{x}$ we obtain the inequality

$$-(2-\log 2)(\log 2)^2+\frac{4}{\sqrt[4]{x}}<0 \tag{63}$$

This inequality is true if $x \geq 50.000$ (see (61)). Therefore (59) imply that the inequality (62) is also true if $x \geq 50.000$.

That is, $g'(x)<0$ if $x \geq 50.000$.

On the other hand $g(50.000)<0$, consequently $g(x)<0$ if $x \geq 50.000$.

That is, the inequality (56) is true if $x \geq 50.000$.

Therefore $f'(x)<0$ if $x \geq 50.000$.

On the other hand, if $5 \leq n \leq 50.000$ we find (using a computer) that $f(n)<1$.

Consequently $f(n)<1$ if $n \geq 5$.

The theorem is proved.

We now can establish the main theorem of this article.

Theorem 5.3 *If $n \geq 4$ the following formula holds*

$$\begin{aligned} p_n &= n\log n+\log(n\log n)(n-Li(n\log n)) \\ &+ \sum_{k=2}^{\infty}\frac{(-1)^k Q_{k-1}(\log(n\log n))}{k!n^{k-1}\log^{k-1} n}(n-Li(n\log n))^k+O(h(n)) \end{aligned}$$

where

$$h(n)=\frac{n\log^2 n}{exp(d\sqrt{\log n})}$$

Proof. It is an inmediate consequence of theorem 5.1 and theorem 5.2.

References

[1] Bromwich, T. J. I' A, *An introduction to the theory of infinite series,* second edition, Macmillan, London, 1942.

Received: December 12, 2007