arXiv:1203.5413v2 [math.NT] 5 Apr 2012

THE n-TH PRIME ASYMPTOTICALLY

JUAN ARIAS DE REYNA AND JÉRÉMY TOULISSE

Abstract. A new derivation of the classic asymptotic expansion of the n-th prime is presented. A fast algorithm for the computation of its terms is also given, which will be an improvement of that by Salvy (1994).

Realistic bounds for the error with $\operatorname{li}^{-1}(n)$, after having retained the first m terms, for $1 \le m \le 11$, are given. Finally, assuming the Riemann Hypothesis, we give estimations of the best possible r_3 such that, for $n \ge r_3$, we have $p_n > s_3(n)$ where $s_3(n)$ is the sum of the first four terms of the asymptotic expansion.

1. Introduction.

1.1. **Historical note.** Chebyshev failed to fully prove the Prime Number Theorem (PNT), but he obtained some notable approximations. For example, he proved that for every natural number n: if the limit

$$\lim_{x\to\infty} \frac{\log^n x}{x}\big(\pi(x) - \operatorname{li}(x)\big)$$

exists, then this limit must be equal to 0.

The question was decided by de la Vallée Poussin (1899) when he gave his bound on the error in the PNT: The above limits exist and equal 0.

In 1894, Pervushin, a priest in Perm, published several formulae obtained empirically about prime numbers[1]. One of these formulae gives the following approximation to the n-th prime

$$\frac{p_n}{n} = \log n + \log\log n - 1 + \frac{5}{12\log n} + \frac{1}{24\log^2 n}.$$

Cesàro then published a note [1, 1894] where he asserts that the true formula is

$$\frac{p_n}{n} = \log n + \log\log n - 1 + \frac{\log\log n - 2}{\log n} - \\ - \frac{(\log\log n)^2 - 6\log\log n + 11}{2\log^2 n} + o(\log^{-2} n).$$

Despite no mention by Cesàro in [1], the editors of his collected works added a note to [1] pointing out that certain formulae quoted by Cesàro, since they followed from the results of Chebyshev, were only established under the assumption of the existence of the implied limits. It therefore remains unsurprising that Hilbert, in the Jahrbuch[2] stated that Césaro did not prove his formula.

[1]Ivan Mikheevich Pervushin (1827-1900) (Иван Михеевич Первушин). No small achievement if we note that he had only a table of primes up to 3 000 000.

[2]*Jahrbuch über die Fortschritte der Mathematik (1868–1942)*, a forerunner for the *Zentralblatt für Mathematik*, at present digitalized at http://www.emis.de/MATH/JFM/JFM.html.

Landau [7, 1907] several years later was better informed: a formula, like that of Cesàro, would imply the PNT, which had yet to be proved at Cesàro's time. However, using the results of Chebyshev, Cesàro may claim that if there is some formula for p_n correct to the order $n(\log n)^{-2}$, then it must coincide with his formula.

Cipolla [3, 1902] obtained an infinite asymptotic expansion for p_n and gave a recursive formula to compute its terms. He published after the results of de la Vallée Poussin but it seems that he was unaware of these results, so that gave his proof under the same hypotheses as Cesáro. So uninformed was he that he attempted to prove some false formulae of Pervushin already *corrected* by Torelli [22]

$$p_{n+1}-p_n=\log n+\log\log n+\frac{\log\log n-1}{\log n}+\mathcal{O}\Bigl(\frac{\log\log n}{\log n}\Bigr)^2$$

with an *impeccable proof* that if such a formula exists, then it must be this formula. (Such a formula would refute the twin prime conjecture, and today the above formula is known to be false.)

In his *Handbuch* Landau [8, § 57] obtained by means of the procedure of Cesàro, some approximative formulae for p_n, and explained that the method could give further terms. He also mentioned some recursive formulae without giving any clue for their derivation.

We may say that Pervushin was the first to deal with a formula for p_n, albeit that he gave only the first few terms. Cesàro then proved that in the case such a formula exists, it must be one from which he would be able to derive several terms. Cipolla found a method to write all the terms of the expansion if there is one. Landau saw that the results of de la Vallée Poussin imply that the expansion certainly exists.

The algorithm given by Cipolla is not very convenient for the computation of the terms of the expansion. He iteratively computes the derivative of some polynomials appearing in the expansion but computes the constant terms as determinants of increasing order. Robin [17, 1988] considers the problem of computing these and other similar expansions, leaving the problem of computing the constant terms of the polynomials as an open problem. Later Salvy [20, 1994] gives a satisfactory algorithm. This algorithm needs $\mathcal{O}(n^{7/2}\sqrt{\log n})$ coefficient operations to compute all the polynomials up to the n-th polynomial.

The asymptotic expansion of p_n also plays a role in the study of $g(n)$, which is the maximum order of any element in the symmetric group S_n. In fact, $\log g(n)$ has the same asymptotic expansion as $\sqrt{\operatorname{li}^{-1}(x)}$ [11].

There are many results giving true bounds on p_n, for example we mention $p_n\ge n\log n$ [18, 1939], and $p_n\ge n(\log n+\log\log n-1)$ [4, 1999] both for $n\ge 2$ (with partial results given in [19], [16], [12], [4]). In [5] it is also proved that

$$p_n\le n\Bigl(\log n+\log\log n-1+\frac{\log\log n-2}{\log n}\Bigr),\qquad n\ge 688\,383.$$

1.2. **Organization of the paper.** In this paper we present a new derivation of the asymptotic expansion for p_n and obtain explicit bounds for the error.

First, it must be said that the asymptotic expansion has, in a certain sense, nothing to do with prime numbers: it is an asymptotic expansion of $\operatorname{ali}(x):=\operatorname{li}^{-1}x$ which is the inverse of the usual logarithmic integral function.

In Section 3 a proof of the existence of the expansion is given, following the path of Cesàro, since it cannot be found elsewhere, although it is frequently claimed it can be done. This Section is not needed in the rest of the paper.

In Section 4, a new formal derivation of the expansion is given. We obtain a new algorithm to compute the polynomials (Theorem 4.9). This is simpler than that given by Salvy [20]. Our algorithm allows all the polynomials up to the n-th one to be computed in $\boldsymbol{\mathcal{O}}(n^2)$ coefficient operations (Theorem 4.11). It must be said that these polynomials have $\boldsymbol{\mathcal{O}}(n^2)$ coefficients.

In Section 5, independently of Section 3, we prove that the formal expansion given in Section 4 is in fact the asymptotic expansion of $\operatorname{ali}(x)$ and gives realistic bounds on the error (Theorem 5.16 and 5.17).

In Section 6, the results are applied to p_n the n-th prime. Using the results of de la Vallée Poussin it can be shown that the asymptotic expansion of $\operatorname{ali}(n)$ is also an asymptotic expansion for p_n.

By assuming the Riemann Hypothesis, we found (Theorem 6.2) that

$$|p_n - \operatorname{ali}(n)| \leq \frac{1}{\pi}\sqrt{n}\,(\log n)^{\frac{5}{2}}, \qquad n \geq 11.$$

This bound of p_n is better than all the bounds cited above.

We end the paper by motivating why the above bounds have not been extended to further terms of the asymptotic expansion (Theorem 6.4).

Notations: With a certain hesitation we have introduced the notation $\operatorname{ali}(x)$ to denote the inverse function of $\operatorname{li}(x)$.

In Section 5, where explicit bounds are sought, it has been useful to denote by θ a real or complex number of absolute value $|\theta| \leq 1$, which will not always be the same, and depends on all parameters or variables in the corresponding equation.

Acknowledgement: The authors would like to thank Jan van de Lune (Hallum, The Netherlands) for his linguistic assistance in preparing the paper, and his interest in our results.

2. The inverse function of the logarithmic integral.

Usually $\operatorname{li}(x)$ is defined for real x as the principal value of the integral

$$\operatorname{li}(x) = \text{p. v.} \int_0^x \frac{dt}{\log t}.$$

It may be extended to an analytic function over the region $\Omega = \mathbb{C} \smallsetminus (-\infty, 1]$, which is the complex plane with a cut along the real axis $x \leq 1$. The main branch of the logarithm is defined in Ω and does not vanish there. Therefore, $\operatorname{li}(z)$ may be defined in Ω by

$$\operatorname{li}(z) = \operatorname{li}(2) + \int_2^z \frac{dt}{\log t}, \qquad z \in \Omega \tag{1}$$

where we integrate, for example, along the segment from 2 to z.

For real $x > 1$, the function $\operatorname{li}(x)$ is increasing and maps the interval $(1, +\infty)$ onto $(-\infty, +\infty)$, so that we may define the inverse function $\operatorname{ali}\colon \mathbb{R} \to (1, +\infty)$ by

$$\operatorname{li}(\operatorname{ali}(x)) = x. \tag{2}$$

The function $\mathrm{li}(x)$ is analytic on Ω, so that $\mathrm{ali}(x)$ is real analytic. It is clear that we have the following rules of differentiation

$$\frac{d}{dx}\,\mathrm{li}(x)=\frac{1}{\log x},\qquad \frac{d}{dx}\,\mathrm{ali}(x)=\log\mathrm{ali}(x). \tag{3}$$

It is well known that the function $\mathrm{li}(x)$ has an asymptotic expansion:

Theorem 2.1. *For each integer $N\ge 0$*

$$\mathrm{li}(x)=\frac{x}{\log x}\Bigl(1+\sum_{k=1}^{N}\frac{k!}{\log^k x}+\mathcal{O}\Bigl(\frac{1}{\log^{N+1}x}\Bigr)\Bigr),\qquad (x\to+\infty). \tag{4}$$

This may be proved by repeated integration by parts (see [13, p. 190–192]).

3. Asymptotic expansion of $\mathrm{ali}(x)$.

In this section, we prove the following

Theorem 3.1. *For each integer $N\ge 0$*

$$\frac{\mathrm{ali}(e^x)}{xe^x}=1+\sum_{n=1}^{N}\frac{P_{n-1}(\log x)}{x^n}+\mathcal{O}\Bigl(\frac{\log^{N+1}x}{x^{N+1}}\Bigr),\qquad (x\to+\infty) \tag{5}$$

where the $P_{n-1}(z)$ are polynomials of degree $\le n$.

In the case of $N=0$ the sum must be understood as equal to 0.

The theorem says that, for each N, there exists an $x_N>1$ and a constant C_N such that

$$\Bigl|\frac{\mathrm{ali}(e^x)}{xe^x}-1-\sum_{n=1}^{N}\frac{P_{n-1}(\log x)}{x^n}\Bigr|\le C_N\frac{\log^{N+1}x}{x^{N+1}},\qquad (x>x_N).$$

In the course of the proof we will make repeated use of the following

Lemma 3.2. *Let $f(x)$ be a function defined on a neighbourhood of $x=0$ such that*

$$f(x)=a_1x+\cdots+a_Nx^N+\mathcal{O}(x^{N+1}),\qquad (x\to 0) \tag{6}$$

where the a_k are given constants. Assume that $g(x)$ satisfies

$$g(x)=\sum_{n=1}^{N}\frac{p_n(\log x)}{x^n}+\mathcal{O}\Bigl(\frac{\log^{N+1}x}{x^{N+1}}\Bigr),\qquad (x\to+\infty) \tag{7}$$

where the $p_n(z)$ are polynomials of degree $\le n$. Then there exist polynomials $q_k(z)$ of degree $\le k$ such that

$$f(g(x))=\sum_{k=1}^{N}\frac{q_k(\log x)}{x^k}+\mathcal{O}\Bigl(\frac{\log^{N+1}x}{x^{N+1}}\Bigr),\qquad (x\to+\infty). \tag{8}$$

Proof. It is clear that, for each $1\le n\le N$ we have $p_n(\log x)x^{-n}=\mathcal{O}((\log x/x)^n)$. Therefore, $g(x)=\mathcal{O}(\log x/x)$ (this is true even when $N=0$ and there is no p_n). It follows that $\lim_{x\to+\infty}g(x)=0$ and by substitution in (6), we obtain

$$f(g(x))=\sum_{n=1}^{N}a_ng(x)^n+\mathcal{O}\Bigl(\frac{\log^{N+1}x}{x^{N+1}}\Bigr). \tag{9}$$

By expanding the powers $g(x)^n$ by (7) it is easy to obtain an expression of the form

$$g(x)^n = \sum_{k=1}^{N} \frac{p_{n,k}(\log x)}{x^k} + \mathcal{O}\Big(\frac{\log^{N+1} x}{x^{N+1}}\Big), \qquad (x \to +\infty) \tag{10}$$

where each $p_{n,k}(z)$ is a polynomial of degree $\leq k$. By substituting these values in equation (9) and collecting terms with the same power of x, (8) is obtained. □

We will prove Theorem 3.1 by induction. The following theorem yields the first step of this induction.

Theorem 3.3.

$$\frac{\operatorname{ali}(x)}{x \log x} = 1 + \mathcal{O}\Big(\frac{\log\log x}{\log x}\Big), \qquad (x \to +\infty). \tag{11}$$

Proof. From (4) with $N = 0$ we have $\frac{\operatorname{li}(y)\log y}{y} = 1 + \mathcal{O}(\log^{-1} y)$ for $y \to \infty$. Since $\lim_{x\to+\infty} \operatorname{ali}(x) = +\infty$ we may substitute $y = \operatorname{ali}(x)$ and obtain

$$\frac{x \log \operatorname{ali}(x)}{\operatorname{ali}(x)} = 1 + \mathcal{O}\Big(\frac{1}{\log \operatorname{ali}(x)}\Big). \tag{12}$$

By taking logarithms

$$\log x - \log \operatorname{ali}(x) + \log\log \operatorname{ali}(x) = \mathcal{O}\Big(\frac{1}{\log \operatorname{ali}(x)}\Big)$$

we obtain

$$\frac{\log x}{\log \operatorname{ali}(x)} = 1 + \mathcal{O}\Big(\frac{\log\log \operatorname{ali}(x)}{\log \operatorname{ali}(x)}\Big) \tag{13}$$

and it follows that

$$\lim_{x\to+\infty} \frac{\log x}{\log \operatorname{ali}(x)} = 1. \tag{14}$$

By taking log in (13)

$$\log\log x - \log\log \operatorname{ali}(x) = \mathcal{O}\Big(\frac{\log\log \operatorname{ali}(x)}{\log \operatorname{ali}(x)}\Big)$$

we obtain

$$\frac{\log\log x}{\log\log \operatorname{ali}(x)} = 1 + \mathcal{O}\Big(\frac{1}{\log \operatorname{ali}(x)}\Big)$$

so that

$$\lim_{x\to+\infty} \frac{\log\log x}{\log\log \operatorname{ali}(x)} = 1. \tag{15}$$

In view of (14) and (15), we may write (12) and (13) in the form

$$\frac{x \log \operatorname{ali}(x)}{\operatorname{ali}(x)} = 1 + \mathcal{O}\Big(\frac{1}{\log x}\Big), \qquad \frac{\log x}{\log \operatorname{ali}(x)} = 1 + \mathcal{O}\Big(\frac{\log\log x}{\log x}\Big) \tag{16}$$

and by multiplying these two, we obtain

$$\frac{x \log x}{\operatorname{ali}(x)} = 1 + \mathcal{O}\Big(\frac{\log\log x}{\log x}\Big)$$

from which (11) can easily be deduced. □

Proof of Theorem 3.1. We proceed by induction. For $N=0$, our theorem is simply Theorem 3.3 with e^x instead of x.

Hence we assume (5) and try to prove the case $N+1$.

Our objective will be obtained by starting from the expansion of $\operatorname{li}(y)$. By (4)

$$\operatorname{li}(y)=\frac{y}{\log y}\Big(1+\sum_{k=1}^{N+1}\frac{k!}{\log^k y}+\mathcal{O}\Big(\frac{1}{\log^{N+2} y}\Big)\Big).$$

By substituting $y=\operatorname{ali}(e^x)$ and applying (14) we obtain

$$\frac{e^x\log\operatorname{ali}(e^x)}{\operatorname{ali}(e^x)}=1+\sum_{k=1}^{N+1}\frac{k!}{(\log\operatorname{ali}(e^x))^k}+\mathcal{O}\Big(\frac{1}{x^{N+2}}\Big).\tag{17}$$

From our induction hypothesis, the expansion of $\log\operatorname{ali}(x)$ and $(\log\operatorname{ali}(x))^{-k}$ is now sought.

By taking the log of (5) we obtain

$$\log\operatorname{ali}(e^x)=x+\log x+\log\Big\{1+\sum_{n=1}^{N}\frac{P_{n-1}(\log x)}{x^n}+\mathcal{O}\Big(\frac{\log^{N+1}x}{x^{N+1}}\Big)\Big\}.$$

Lemma 3.2 may be applied with

$$\log(1+X)=X-\frac{X^2}{2}+\frac{X^3}{3}-\cdots(-1)^{N+1}\frac{X^N}{N}+\mathcal{O}(X^{N+1})$$

to obtain

$$\log\operatorname{ali}(e^x)=x+\log x+\sum_{n=1}^{N}\frac{Q_{n+1}(\log x)}{x^n}+\mathcal{O}\Big(\frac{\log^{N+1}x}{x^{N+1}}\Big).$$

The reason why we have written Q_{n+1} instead of $Q_n(x)$ is revealed below. The above may be written as

$$\log\operatorname{ali}(e^x)=x\Big\{1+\frac{\log x}{x}+\sum_{n=1}^{N}\frac{Q_{n+1}(\log x)}{x^{n+1}}+\mathcal{O}\Big(\frac{\log^{N+1}x}{x^{N+2}}\Big)\Big\}$$

or

$$\log\operatorname{ali}(e^x)=x\Big\{1+\sum_{n=1}^{N+1}\frac{Q_n(\log x)}{x^n}+\mathcal{O}\Big(\frac{\log^{N+2}x}{x^{N+2}}\Big)\Big\}\tag{18}$$

where the $Q_n(z)$ are polynomials of degree $\le n$. Observe that knowing the expansion of $\operatorname{ali}(e^x)$ up to $(\log x/x)^{N+1}$ has enabled us to obtain $\log\operatorname{ali}(e^x)$ up to $(\log x/x)^{N+2}$; this will be of great importance in what follows.

From (18), for all natural numbers n,

$$\frac{1}{(\log\operatorname{ali}(e^x))^n}=\frac{1}{x^n}\Big\{1+\sum_{k=1}^{N+1}\frac{Q_k(\log x)}{x^k}+\mathcal{O}\Big(\frac{\log^{N+2}x}{x^{N+2}}\Big)\Big\}^{-n}.$$

By applying Lemma 3.2 with

$$(1+x)^{-n}-1=\sum_{r=1}^{N+1}\binom{-n}{r}x^r+\mathcal{O}(x^{-N-2})$$

we obtain

$$\frac{1}{\{\log \operatorname{ali}(e^x)\}^n} = \frac{1}{x^n}\Bigl(1+\sum_{k=1}^{N+1}\frac{V_{n,k}(\log x)}{x^k}+\mathcal{O}\Bigl(\frac{\log^{N+2}x}{x^{N+2}}\Bigr)\Bigr) \tag{19}$$

where the $V_{n,k}(z)$ are polynomials of degree $\le k$.

By substituting these values of $\{\log \operatorname{ali}(e^x)\}^{-n}$ in (17), we obtain

$$\frac{e^x\log \operatorname{ali}(e^x)}{\operatorname{ali}(e^x)} = 1+\sum_{k=1}^{N+1}\frac{U_k(\log x)}{x^k}+\mathcal{O}\Bigl(\frac{\log^{N+2}x}{x^{N+2}}\Bigr).$$

Hence again from (19) with $n=1$

$$\frac{e^x}{\operatorname{ali}(e^x)} = \frac{1}{x}\Bigl\{1+\sum_{k=1}^{N+1}\frac{V_{1,k}(\log x)}{x^k}+\mathcal{O}\Bigl(\frac{\log^{N+2}x}{x^{N+2}}\Bigr)\Bigr\}\times \\ \times\Bigl\{1+\sum_{k=1}^{N+1}\frac{U_k(\log x)}{x^k}+\mathcal{O}\Bigl(\frac{\log^{N+2}x}{x^{N+2}}\Bigr)\Bigr\}$$

from which we derive that there exist polynomials $W_k(z)$ of degree $\le k$ such that

$$\frac{xe^x}{\operatorname{ali}(e^x)} = 1+\sum_{k=1}^{N+1}\frac{W_k(\log x)}{x^k}+\mathcal{O}\Bigl(\frac{\log^{N+2}x}{x^{N+2}}\Bigr).$$

Another application of Lemma 3.2 yields

$$\frac{\operatorname{ali}(e^x)}{xe^x} = 1+\sum_{k=1}^{N+1}\frac{T_k(\log x)}{x^k}+\mathcal{O}\Bigl(\frac{\log^{N+2}x}{x^{N+2}}\Bigr)$$

with polynomials $T_k(z)$ of degree $\le k$. Therefore, we have an asymptotic expansion of type (5) with $N+1$ instead of N. The usual argument of uniqueness of the asymptotic expansion applies here so that $T_k(z)=P_k(z)$ for $1\le k\le N$. □

4. Formal Asymptotic expansion.

First we give some motivation. We have seen that the asymptotic expansion of $\operatorname{ali}(e^x)$ is

$$\operatorname{ali}(e^x) = xe^xV(x,\log x),\quad \text{where}\quad V(x,y) := 1+\sum_{n=1}^{\infty}\frac{P_{n-1}(y)}{x^n}$$

and differentiation yields

$$e^x\log \operatorname{ali}(e^x) = (e^x+xe^x)V+xe^xV_x+e^xV_y.$$

Here $\log \operatorname{ali}(e^x) = \log\bigl(xe^xV(x,\log x)\bigr) = y+x+\log V$, so that

$$y+x+\log V = V+xV+xV_x+V_y$$

which we write as

$$V = 1+\frac{y}{x}-\frac{1}{x}V-V_x-\frac{1}{x}V_y+\frac{1}{x}\log V. \tag{20}$$

This ends our motivation for considering this equation.

Consider now the ring A of the formal power series of the type

$$\sum_{n=0}^{\infty}\frac{q_n(y)}{x^n}$$

where the $q_n(y)$ are polynomials with complex coefficients of degree less than or equal to n. In particular $q_0(y)$ is a constant.

It is clear that A, with the obvious operations, is a ring. The elements with $q_0=0$ form a maximal ideal I. An element $1+u$ with $q_0=1$ is invertible, with inverse $1-u+u^2-\cdots$. It follows that if $a\notin I$, then a is also invertible. Hence I is the unique maximal ideal and A is a local ring. If $a\in A$ is a non-vanishing element, then there exists a least natural number n with $q_n(y)\neq 0$. We define $\deg(a)=n$ in this case, with $\deg(0)=\infty$.

As is usual in local rings, (see [9]) we may define a topology induced by the norm $\|a\|=2^{-\deg(a)}$, which, with the associated metric, induces a complete metric space. Indeed A is isomorphic to $\mathbb{C}[[X,Y]]$, by means of the application that sends $X\mapsto x^{-1}$, $Y\mapsto yx^{-1}$.

Given $a\in A$ with $a=\sum_{n=0}^{\infty}\frac{q_n(y)}{x^n}$, we define two derivates

$$a_x=-\sum_{n=1}^{\infty}\frac{nq_n(y)}{x^{n+1}}\quad\text{and}\quad a_y=\sum_{n=1}^{\infty}\frac{q_n'(y)}{x^n}.$$

Finally the set $U\subset A$ of elements with $q_0=1$ form a multiplicative subgroup of A^* (the group of invertible elements of A). For $1+u\in U$, we define

$$\log(1+u)=\sum_{k=1}^{\infty}(-1)^{k+1}\frac{u^k}{k}$$

which is a series that is easily shown to converge since $u^k\in I^k$.

We are now ready to prove the following

Theorem 4.1. *The equation* (20) *has one and only one solution in the ring A.*

Proof. For $V\in U$, we define $T(V)$ as

$$T(V):=1+\frac{y}{x}-\frac{1}{x}V-V_x-\frac{1}{x}V_y+\frac{1}{x}\log V.$$

It is clear that $T(V)\in U$. We may apply Banach's fixed-point theorem. Indeed, we have $\deg(T(V)-T(W))\le 1+\deg(V-W)$, so that $\|T(V)-T(W)\|\le\frac{1}{2}\|V-W\|$.

By Banach's theorem there is a unique solution to $V=T(V)$. We may obtain this solution as the limit of the sequence $T^n(1)$. In fact, since $\deg(T(V)-T(W))\le 1+\deg(V-W)$, in each iteration we obtain one further term of the expansion. In this way, it is easy to prove that the solution is

$$V=1+\frac{y-1}{x}+\frac{y-2}{x^2}+\cdots$$

However, we are going to find more direct methods to compute the terms of the expansion. □

Definition 4.2. Let V be the unique solution to equation (20). Since it is in A, it has the form

$$V(x,y)=1+\sum_{n=1}^{\infty}\frac{P_{n-1}(y)}{x^n}\tag{21}$$

where for $n \geq 0$, $P_{n-1}(y)$ is a polynomial of degree $\leq n$.

In the following sections, we prove that V yields the asymptotic expansion of $\mathrm{ali}(e^x)$. For this proof the following property is crucial.

Theorem 4.3. *For $N \geq 1$, let*

$$W(x,y) = W_N(x,y) := 1 + \sum_{k=1}^{N} \frac{P_{n-1}(y)}{x^n}. \tag{22}$$

Then

$$W - 1 - \frac{y}{x} + \frac{1}{x}W + W_x + \frac{1}{x}W_y - \frac{1}{x}\log W = -\frac{P_N(y)}{x^{N+1}} + \frac{u(y)}{x^{N+2}} + \frac{v(y)}{x^{N+3}} + \cdots \tag{23}$$

Proof. By the definition of the P_n we know that $\deg(V - W) \geq N+1$. Therefore, $\deg(V - T(W)) \geq N+2$. That is

$$V - T(W) = \frac{u_0(y)}{x^{N+2}} + \frac{v_0(y)}{x^{N+3}} + \cdots$$

where u_0, v_0 are polynomials. We also have

$$V = W + \sum_{n=N+1}^{\infty} \frac{P_{n-1}(y)}{x^n}$$

so that

$$W - T(W) = \frac{u_0(y)}{x^{N+2}} + \frac{v_0(y)}{x^{N+3}} + \cdots - \sum_{n=N+1}^{\infty} \frac{P_{n-1}}{x^n}.$$

That is,

$$W - T(W) = -\frac{P_N(y)}{x^{N+1}} + \frac{u(y)}{x^{N+2}} + \frac{v(y)}{x^{N+3}} + \cdots$$

for certain polynomials u, v, … □

In the sequel V will denote the unique solution to (20). The element $\log V$ belongs to A, so that there are polynomials $Q_n(y)$ of degree less than or equal to n such that

$$\log V = \sum_{n=1}^{\infty} \frac{Q_n(y)}{x^n}. \tag{24}$$

From equation (20), we may obtain $\log V$ in terms of V and its derivatives. It is easy to obtain from this expression the following relation

$$Q_n(y) = P_n(y) - (n-1)P_{n-1}(y) + P'_{n-1}(y), \qquad (n \geq 1). \tag{25}$$

Theorem 4.4. *The polynomials $P_n(y)$ that appear in the unique solution* (21) *to equation* (20) *may be computed by the following recurrence relations:*

$$P_0 = y - 1, \quad \textit{and for } n \geq 1$$

$$P_n = nP_{n-1} - P'_{n-1} + \frac{1}{n}\sum_{k=1}^{n-1} k\big\{(k-1)P_{k-1} - P_k - P'_{k-1}\big\}P_{n-k-1}. \tag{26}$$

Proof. By differentiating (24) with respect to x, we obtain

$$\Big(\sum_{n=1}^\infty \frac{nQ_n(y)}{x^{n+1}}\Big)\Big(1+\sum_{n=1}^\infty \frac{P_{n-1}(y)}{x^n}\Big) = \sum_{n=1}^\infty \frac{nP_{n-1}(y)}{x^{n+1}}.$$

By equating the coefficients of x^{-n-1}, we obtain

$$nP_{n-1} = nQ_n + \sum_{k=1}^{n-1} kQ_kP_{n-k-1}, \qquad (n\ge 2). \tag{27}$$

Now we substitute the values of the Q_n given in (25)

$$nP_{n-1} = nP_n - n(n-1)P_{n-1} + nP'_{n-1} + \sum_{k=1}^{n-1} kQ_kP_{n-k-1}$$

so that

$$nP_n = n^2P_{n-1} - nP'_{n-1} - \sum_{k=1}^{n-1} k\big\{P_k-(k-1)P_{k-1}+P'_{k-1}\big\}P_{n-k-1}.$$

□

From this expression it is very easy to compute the first terms of the expansions

$$V = 1+\frac{y-1}{x}+\frac{y-2}{x^2}-\frac{y^2-6y+11}{2x^3}+\frac{2y^3-21y^2+84y-131}{6x^4}- \\ -\frac{3y^4-46y^3+294y^2-954y+1333}{12x^5}+\cdots,$$

$$\log V = \frac{y-1}{x}-\frac{y^2-4y+5}{2x^2}+\frac{2y^3-15y^2+42y-47}{6x^3}- \\ -\frac{3y^4-34y^3+156y^2-366y+379}{12x^4}+\cdots$$

Theorem 4.5. *We have*

(a) *For $n\ge 1$, the degree of P_n is less than or equal to n.*
(b) *$n!\,P_n(y)$ has integer coefficients.*

Proof. The equation (20) may be written

$$V-1-\frac{y}{x} = \frac{1}{x}\big(\log V - V - xV_x - V_y\big).$$

Since $xV_x\in A$, it is clear that

$$xV-x-y = -1+\sum_{n=1}^\infty \frac{P_n(y)}{x^n}\in A.$$

This implies that the degree of P_n is less than or equal to n.

We prove (b) by induction. The first few P_n satisfy this property. We define $p_k := k!P_k$ so that the recurrence relation (26) may be written as

$$p_n = n^2p_{n-1}-np'_{n-1}+ \\ +(n-1)\sum_{k=1}^{n-1}\binom{n-2}{k-1}\big\{k(k-1)p_{k-1}-p_k-kp'_{k-1}\big\}p_{n-k-1}.$$

Hence, by induction, all p_n have integer coefficients. □

The most significant contribution by Cipolla is his proof of a recurrence for the coefficients $a_{n,k}$ of P_n (see (30)), which is better than the recurrence given in (26). We intend to give a slightly different proof. The result of Cipolla is equivalent to the following surprising fact: The solution V of equation (20) formally satisfies the following linear partial differential equation:

$$V = (x-1)V_y - xV_x. \tag{28}$$

This equation can easily be deduced from the following Theorem.

Theorem 4.6. *For $n \geq 1$, we have*

$$\begin{aligned} (n-1)P_{n-1}(y) &= P'_{n-1}(y) - P'_n(y), \qquad (n \geq 1) \\ (n-1)Q_{n-1}(y) &= Q'_{n-1}(y) - Q'_n(y), \qquad (n \geq 2). \end{aligned} \tag{29}$$

Proof. We will proceed by induction. For $n \leq 3$ it can be verified that these equalities are satisfied.

We now assume that (29) is satisfied for $n \leq N$, and we will show that these equations are true for $n = N+1$.

By differentiating (24) with respect to y we get

$$\Big(\sum_{n=1}^{\infty} \frac{Q'_n(y)}{x^n}\Big)\Big(1 + \sum_{n=1}^{\infty} \frac{P_{n-1}(y)}{x^n}\Big) = \sum_{n=1}^{\infty} \frac{P'_{n-1}(y)}{x^n}$$

so that by equating the coefficients of x^{-N-1} and of x^{-N} we obtain

$$Q'_{N+1} = P'_N - \sum_{k=0}^{N-1} P_k Q'_{N-k}, \qquad Q'_N = P'_{N-1} - \sum_{k=0}^{N-2} P_k Q'_{N-k-1}.$$

Subtracting these equations we get

$$Q'_{N+1} - Q'_N = P'_N - P'_{N-1} - \sum_{k=0}^{N-2} P_k(Q'_{N-k} - Q'_{N-k-1}) - P_{N-1}$$

and by the induction hypothesis this is equal to

$$-(N-1)P_{N-1} + \sum_{k=0}^{N-2} P_k \cdot (N-k-1)Q_{N-k-1} - P_{N-1} = $$

$$= -NP_{N-1} + \sum_{k=1}^{N-1} kQ_k P_{N-k-1}.$$

By (27) this is equal to $NP_{N-1} - NQ_N$ so that we obtain

$$Q'_{N+1} - Q'_N = -NQ_N.$$

This is the second equation of (29) for $n = N+1$. In order to achieve the result for the first equation, observe that from (25) we get

$$\begin{aligned} NQ_N &= NP_N - N(N-1)P_{N-1} + NP'_{N-1} \\ -Q'_N &= -P'_N + (N-1)P'_{N-1} - P''_{N-1} \\ Q'_{N+1} &= P'_{N+1} - NP'_N + P''_N. \end{aligned}$$

By adding these equations we obtain

$$0 = NP_N - P'_N + P'_{N+1} + N\bigl\{P'_{N-1} - P'_N - (N-1)P_{N-1}\bigr\} - \\ - \bigl\{P''_{N-1} - P''_N - (N-1)P'_{N-1}\bigr\} = NP_N - P'_N + P'_{N+1}$$

which is the first equation of (29) for $n = N + 1$. □

We define the coefficients $a_{n,k}$ implicitly by

$$P_n(y) = \frac{(-1)^{n+1}}{n!}\Bigl(a_{n,0}y^n - a_{n,1}y^{n-1} + \cdots (-1)^n a_{n,n}\Bigr) = \\ = \frac{(-1)^{n+1}}{n!}\sum_{k=0}^{n}(-1)^k\, a_{n,k}\, y^{n-k}, \qquad (n \geq 1). \tag{30}$$

Analogously, Q_n is of a degree less than or equal to n, and we define the coefficients $b_{n,k}$ implicitly by

$$Q_n(y) = \frac{(-1)^{n+1}}{n!}\sum_{k=0}^{n}(-1)^k\, b_{n,k}\, y^{n-k}, \qquad (n \geq 1). \tag{31}$$

Remark 4.7. $P_0(y)$ has degree 1, which is not given by (30). However, we can extend the definition of $a(n,k)$ in such a way that, for $n \geq 1$ we have $a(n,k) = 0$ for $k < 0$ or $k > n$. Then a formula such as (30) also holds for $n = 0$ if we add up the values from $k = -1$ to $k = n$ and define $a(0,0) = 1$, $a(0,-1) = 1$ and $a(0,k) = 0$ for other values of k

$$P_n(y) = \frac{(-1)^{n+1}}{n!}\sum_{k=-1}^{n}(-1)^k\, a_{n,k}\, y^{n-k}, \qquad (n \geq 0). \tag{30 bis}$$

Note that $Q_0(y)$ remains undefined.

Theorem 4.8. *For $1 \leq n$ and $0 \leq k < n$, we have (when defined)*

$$a_{n,k} = na_{n-1,k-1} + \frac{n(n-1)}{n-k}a_{n-1,k}, \quad b_{n,k} = nb_{n-1,k-1} + \frac{n(n-1)}{n-k}b_{n-1,k}. \tag{32}$$

For $1 \leq n$ and $0 \leq k \leq n$, we have

$$a_{n,k} = b_{n,k} + (n-k+1)a_{n,k-1}. \tag{33}$$

For $n \geq 1$, we have

$$b_{n,n} = na_{n-1,n-1} + \sum_{k=1}^{n-1}\binom{n-1}{k}\, k\, b_{k,k}\, a_{n-k-1,n-k-1}. \tag{34}$$

Proof. (32) is obtained by equating the coefficients of y^{n-k-1} in the first equation in (29). In this way, we obtain

$$(n-1)\frac{(-1)^{n+k}}{(n-1)!}a_{n-1,k} = \\ = (n-k)\frac{(-1)^{n+k+1}}{(n-1)!}a_{n-1,k-1} - (n-k)\frac{(-1)^{n+k+1}}{n!}a_{n,k}.$$

If $n \neq k$, then the equation for $a_{n,k}$ in (32) is obtained. The other equation in $b_{n,k}$ is obtained analogously from the second equation in (29).

To prove (33), observe that by (25), $Q_n = P_n - (n-1)P_{n-1} + P'_{n-1}$, and from (29) it follows that

$$Q_n = P_n + P'_n, \qquad (n \geq 1). \tag{35}$$

Now by equating the coefficient of y^{n-k} in both members of this equality we obtain (33).

Finally (34) follows from (27). Recall that $-\frac{a_{n,n}}{n!}$ and $-\frac{b_{n,n}}{n!}$ are respectively the values of $P_n(0)$, and $Q_n(0)$. Hence, by setting $y = 0$ in (27), we obtain (34) through multiplication by $(n-1)!$ and the reordering of the terms. □

The main problem now is that equations (32) do not allow us to compute the coefficients $a_{n,n}$. Cipolla gives an algorithm to simultaneously compute the coefficients $a_{n,k}$ and $b_{n,k}$ based on Theorem 4.8. In the procedure of Cipolla, these key coefficients $a_{n,n}$ are recursively computed using all the previous coefficients. We prefer a method that computes $A_n := a_{n,n}$ and $B_n := a_{n,n-1}$ separately and then compute the remaining coefficients by using (32).

Theorem 4.9. *In order to compute the numbers $a_{n,k}$, we may first compute the sequences $A_n := a_{n,n}$ and $B_n := a_{n,n-1}$ by the recursions*

$$A_0 = 1, \quad A_1 = 2, \quad B_0 = 1, \quad B_1 = 1, \tag{36}$$

$$B_n = nB_{n-1} + n(n-1)A_{n-1} \tag{37}$$

$$A_n = n^2 A_{n-1} + nB_{n-1} - (n-1)\sum_{k=1}^{n-1} \binom{n-2}{k-1} \big\{k(k-1)A_{k-1} - A_k + kB_{k-1}\big\} A_{n-k-1}. \tag{38}$$

After this one we may obtain $a(n,k) := a_{n,k}$. Setting

$$a(0,0) = 1, \quad a(0,-1) = 1, \quad a(1,0) = 1, \quad a(1,1) = 2$$

and all other $a(0,k)$ and $a(1,k) = 0$. Then, for $n \geq 2$, put

$$a(n,n) = A_n,$$

$$a(n,k) = na(n-1,k-1) + \frac{n(n-1)}{n-k} a(n-1,k), \quad (0 \leq k < n) \tag{39}$$

where $a(n,k) = 0$ for $k < 0$ or $k > n$.

Finally, we may obtain the $b(n,k) := b_{n,k}$ from

$$b(n,k) = a(n,k) - (n-k+1)a(n,k-1). \tag{40}$$

Proof. The constant term of P_n is $-\frac{A_n}{n!}$ and the coefficient of y in P_n is $\frac{B_n}{n!}$, so that equation (37) follows from the first equation in (29) taking it with $y = 0$.

In the same way, (38) follows from (26), and (39) is the first equation in (32).

Equation (40) for the $b(n,k)$ follows easily from (35). □

The array of coefficients $a(n,k)$ for $0 \leq n, k \leq 7$, reads

1	0	0	0	0	0	0	0
1	2	0	0	0	0	0	0
1	6	11	0	0	0	0	0
2	21	84	131	0	0	0	0
6	92	588	1908	2666	0	0	0
24	490	4380	22020	62860	81534	0	0
120	3084	35790	246480	1075020	2823180	3478014	0
720	22428	322224	2838570	16775640	66811920	165838848	196993194

and the $b(n,k)$ for $1 \le n \le 7$ and $0 \le k \le 7$ are

1	1	0	0	0	0	0	0
1	4	5	0	0	0	0	0
2	15	42	47	0	0	0	0
6	68	312	732	758	0	0	0
24	370	2420	8880	18820	18674	0	0
120	2364	20370	103320	335580	673140	654834	0
720	17388	187656	1227450	5421360	16485000	32215008	31154346

Theorem 4.10. (a) *The coefficients $b(n,k)$ are integers.*
(b) *$a(n,k) \ge 0$ and $b(n,k) \ge 0$.*
(c) *$a(n,k-1) \le a(n,k)$ for $1 \le k \le n$.*
(d) *For $n \ge 1$, $a(n,0) = (n-1)!$.*

Proof. (a) We have proved in Theorem 4.5 that the numbers $a(n,k)$ are integers, so that from (40), the coefficients $b(n,k)$ are also integers.

(b) We proceed by induction on n. Assuming that we have proved that $a(m,k)$ and $b(m,k)$ are positive for $m < n$, it follows from (32) that $a(n,k)$ and $b(n,k)$ are positive for $0 \le k < n$. Then (34) implies that $b(n,n) \ge 0$, and (33) with $k = n$ proves that $a(n,n) \ge 0$.

(c) This is a simple consequence of (33).

(d) The equation follows from (39) by induction. □

Theorem 4.11. *By means of the rule in Theorem 4.9, one may compute all coefficients $a_{n,k}$ of the polynomials $P_n(y)$ for $1 \le n \le N$ in $\mathcal{O}(N^2)$ coefficient operations.*

Proof. We count the operations needed, following the indications in Theorem 4.9, to compute every $a_{n,k}$ for $0 \le n \le N$ and $0 \le k \le n$.

First we must compute the numbers $\binom{m}{j}$ for $0 \le m \le N-2$. Using the scheme of the usual triangle, we need to carry out $\sum_{k=1}^{N-3} k$ additions, which involves $(N-2)(N-3)/2$ operations.

The numbers B_n must now be computed for $2 \le n \le N$ by means of the formula

$$B_n = n * (B_{n-1} + (n-1) * A_{n-1}).$$

Each B_n requires 4 operations, therefore a total of $4(N-1)$ operations are needed. We compute the A_n for $2 \le n \le N$ using the formula

$$A_n = n * n * A_{n-1} + n * B_{n-1} - (n-1) * \\ * \sum_{k=1}^{n-1} \binom{n-2}{k-1} * \{k * (k-1) * A_{k-1} - A_k + k * B_{k-1}\} * A_{n-k-1}.$$

Hence A_n requires $7+\sum_{k=1}^{n-1} 8 = 8n-1$ operations. All A_n together take $\sum_{n=2}^{N}(8n-1) = 4N^2+3N-7$ operations. These numbers are the $a_{n,n}$. The $a_{0,k}$ and $a_{1,k}$ require no operations. Finally we compute for $0 \le k < n$

$$a_{n,k} = n * \bigl\{a_{n-1,k-1} + (n-1) * a_{n-1,k}/(n-k)\bigr\}.$$

Therefore, $a_{n,k}$ takes 6 operations. For each n, every $a_{n,k}$ for $1 \le k < n$ takes $6(n-1)$ operations. And each $a_{n,k}$ for $2 \le n \le N$ takes $\sum_{n=2}^{N} 6(n-1) = 3N(N-1)$.

The total cost in number of operations is therefore

$$\frac{(N-2)(N-3)}{2} + 4(N-1) + 4N^2 + 3N - 7 + 3N(N-1) = \\ = \frac{1}{2}(15N^2+3N-16).$$

□

5. Bounds for the asymptotic expansion.

5.1. **The sequence (a_n).** First we define a sequence of numbers as the coefficients of a formal expansion in A.

Lemma 5.1. *There exists a sequence of integers (a_n) such that*

$$\log\Bigl(1-\sum_{n=1}^{\infty}\frac{n!}{x^n}\Bigr)^{-1} = \sum_{n=1}^{\infty}\frac{a_n}{n}\frac{1}{x^n}. \tag{41}$$

The coefficients may be computed by the recursion

$$a_1 = 1, \quad a_n = n!\cdot n + \sum_{k=1}^{n-1} k!\, a_{n-k}. \tag{42}$$

Proof. It is clear that $u = 1-\sum n! x^{-n} \in U \subset A$, so that $u^{-1} \in U$ and $\log u^{-1}$ are well defined. To obtain the recursion we differentiate (41) to obtain

$$-\sum_{n=1}^{\infty}\frac{n\cdot n!}{x^{n+1}} = -\Bigl(\sum_{n=1}^{\infty}\frac{a_n}{x^{n+1}}\Bigr)\Bigl(1-\sum_{n=1}^{\infty}\frac{n!}{x^n}\Bigr).$$

Equation (42) is obtained by equating the coefficients of x^{-n-1}. The recurrence (42) proves that a_n is a natural number for each $n \ge 1$. □

The first terms of the sequence $(a_n)_{n=1}^{\infty}$ are

$$1, 5, 25, 137, 841, 5825, 45529, 399713, 3911785, 42302225, \dots$$

Lemma 5.2. *For each natural number n we have*

$$a_n \le 2n \cdot n!. \tag{43}$$

Proof. We may verify this property for a_1, a_2, a_3 and a_4 directly. For $n \ge 4$ we proceed by induction. Assume the inequality for a_k with $k < n$, so that by (42)

$$1 \le \frac{a_n}{n!\cdot n} \le 1 + \sum_{k=1}^{n-1}\frac{a_{n-k}}{(n-k)!\cdot(n-k)}\frac{n-k}{n}\binom{n}{k}^{-1} \le \\ \le 1 + 2\Bigl(\frac{1}{n} + \sum_{k=2}^{n-2}\binom{n}{k}^{-1} + \frac{1}{n^2}\Bigr) \le 1 + 2\Bigl(\frac{1}{n} + \frac{1}{n^2} + (n-3)\frac{2}{n(n-1)}\Bigr).$$

For $n \geq 4$, it is easy to see that this is ≤ 2. □

Lemma 5.3. *For each natural number N there is a positive constant c_N such that*

$$x\Big(1-\sum_{n=1}^{N}\frac{n!}{x^n}\Big)\geq 1,\qquad x\geq c_N. \tag{44}$$

Proof. It is clear that the left-hand side of (44) is increasing and tends to $+\infty$ when $x\to+\infty$, from which the existence of c_N is clear.

The value of c_N may be determined as the solution of the equation

$$x\Big(1-\sum_{n=1}^{N}\frac{n!}{x^n}\Big)=1,\qquad x>1. \tag{45}$$

In this way we found the following values.

c_1	2	c_6	4.15213	c_{11}	5.61664	c_{20}	8.70335
c_2	2.73205	c_7	4.43119	c_{12}	5.93649	c_{30}	12.34925
c_3	3.20701	c_8	4.71412	c_{13}	6.26449	c_{40}	16.03475
c_4	3.56383	c_9	5.00517	c_{14}	6.59947	c_{50}	19.72833
c_5	3.86841	c_{10}	5.30597	c_{15}	6.94035	c_{60}	23.42351

□

Remark 5.4. Notice that for $x\geq c_N$ the sum in (44) is positive and less than 1.

Proposition 5.5. *For each natural number N there exists $d_N>0$ such that, for $x\in\mathbb{C}$ with $|x|\geq d_N$, there exists θ with $|\theta|\leq 1$ such that*

$$\log\Big(1-\sum_{n=1}^{N}\frac{n!}{x^n}\Big)^{-1}=\sum_{n=1}^{N}\frac{a_n}{n}\frac{1}{x^n}+\theta\frac{a_{N+1}}{N+1}\frac{1}{x^{N+1}},\qquad |x|>d_N. \tag{46}$$

Proof. By comparing the expansions (41) and

$$\log\Big(1-\sum_{n=1}^{N}\frac{n!}{x^n}\Big)^{-1}=\sum_{n=1}^{N}\frac{a_n}{n}\frac{1}{x^n}+\sum_{n=N+1}^{\infty}\frac{b_n}{n}\frac{1}{x^n} \tag{47}$$

it is clear that $\frac{b_{N+1}}{N+1}+(N+1)!=\frac{a_{N+1}}{N+1}$, so that $b_{N+1}<a_{N+1}$.

The above expansion is convergent for all sufficiently large $|x|$, so that

$$\sum_{n=N+1}^{\infty}\frac{b_n}{n}\frac{1}{x^n}=\frac{b_{N+1}}{N+1}\frac{1}{x^{N+1}}g_N(x)$$

where $\lim_{x\to\infty}g_N(x)=1$. Hence there exist sufficiently large d_N such that

$$|b_{N+1}g_N(x)|<a_{N+1},\qquad |x|>d_N.$$

This ends the proof of the existence of d_N.

We have

$$\Big(\log\Big(1-\sum_{n=1}^{N}\frac{n!}{x^n}\Big)^{-1}-\sum_{n=1}^{N}\frac{a_n}{n}\frac{1}{x^n}\Big)\frac{(N+1)x^{N+1}}{a_{N+1}}=$$

$$=\frac{(N+1)}{a_{N+1}}\sum_{n=N+1}^{\infty}\frac{b_n}{n}\frac{1}{x^{n-N-1}}.$$

Since all a_n and b_n are positive, this is a decreasing function for $x\to+\infty$, and the lowest value of d_N will be the unique solution of

$$\Bigl(\log\Bigl(1-\sum_{n=1}^N\frac{n!}{x^n}\Bigr)^{-1}-\sum_{n=1}^N\frac{a_n}{n}\frac{1}{x^n}\Bigr)\frac{(N+1)x^{N+1}}{a_{N+1}}=1.$$

We obtain the following table of values

d_1	1.03922	d_6	4.54145	d_{11}	5.73661	d_{20}	8.73298
d_2	2.38568	d_7	4.75734	d_{12}	6.03061	d_{30}	12.37349
d_3	3.33232	d_8	4.97336	d_{13}	6.33969	d_{40}	16.05983
d_4	3.92171	d_9	5.20626	d_{14}	6.66091	d_{50}	19.75448
d_5	4.28707	d_{10}	5.46090	d_{15}	6.99175	d_{60}	23.45053

□

Remark 5.6. The numbers d_N in Lemma 5.5 are very similar to the numbers c_N of Lemma 5.3. This is no more than an experimental observation, but since the c_N numbers are easy to compute and d_N are somewhat elusive, it has been useful to start from c_N as an approximation to d_N in order to compute d_N.

5.2. **Some inequalities.**

Lemma 5.7. *For $u\ge2$ we have $\log\operatorname{ali}(u)\le2\log u$. For $u\ge e^2$ we have $\operatorname{ali}(u)\le 2u\log u$.*

Proof. The first inequality is equivalent to $\operatorname{ali}(u)\le u^2$. Since $\operatorname{li}(x)$ is strictly increasing, the inequality is equivalent to $u\le\operatorname{li}(u^2)$.

For $u>2$ we have $\operatorname{li}(u)>\operatorname{li}(2)=1.04516\dots$ so that

$$\operatorname{li}(u^2)=\operatorname{li}(u)+\int_u^{u^2}\frac{dt}{\log t}>1+\frac{u^2-u}{\log u^2}.$$

Hence, our inequality follows from $\frac{u^2-u}{\log u^2}>u-1$, that is from $u>2\log u$. However, this last inequality is certainly true for $u>2$.

The second inequality is equivalent to $u\le\operatorname{li}(2u\log u)$ and has a similar easy proof. □

Lemma 5.8. *For all integers $n\ge1$ we have*

$$\int_{e^{f_n}}^u\frac{(\log\log t)^n}{\log^{n+1}t}\,dt\le4u\frac{(\log\log u)^n}{\log^{n+1}u},\qquad(u\ge e^{f_n}) \tag{48}$$

where $f_n=4(n+1)/3$.

Proof. Notice that $f_n>1$. For $t\ge e$ the function $\log\log t$ is positive and increasing so that

$$\int_{e^{f_n}}^u\frac{(\log\log t)^n}{\log^{n+1}t}\,dt\le(\log\log u)^n\int_{e^{f_n}}^u\frac{dt}{\log^{n+1}t}.$$

It remains to be shown that

$$\int_{e^{f_n}}^u\frac{dt}{\log^{n+1}t}\le\frac{4u}{\log^{n+1}u},\qquad(u\ge e^{f_n}).$$

Replacing u by e^x this is equivalent to

$$\int_{f_n}^x\frac{e^t}{t^{n+1}}\,dt\le\frac{4e^x}{x^{n+1}},\qquad(x\ge f_n).$$

For the function

$$G(x) := \frac{4e^x}{x^{n+1}} - \int_{f_n}^x \frac{e^t}{t^{n+1}}\,dt$$

we have

$$G'(x) = \frac{e^x}{x^{n+1}}\Big(4 - \frac{4(n+1)}{x} - 1\Big)$$

so that for $x > 4(n+1)/3$ we obtain $G'(x) > 0$. Since $G(f_n) > 0$ we have $G(x) > 0$ for all $x > f_n$. $\square$

Theorem 5.9. *The polynomials $P_n(y)$ defined in* (21) *satisy the inequalities*

$$|P_n(y)| \le 3 \cdot n!\, y^n, \qquad y \ge 2, \quad n \ge 1 \tag{49}$$

and $|P_0(y)| \le y$ for $y \ge 2$.

Proof. Since $P_0(y) = y - 1$, the second assertion is trivial.

Given $r > 0$, for each polynomial $P(x) = \sum_{n=0}^N a_n x^n$ we define

$$\|P\| = \sum_{n=0}^N |a_n| r^n.$$

It is easy to show that

$$\|P + Q\| \le \|P\| + \|Q\|, \qquad \|PQ\| \le \|P\| \cdot \|Q\|$$

and that for the derivative of a polynomial of degree $\le N$

$$\|P'\| = \sum_{n=0}^N n|a_n|r^{n-1} \le \frac{N}{r}\sum_{n=0}^N |a_n| r^n = \frac{N}{r}\|P\|.$$

For $y \ge r$ we have the inequality

$$|P(y)| = \Big|\sum_{n=0}^N a_n y^n\Big| \le \sum_{n=0}^N |a_n| y^n \le \sum_{n=0}^N |a_n| r^n (y/r)^n \le (y/r)^N \|P\|.$$

Hence, our Theorem follows if it can be shown that for $n \ge 1$ we have $\|P_n\| \le 3 \cdot 2^n\, n!$ (for $r = 2$).

Define $S_n := \|P_n\|$. By (30) we have $S_n = -P_n(-2)$, and it can be shown that $S_n \le 3 \cdot 2^n\, n!$ for $0 \le n \le 15$.

For $n > 15$ it follows from (26) and the aforementioned properties of $\|P\|$ that

$$S_n \le nS_{n-1} + \frac{n}{2}S_{n-1} + \frac{1}{n}\sum_{k=1}^{n-1} k\Big((k-1)S_{k-1} + S_k + \frac{k}{2}S_{k-1}\Big)S_{n-k-1}.$$

It follows that $S_n \le T_n$ where $T_n := S_n \le 3 \cdot 3^n\, n!$ for $0 \le n \le 15$, and that for $n > 15$

$$T_n := \frac{3n}{2}T_{n-1} + \frac{1}{n}\sum_{k=1}^{n-1}\Big(kT_k + \frac{k(3k-2)}{2}T_{k-1}\Big)T_{n-k-1}.$$

Now we proceed by induction. For $n > 15$ and assuming that we have proved $T_k \le 3\cdot 2^k\, k!$ for $k < n$, we obtain

$$T_n \le \frac{9n}{2}2^{n-1}(n-1)!+ \\ +\frac{9}{n}\sum_{k=1}^{n-1}\Big(k2^k k! + \frac{k(3k-2)}{2}2^{k-1}(k-1)!\Big)2^{n-k-1}(n-k-1)!.$$

Hence

$$\frac{T_n}{3\cdot 2^n\, n!} \le \frac{3}{4}+\frac{3}{n}\sum_{k=1}^{n-1}\Big(\frac{k\cdot k!(n-k-1)!}{2\cdot n!}+\frac{(3k-2)k!(n-k-1)!}{8\cdot n!}\Big) \le \\ \le \frac{3}{4}+\frac{3}{8n^2}\sum_{k=1}^{n-1}\frac{7k-2}{\binom{n-1}{k}} \le \frac{3}{4}+\frac{3(7n-9)}{8n^2}+\frac{3}{8n^2}\sum_{k=1}^{n-2}\frac{7k-2}{\binom{n-1}{k}}.$$

Therefore, by using the symmetry of the combinatorial numbers, we obtain

$$\frac{T_n}{3\cdot 2^n\, n!} \le \frac{3}{4}+\frac{3(7n-9)}{8n^2}+\frac{3}{16n^2}\sum_{k=1}^{n-2}\frac{7n-11}{\binom{n-1}{k}} \le \\ \le \frac{3}{4}+\frac{3(7n-9)}{8n^2}+\frac{3}{16n^2}\cdot(n-2)\frac{7n-11}{n-1} \le \\ \le \frac{3}{4}+\frac{3(7n-9)}{8n^2}+\frac{3}{16n^2}\cdot(7n-11) = 1-\frac{n(4n-63)+87}{16n^2} < 1$$

for $n > 15$. □

Corollary 5.10. *We have*

$$|P_{n-1}(y)| \le n!\, y^n, \qquad n \ge 1, \quad y \ge 2. \tag{50}$$

Proof. This follows easily from the above Theorem. □

5.3. **Main inequalities.** To simplify our formulae we introduce some notation. First we set $r_n := 3\cdot n!$ so that, for $n \ge 1$, we have $|P_n(y)| \le r_n y^n$ when $y > 2$.

Let c_n and d_n be the constants introduced in Lemma 5.3 and Proposition 5.5. Let α_n be equal to $\max(e, c_n, d_n)$ and let $\beta_n \ge e$ be the solution of the equation

$$\frac{x}{\log x} = \alpha_n. \tag{51}$$

(The function $\frac{t}{\log t}$ is increasing for $t \ge e$).

Finally, define $x_n := \max(\beta_n, f_n, e^2)$, where f_n is defined in Lemma 5.8.

Proposition 5.11. *Let x be a real number such that $x \ge x_n$, and set $y := \log x$. Then*

$$y \ge 2, \qquad x \ge c_n y, \qquad x \ge d_n y, \qquad x \ge f_n.$$

Proof. Since $x \ge x_n = \max(\beta_n, f_n, e^2)$ we have $x \ge e^2$, so that $y = \log x \ge 2$.

We also have $x \ge \beta_n \ge e$. Since $\frac{t}{\log t}$ is an increasing function for $t \ge e$ we obtain $\frac{x}{\log x} \ge \frac{\beta_n}{\log \beta_n} = \alpha_n = \max(e, c_n, d_n)$. Therefore, $\frac{x}{y} \ge c_n$ and $\frac{x}{y} \ge d_n$ as required. □

We insert a table of the constants x_n.

x_1	7.38906	x_6	10.81135	x_{11}	16.00000	x_{20}	29.57923
x_2	7.38906	x_7	11.70187	x_{12}	17.33333	x_{30}	47.86556
x_3	7.38906	x_8	12.60164	x_{13}	18.66667	x_{40}	67.69154
x_4	8.29874	x_9	13.58167	x_{14}	20.00000	x_{50}	88.57644
x_5	9.77283	x_{10}	14.66667	x_{15}	21.42740	x_{60}	110.29065

For each natural number N we set

$$W_N = 1 + \sum_{n=1}^{N} \frac{P_{n-1}(y)}{x^n} \tag{52}$$

and frequently we write $W := W_N$ when N is fixed.

Proposition 5.12. *For $N \geq 1$ let $W = W_N$ (as in (52)). Then for $x \geq x_N$ and $y = \log x$ there exists θ with $|\theta| \leq 1$ such that*

$$W + xW + xW_x + W_y - x - y - \log W = \theta \cdot r_{N+1} \frac{y^N}{x^N}. \tag{53}$$

Proof. Denote by $T = T(x, y)$ the value of $W + xW + xW_x + W_y - x - y - \log W$. Then we have

$$T = (1 + x) \sum_{n=0}^{N} \frac{P_{n-1}(y)}{x^n} - \sum_{n=1}^{N} \frac{nP_{n-1}(y)}{x^n} + \sum_{n=1}^{N} \frac{P'_{n-1}(y)}{x^n} - \\ - x - y + \log\Big(1 + \sum_{n=1}^{N} \frac{P_{n-1}(y)}{x^n}\Big)^{-1}.$$

From (24) we have the expansion

$$\log\Big(1 + \sum_{n=1}^{\infty} \frac{P_{n-1}(y)}{x^n}\Big)^{-1} = -\sum_{n=1}^{\infty} \frac{Q_n(y)}{x^n}. \tag{54}$$

From Proposition 5.11 we know that $y = \log x > 2$ and $x \geq yd_N$. From (50), for $y > 2$, we have $|P_{n-1}(y)| \leq n! y^n$ so that we have the majorant

$$\log\Big(1 + \sum_{n=1}^{N} \frac{P_{n-1}(y)}{x^n}\Big)^{-1} \ll \log\Big(1 - \sum_{n=1}^{N} \frac{n!}{(x/y)^n}\Big)^{-1} \tag{55}$$

(by considering this expression as a power series in x^{-1}, and y as a parameter). From (54) and (55), we obtain

$$\log\Big(1 + \sum_{n=1}^{N} \frac{P_n(y)}{x^n}\Big)^{-1} = -\sum_{n=1}^{N} \frac{Q_n(y)}{x^n} + S_N(x, y) \tag{56}$$

where $S_N(x, y)$ is a power series majorized by the Taylor expansion of

$$\log\Big(1 - \sum_{n=1}^{N} \frac{n!}{(x/y)^n}\Big)^{-1} - \sum_{n=1}^{N} \frac{a_n}{n} \frac{1}{(x/y)^n}$$

(compare equation (47)).

By applying Proposition 5.5 we deduce that, for $x > yd_N$, there exists θ with $|\theta| \le 1$ and

$$S_N(x,y) = \theta \frac{a_{N+1}}{N+1}\frac{y^{N+1}}{x^{N+1}}. \tag{57}$$

If we substitute (56) in the expression for T, then by Theorem 4.3, all the terms in x^{-n} with $n < N$ cancel out, and the terms in x^{-N} add up to $-P_N(y)x^{-N}$. It follows that

$$T = -\frac{P_N(y)}{x^N} + S_N(x,y). \tag{58}$$

Therefore, since $y > 2$, we have

$$|T| \le r_N \frac{y^N}{x^N} + \frac{a_{N+1}}{N+1}\frac{y^{N+1}}{x^{N+1}}$$

so that from (43) ,

$$|T| \le \frac{y^N}{x^N}\Big(3\cdot N! + 2\cdot(N+1)!\frac{\log x}{x}\Big) \le \frac{3\cdot(N+1)!y^N}{x^N} = r_{N+1}\frac{y^N}{x^N}$$

where $N \ge 1$ and $\frac{3}{2} + \frac{2\log x}{x} \le 3$ for $x \ge e^2$ are applied. □

Proposition 5.13. *For each natural number N let $u_N = e^{x_N}$. Then there exists $v_N > u_N$ such that*

$$\operatorname{li}(f_N(u)) - u = \theta \cdot 13(N+1)!\,\frac{u(\log\log u)^N}{\log^{N+1} u}, \qquad (u > v_N) \tag{59}$$

where $|\theta| \le 1$ and

$$f_N(e^x) := xe^x W_N(x, \log x) = xe^x\Big(1 + \sum_{n=1}^{N}\frac{P_{n-1}(\log x)}{x^n}\Big). \tag{60}$$

Proof. To simplify the notation, the abbreviation $W(x,y) = W_N(x,y)$ is used. Differentiating (60) we obtain

$$\frac{d}{dx}\big(\operatorname{li}(f_N(e^x)) - e^x\big) =$$
$$= \frac{1}{\log(f_N(e^x))}\Big\{e^x W + xe^x W + xe^x\Big(W_x + \frac{1}{x}W_y\Big)\Big\} - e^x.$$

Assume that $x \ge x_N$, so that $x \ge d_N \log x$ and $x \ge e^2$. We may apply (53) to obtain

$$\frac{d}{dx}\big(\operatorname{li}(f_N(e^x)) - e^x\big) = \frac{e^x}{\log(f_N(e^x))}\Big\{W + xW + xW_x + W_y\Big\} - e^x =$$
$$= \frac{e^x}{\log(f_N(e^x))}\Big\{x + \log x + \log W + \theta r_{N+1}\frac{\log^N x}{x^N}\Big\} - e^x.$$

This may be simplified to

$$\frac{d}{dx}\big(\operatorname{li}(f_N(e^x)) - e^x\big) = \frac{e^x}{\log(f_N(e^x))}\cdot\theta\frac{r_{N+1}\log^N x}{x^N}.$$

Since $x > x_N$ we have $x \geq yc_N$, so that by Lemma 5.3

$$\Big|x\Big(1+\sum_{n=1}^N \frac{P_{n-1}(\log x)}{x^n}\Big)\Big| \geq x\Big(1-\sum_{n=1}^N \frac{n!}{(x/y)^n}\Big) \geq 1$$

that is $xW_N(x, \log x) \geq 1$, so that $\log(f_N(e^x)) \geq x$. Hence, for $x \geq x_N$ (with another θ), we have

$$\frac{d}{dx}\big(\operatorname{li}(f_N(e^x)) - e^x\big) = \theta \frac{r_{N+1}\log^N x}{x^{N+1}} e^x.$$

Defining $H_N(u) := \operatorname{li}(f_N(u)) - u$ the above equation is equivalent to

$$H_N'(e^x) = \theta \frac{r_{N+1}\log^N x}{x^{N+1}}, \qquad (x \geq x_N)$$

and, since $u_N := e^{x_N}$,

$$H_N'(u) = \theta \frac{r_{N+1}(\log\log u)^N}{\log^{N+1} u}, \qquad (u \geq u_N).$$

Lemma 5.8 can be applied since $x_N \geq f_N$, so that $u \geq u_N \geq e^{f_N}$. Hence, integrating over the interval (u_N, u) we get

$$H_N(u) = H_N(u_N) + \theta \frac{4r_{N+1}\, u(\log\log u)^N}{\log^{N+1} u}, \qquad (u \geq u_N).$$

The function $(N+1)!\frac{u}{\log^{N+1} u}\cdot(\log\log u)^N$ is increasing (as product of two positive increasing functions) for $u > e^{f_n}$, so that there exists $v_N > u_N$ for which this function is greater than $H_N(u_N)$, so that

$$H_N(u) = \theta \frac{13\cdot(N+1)!\, u(\log\log u)^N}{\log^{N+1} u}, \qquad (u \geq v_N).$$

□

Remark 5.14. For the values of n appearing in our tables, the equality $u_n = v_n$ holds, since, in these cases,

$$H_n(u_n) \leq \frac{(n+1)!\, u_n(\log\log u_n)^n}{\log^{n+1} u_n}.$$

Lemma 5.15. *For any natural number N, and $u > e^{x_N}$ we have $\log f_N(u) < 2\log u$.*

Proof. First observe that the hypothesis $u > e^{x_N}$ implies (with $u = e^x$) that $x > x_N$, so that $\log x > 2$ and $x > c_N \log x$. (Proposition 5.11).

The inequality $\log f_N(u) < 2\log u$ is equivalent to $f_N(u) < u^2$, and together with $u = e^x$ it is equivalent to

$$xe^x\Big(1+\sum_{n=1}^N \frac{P_{n-1}(\log x)}{x^n}\Big) < e^{2x}.$$

From Corollary 5.10, since $x \geq e^2$ and $x \geq c_N y$, and by Remark 5.14,

$$x\Big(1+\sum_{n=1}^N \frac{P_{n-1}(\log x)}{x^n}\Big) \leq x\Big(1+\sum_{n=1}^N \frac{n!}{(x/y)^n}\Big) < x\Big(1-\sum_{n=1}^N \frac{n!}{(x/y)^n}\Big)^{-1}.$$

Hence our inequality follows from

$$x < \frac{ye^x}{x}\cdot\frac{x}{y}\Big(1-\sum_{n=1}^N\frac{n!}{(x/y)^n}\Big), \qquad y=\log x.$$

Since we assume that $x\ge yc_N$, the second factor is greater than 1, so that

$$\frac{ye^x}{x}\cdot\frac{x}{y}\Big(1-\sum_{n=1}^N\frac{n!}{(x/y)^n}\Big) > \frac{ye^x}{x}$$

Finally, it is easy to prove that $e^x\log x > x^2$ for $x>e^2$. □

The asymptotic expansion with bounds can now be proved.

Theorem 5.16. *For each integer $N\ge 1$*

$$\operatorname{ali}(u) = f_N(u) + 26\theta(N+1)!\,u\Big(\frac{\log\log u}{\log u}\Big)^N, \qquad (u\ge v_N), \tag{61}$$

where v_N is the number defined in Proposition 5.13 .

Proof. Since $\operatorname{li}(\operatorname{ali}(u)) = u$, Proposition 5.13 yields, for $u>v_N$,

$$\operatorname{li}(f_N(u)) - \operatorname{li}(\operatorname{ali}(u)) = \int_{\operatorname{ali}(u)}^{f_N(u)}\frac{dt}{\log t} = 13\theta(N+1)!\,u\frac{(\log\log u)^N}{\log^{N+1}u}.$$

Since $v_N\ge u_N = e^{x_N}$, $u\ge v_N$ implies $\log u\ge 2$, hence $u\ge 2$.

From Lemma 5.7, $\log\operatorname{ali}(u)\le 2\log u$, for $u>2$. Analogously, Lemma 5.15 implies that $\log f_N(u)\le 2\log u$, for $u>e^{x_N}$. Therefore, for $u>v_N$, we have

$$\frac{|\operatorname{ali}(u)-f_N(u)|}{2\log u}\le\Big|\int_{\operatorname{ali}(u)}^{f_N(u)}\frac{dt}{\log t}\Big|.$$

It follows that there exists θ' with $|\theta'|\le 1$ such that

$$\operatorname{ali}(u)-f_N(u) = \theta'(2\log u)\int_{\operatorname{ali}(u)}^{f_N(u)}\frac{dt}{\log t}$$

and the result follows easily. □

The actual error appears to be much smaller than that given in Theorem 5.16. However, as usual with asymptotic expansions, having a true bound allows realistic bounds to be given of the remainder for specific values of N.

The true error after N terms of an asymptotic expansion, while the terms are decreasing in magnitude, is often of the size of the first omitted term. In our case, the magnitude of the term $P_N(\log x)x^{-N-1}$ depends on the polynomial $P_N(\log x)$.

Numerically, it appears that for $n\ge 3$:

$$|P_n(y)|\le\Big(\frac{n}{e\log n}\Big)^n y^n, \qquad (y>2\log n) \tag{62}$$

although we have not been able to prove this.

From Theorem 5.16, more realistic bounds can be obtained for the first values of N. This is done in the following Theorem.

Theorem 5.17. *For $2 \le N \le 11$, we have*

$$\frac{\operatorname{ali}(e^x)}{xe^x} = 1 + \sum_{n=1}^{N} \frac{P_{n-1}(\log x)}{x^n} + \theta \cdot 20 \cdot \Big(\frac{N}{e \log N}\Big)^N \cdot \frac{\log^N x}{x^{N+1}}, \qquad (x > z_N) \tag{63}$$

where

$$z_2 = 1.50, \quad z_3 = 2.34, \quad z_4 = 3.32, \quad z_5 = 4.33, \quad z_6 = 5.36,$$
$$z_7 = 6.39, \quad z_8 = 7.43, \quad z_9 = 8.46, \quad z_{10} = 9.50, \quad z_{11} = 10.53.$$

Proof. By taking $N = 10$ in Theorem 5.16, we have, for $u = e^x > e^{x_{10}}$, (recall also Remark **??**)

$$\frac{\operatorname{ali}(e^x)}{xe^x} = 1 + \sum_{n=1}^{10} \frac{P_{n-1}(\log x)}{x^n} + \theta R \frac{\log^{10} x}{x^{11}}, \qquad (x > x_{10})$$

with $R = 26 \cdot 11! = 1\,037\,836\,800$.

We compute the maximum[3] M_n of $|P_{n-1}(\log x)/\log^{n-1} x|$ for $x > x_{10}$, so that for any $2 \le N \le 10$, we have

$$\frac{\operatorname{ali}(e^x)}{xe^x} = 1 + \sum_{n=1}^{N} \frac{P_{n-1}(\log x)}{x^n} + \\ + \frac{\log^N x}{x^{N+1}} \Big(\sum_{n=N+1}^{10} \frac{P_{n-1}(\log x)}{\log^{n-1} x} \frac{\log^{n-N-1} x}{x^{n-N-1}} + \theta R \frac{\log^{10-N} x}{x^{10-N}} \Big)$$

so that

$$\frac{\operatorname{ali}(e^x)}{xe^x} = 1 + \sum_{n=1}^{N} \frac{P_{n-1}(\log x)}{x^n} + \\ + \theta \frac{\log^N x}{x^{N+1}} \Big(\sum_{n=N+1}^{10} \frac{M_n \log^{n-N-1} x}{x^{n-N-1}} + R \frac{\log^{10-N} x}{x^{10-N}} \Big).$$

We determine a value $z'_N > x_{10}$ such that, for $x = z'_N$,

$$\Big(\sum_{n=N+1}^{10} \frac{M_n \log^{n-N-1} x}{x^{n-N-1}} + R \frac{\log^{10-N} x}{x^{10-N}} \Big) < 20 \Big(\frac{N}{e \log N} \Big)^N.$$

Since this is a decreasing function of x, we obtain for $x > z'_N$

$$\frac{\operatorname{ali}(e^x)}{xe^x} = 1 + \sum_{n=1}^{N} \frac{P_{n-1}(\log x)}{x^n} + \theta \cdot 20 \Big(\frac{N}{e \log N}\Big)^N \frac{\log^N x}{x^{N+1}}. \tag{64}$$

We consider the function

$$\Big(\frac{\operatorname{ali}(e^x)}{xe^x} - 1 - \sum_{n=1}^{N} \frac{P_{n-1}(\log x)}{x^n} \Big) \frac{x^{N+1}}{\log^N x}$$

on the interval $(1.3, z'_N)$, to determine the least value of z_N for which (64) is true.

[3] $M_2 = 1$, $M_3 = 1/2$, $M_4 = 1/3$, $M_5 = 0.250636$, $M_6 = 0.526887$, $M_7 = 1.300565$, $M_8 = 3.719653$, $M_9 = 12.070813$, $M_{10} = 43.788782$. This last maximum would be much smaller if the maximum were taken from a point slightly greater than x_{10}.

In this way we find: $z_2' = 32$ and then $z_2 = 1.5$; $z_3' = 49.5$ and then $z_3 = 2.3395$; $z_4' = 82$ and then $z_4 = 3.3114$; $z_5' = 155$ and then $z_5 = 4.3237$.

If we take $N = 20$ in Theorem 5.16, we obtain $z_6' = 113$, $z_7' = 143$, $z_8' = 187$, $z_9' = 251$ $z_{10}' = 353$, $z_{11}' = 528$ from which $z_6 = 5.3514$, $z_7 = 6.3851$, $z_8 = 7.4208$, $z_9 = 8.4566$, $z_{10} = 9.4914$ and $z_{11} = 10.5251$ are obtained. □

Remark 5.18. We have proved (63) only for $2 \leq N \leq 11$, although something similar appears to be true for the general case. If (63) were true for all n, then for a large $u = e^x$ we could take $N \approx x$ terms in the expansion and in this way the error would be $\approx \frac{20}{xe^x}$, so that $\operatorname{ali}(u)$ could be computed with an error less than ≈ 20.

In fact, for several values of u, the terms of the expansion have been computed up to the point where these terms start to increase. Always the computation is terminated when $N \approx x$ and the error appears to be bounded. (For example, with $u = 10^{100}$, we compute 230 terms of the expansion, which coincides with $\log 10^{100} \approx 230.259$. The approximate value obtained for $\operatorname{ali}(u)$ has an absolute error equal to 40.94738, which can be compared with the fact that $\operatorname{ali}(10^{100})$ has 103 digits).

6. Applications to p_n.

6.1. **Asymptotic expansion of p_n.** Inequalities for the n-th prime number can be found in [18], [16], [12], [4]. In fact, from $\pi(x) = \operatorname{li}(x) + \mathcal{O}(r(x))$ we may obtain $p_n = \operatorname{ali}(n) + \mathcal{O}(r(n\log n)\log n)$, if $r(x)/x$ is sufficiently small. For example, in [12], it is noticed that from a result of Massias [10], it follows that

$$p_n = \operatorname{ali}(n) + \mathcal{O}(ne^{-c\sqrt{\log n}}) \tag{65}$$

so that the asymptotic expansion of $\operatorname{ali}(n)$ is also an asymptotic expansion for p_n, that is,

$$p_n = n\log n\Big(1 + \sum_{k=1}^{N} \frac{P_{k-1}(\log\log n)}{\log^k n}\Big) + \mathcal{O}\Big(n\Big(\frac{\log\log n}{\log n}\Big)^N\Big). \tag{66}$$

By assuming the Riemann hypothesis, Schoenfeld [21] has proved

$$|\pi(x) - \operatorname{li}(x)| < \frac{1}{8\pi}\sqrt{x}\log x, \qquad (x > 2657). \tag{67}$$

This result will be used to obtain (under RH) some precise bounds for p_n.

Lemma 6.1. *We have $\sqrt{x}(\log x)^{\frac{5}{2}} < \operatorname{ali}(x)$ for $x > 94$.*

Proof. The inequality is equivalent to

$$\operatorname{li}(\sqrt{x}(\log x)^{\frac{5}{2}}) < x.$$

Differentiating the function $f(x) := x - \operatorname{li}(\sqrt{x}(\log x)^{\frac{5}{2}})$ we get

$$f'(x) = 1 - \frac{1}{\log(\sqrt{x}(\log x)^{\frac{5}{2}})}\Big(\frac{\log^{5/2} x}{2\sqrt{x}} + \frac{5\log^{3/2} x}{2\sqrt{x}}\Big).$$

Hence this derivative is positive if and only if

$$\log^{3/2} x\Big(\frac{\log x}{2} + \frac{5}{2}\Big) < \sqrt{x}\Big(\frac{\log x}{2} + \frac{5}{2}\log\log x\Big).$$

For $x > 94$ we have $(\log x)^{\frac{3}{2}} < \sqrt{x}$, so that $f'(x) > 0$ for $x > 94$.

Finally, one may verify that $f(94) > 0$. Hence $f(x) > 0$ for $x > 94$. □

Theorem 6.2. *The Riemann hypothesis is equivalent to the assertion*

$$|p_n - \operatorname{ali}(n)| < \frac{1}{\pi}\sqrt{n}\log^{\frac52} n \qquad \text{for all } n \geq 11. \tag{68}$$

Proof. First we assume the Riemann Hypothesis and prove (68). Let $r(x) := \frac{1}{8\pi}\sqrt{x}\log x$, $f(x) := \operatorname{li}(x) - r(x)$, and $g(x) := \operatorname{li}(x) + r(x)$. For $x > 1$ we have $f(x) < \operatorname{li}(x) < g(x)$, where the three functions are strictly increasing. From (67) for $x > 2657$, we also have $f(x) < \pi(x) < g(x)$.

The inverse functions satisfy $g^{-1}(y) < \operatorname{ali}(y) < f^{-1}(y)$, and if $y = n > \pi(2657) = 384$ is a natural number, then $g^{-1}(n) < p_n < f^{-1}(n)$. It follows that the distance from $\operatorname{ali}(n)$ to p_n is bounded by

$$|p_n - \operatorname{ali}(n)| \leq \max\bigl(f^{-1}(n) - \operatorname{ali}(n), \operatorname{ali}(n) - g^{-1}(n)\bigr).$$

Hence, we have to bound $f^{-1}(y) - \operatorname{ali}(y)$ and $\operatorname{ali}(y) - g^{-1}(y)$.

We consider y as a parameter and set $\alpha = \operatorname{ali}(y)$, so that $\operatorname{li}(\alpha) = \operatorname{li}(\operatorname{ali}(y)) = y$.

Consider the function $u(\xi) := f(\xi) - \operatorname{li}(\alpha) = f(\xi) - y$, which is strictly increasing and satisfies

$$u(\xi) = \operatorname{li}(\xi) - r(\xi) - \operatorname{li}(\alpha) = \int_\alpha^\xi \frac{dt}{\log t} - r(\xi).$$

Therefore, $u(\alpha) = -r(\alpha) < 0$ and

$$u(f^{-1}(y)) = f(f^{-1}(y)) - \operatorname{li}(\alpha) = y - \operatorname{li}(\operatorname{ali}(y)) = 0.$$

If a point b is found where $u(b) > 0$, then $\alpha < f^{-1}(y) < b$, so that $b - \alpha > f^{-1}(y) - \alpha$ and one of the required bounds is obtained.

Therefore, we try $b = \alpha + c\sqrt{y}(\log y)^{\frac52}$ with $c < 1$. We have

$$u(\alpha + c\sqrt{y}(\log y)^{\frac52}) = \int_\alpha^{\alpha + c\sqrt{y}(\log y)^{\frac52}} \frac{dt}{\log t} - r(\alpha + c\sqrt{y}(\log y)^{\frac52}) >$$
$$> \frac{c\sqrt{y}(\log y)^{\frac52}}{\log(\alpha + c\sqrt{y}(\log y)^{\frac52})} - r(\alpha + c\sqrt{y}(\log y)^{\frac52}).$$

From Lemma 6.1 for $y > 94$, we have $\sqrt{y}(\log y)^{\frac52} < \operatorname{ali}(y) = \alpha$, so that

$$u(\alpha + c\sqrt{y}(\log y)^{\frac52}) > \frac{c\sqrt{y}(\log y)^{\frac52}}{\log(2\alpha)} - r(2\alpha)$$
$$= \frac{c\sqrt{y}(\log y)^{\frac52} - \frac{1}{8\pi}\sqrt{2\alpha}\log^2(2\alpha)}{\log(2\alpha)}.$$

We want to show that this expression is positive. For $y > 94$, we have $\alpha = \operatorname{ali}(y) < 2y\log y$ (by Lemma 5.7), so that $\alpha < 2y\log y < 4y\log y < y^2$ (for $y > 94$), which yields (with $c = 1/\pi$)

$$c\sqrt{y}(\log y)^{\frac52} - \frac{1}{8\pi}\sqrt{2\alpha}\log^2(2\alpha) >$$
$$> c\sqrt{y}(\log y)^{\frac52} - \frac{1}{8\pi}\sqrt{4y\log y}\log^2(4y\log y) >$$
$$> c\sqrt{y}(\log y)^{\frac52} - \frac{1}{8\pi}\sqrt{4y\log y}\log^2(y^2)) = 0.$$

Hence, we have proved that $f^{-1}(y) - \alpha < \frac{1}{\pi}\sqrt{y}(\log y)^{5/2}$ for $y > 94$.

To bound $\alpha - g^{-1}(y)$, we consider the function $v(\xi) := g(\xi) - \mathrm{li}(\alpha) = g(\xi) - y$. Then

$$v(\xi) = r(\xi) - \int_\xi^\alpha \frac{dt}{\log t}$$

and $v(\alpha) = r(\alpha) > 0$, $v(g^{-1}(y)) = g(g^{-1}(y)) - y = 0$. If a value b is found such that $v(b) < 0$, it will follow that $\alpha - g^{-1}(y) < \alpha - b$.

Choose $b = \alpha - c\sqrt{y}(\log y)^{5/2}$ with $c = \frac{1}{\pi}$. We claim that $v(b) < 0$. We have

$$v(b) = r(b) - \int_b^\alpha \frac{dt}{\log t} < r(b) - \frac{\alpha - b}{\log \alpha}$$

and our claim will follow from $r(b)\log\alpha - c\sqrt{y}(\log y)^{5/2} < 0$. Finally, since $b < \alpha = \mathrm{ali}(y) < 2y\log y$, and by Lemma 5.7,

$$r(b)\log\alpha < \frac{1}{8\pi}\sqrt{\mathrm{ali}(y)}(\log \mathrm{ali}(y))^2 < \frac{1}{8\pi}\sqrt{2y\log y}(2\log y)^2$$

which proves our claim.

Hence, (by assuming RH), we have proved that $|p_n - \mathrm{ali}(n)| < \frac{1}{\pi}\sqrt{n}(\log n)^{\frac{5}{2}}$ for $n > 384 > 94$. By verifying all $1 \le n \le 385$, we find that the inequality holds except for $n < 11$.

The reverse implication is simple. From $|p_n - \mathrm{ali}(n)| = \boldsymbol{\mathcal{O}}(n^{\frac{1}{2}+\varepsilon})$ for any $\varepsilon > 0$, we may derive that $\pi(x) = \mathrm{li}(x) + \boldsymbol{\mathcal{O}}(x^{\frac{1}{2}+\varepsilon})$. It is well known that this is equivalent to the Riemann Hypothesis. □

Remark 6.3. The inequality (68) is only proved by assuming the Riemann Hypothesis, but is stronger than those contained in [18], [16], [12], [4]. Inequality (68) gives approximately half of the digits of p_n. If our conjecture that (63) is true for all N is also assumed, then the asymptotic expansion gives about half of the digits of p_n.

6.2. **Inequalities for the n-th prime.** Let

$$s_N(n) = n\log n\Big(1 + \sum_{k=1}^{N} \frac{P_{k-1}(\log\log n)}{\log^k n}\Big)$$

where $s_0 = n\log n$.

Cipolla noted that for $k \ge 1$, $P_k(y) = (-1)^{k+1}\frac{y^k}{k} + \cdots$, and $P_0(y) = y - 1$. Hence, except for the first term, eventually the sign of the k-th term $P_{k-1}(\log\log n)\log^{-k} n$ becomes $(-1)^k$. The asymptotic expansion implies that there exist r_N such that

$$\begin{aligned} p_n > s_0(n), \quad n > r_0, \qquad p_n > s_1(n), \quad n > r_1, \\ p_n < s_{2N}(n), \quad n > r_{2N}, \qquad p_n > s_{2N+1}(n), \quad n > r_{2N+1}. \end{aligned} \tag{69}$$

In fact, $r_0 = 2$ is the main result in [18], $r_1 = 2$ is proved in [4] and $r_2 = 688\,383$ is proved in [5]. The value of r_N for $N \ge 3$ has not been determined. See Theorem 6.4 for an estimation of r_3 by assuming RH.

The above reasoning may give the impression that the terms of the asymptotic expansion of $\mathrm{ali}(u)$ are alternating in sign, starting from the second term. However this is not true. For example, computing the first 230 terms for $\mathrm{ali}(10^{100})$, we found only three positive terms P_0/x, P_1/x^2, and P_3/x^4. In fact, the sign of the k-th term is that of $P_{k-1}(\log\log n)$. Thus we are interested in the sign of these polynomials.

The polynomials $P_N(y)$, for $1 \le N \le 23$ of odd index, have one and only one real root, which is positive. Starting from $P_1(y)$ which vanishes at $y = 2$, these roots are

$$2, \quad 4.23415, \quad 5.83131, \quad 7.43591, \quad 9.07979, \quad 10.6881, \quad 12.2538,$$
$$13.7876, \quad 15.2977, \quad 16.79, \quad 18.2683, \quad 19.7353.$$

The polynomials P_2, P_4 and P_6 have no real roots, and all $P_8, \dots, P_{22}$ have two positive real roots. These pairs of roots are:

$$(6.4306, 8.2185), \quad (7.16158, 9.88528), \quad (7.90293, 11.4752),$$
$$(8.63359, 13.0241), \quad (9.3507, 14.5452), \quad (10.055, 16.0458),$$
$$(10.7478, 17.5307), \quad (11.4307, 19.003).$$

For example, $P_9(\log\log n)$ is positive only for $n > \exp(e^{9.07\cdots})$, which is a very big number.

The even terms at first sight appear *negative*. However $P_{10}(\log\log n)$, for example, is negative except in the interval $\exp(e^{7.16\cdots}) < n < \exp(e^{9.88\cdots})$.

In a certain sense, the inequalities (69) are the wrong inequalities. These inequalities would hold only for very large values of r_N, especially when we want a lower bound of $p_n > s_{2N+1}$ (except for the three known cases). We estimate r_3.

Theorem 6.4. *Let r_3 be the smallest number such that*

$$p_n > s_3 := n\log n + n(\log\log n - 1) + n\frac{\log\log n - 2}{\log n} - n\frac{(\log\log n)^2 - 6\log\log n + 11}{2\log^2 n}, \qquad n \ge r_3. \tag{70}$$

Then, if the Riemann Hypothesis is assumed,

$$39 \times 10^{29} < r_3 \le 39.58 \times 10^{29}.$$

Proof. By Theorem 6.2, there exists θ_1 with $|\theta_1| \le 1$ such that

$$p_n = \operatorname{ali}(n) + \theta_1\frac{\sqrt{n}}{\pi}(\log n)^{\frac52}, \qquad n \ge 11.$$

By Theorem 5.17, with $5 \le N \le 10$ and setting $n = e^x$, $x = \log n$, and $y = \log\log n$, we have for $n > e^{z_N}$

$$\operatorname{ali}(n) = xe^x\Big(1 + \frac{y-1}{x} + \frac{y-2}{x^2} - \frac{y^2 - 6y + 11}{2x^3} + \frac{P_3(y)}{x^4} + \sum_{k=5}^N \frac{P_{k-1}(y)}{x^k} + \theta_2 c_N\frac{y^N}{x^{N+1}}\Big).$$

The inequality of the Theorem is obtained if

$$xe^x\Big(\frac{P_3(y)}{x^4} + \sum_{k=5}^N \frac{P_{k-1}(y)}{x^k} - c_N\frac{y^N}{x^{N+1}}\Big) - \frac{e^{x/2}}{\pi}x^{\frac52} > 0.$$

This is equivalent to

$$P_3(y) + \sum_{k=5}^N P_{k-1}(y)e^{-(k-4)y} > c_N y^N e^{-(N-3)y} + \frac{1}{\pi}e^{\frac{11}{2}y}e^{-\frac12 e^y}.$$

Since $P_3(y) \to +\infty$ and all the other terms tend to 0 as $y \to +\infty$, it is clear that the inequality is true for $y > y_0$. With $N = 10$, we find $y_0 = 4.254946453\dots$ The inequality $p_n < s_n$ is true for $n \geq 3.95702224148845656 \times 10^{30}$. This proves that $r_3 \leq 39.58 \times 10^{29}$.

In order to show that $r_3 > 39 \times 10^{29}$, we directly show that, for $n = 39 \times 10^{29}$, the opposite inequality $p_n < s_3$ is obtained.

We compute

$$s_3 = 2.87527\,18639\,02974\,79681\,42399\,35057\,89294\,02005\,87915 \times 10^{32}.$$

Now we can compute $\operatorname{ali}(n)$, for which we already have obtained a good approximation through the asymptotic expansion, and then apply the Newton method

$$\operatorname{ali}(n) = 2.87527\,18639\,02495\,21516\,14800\,14732\,45414\,39731 \times 10^{32}.$$

Therefore, from Theorem 6.2, we obtain $p_n < \operatorname{ali}(n) + \frac{1}{\pi}\sqrt{n}\log^{\frac{5}{2}} n$, so that

$$p_n < 2.87527\,18639\,02756\,97808\,39055\,05640\,30082\,86370\,11482 \times 10^{32}$$

and we can conclude that $p_n < s_3$. □

References

[1] E. Cesàro, *Sur une formule empirique de M. Pervouchine*, C. R. Math. Acad. Sci. Paris **119** (1894), 848–849. (also in [2]).

[2] E. Cesàro, *Opere Scelte*, Vol. I, Parte Seconda, Edizioni Cremonese, Roma, 1965.

[3] M. Cipolla, *La determinazione assintotica dell' n^{imo} numero primo*, Rend. Accad. Sci. Fis-Mat. Napoli (3) **8** (1902), 132–166.

[4] P. Dusart, *The k-th prime is greater than $k(\log k + \log\log k - 1)$ for $k \geq 2$*, Math. Comp. **68** (1999), 411-415.

[5] P. Dusart, *Estimates of some functions over primes without R.H.*, arXiv 1002:0442 (2010). `http://front.math.ucdavis.edu/1002.0442`.

[6] A. E. Ingham, *The Distribution of Prime Numbers*, Cambridge University Press, 1932.

[7] E. Landau, *Über die Multiplikation Dirichlet'scher Reihen*, Rend. Circ. Matem. Palermo, **24** (1907), 81–159.

[8] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, Teubner, Leipzig 1909. [Reprint: Chelsea 1953].

[9] S. Lang, *Algebra*, Revised 3rd edition, Springer, New York 2002.

[10] J.-P. Massias, *Ordre maximum d'un élément du groupe symétrique et applications*, Thèse de 3ème cycle, Limoges, France, 1985.

[11] J.-P. Massias, J. L. Nicolas & G. Robin, *Evaluation asymptotique de l'ordre maximum d'un élément du groupe symétrique*, Acta Arithmetica **50** (1988), 221–242

[12] J.-P. Massias & G. Robin, *Bornes effectives pour certaines fonctions concernant les nombres premiers*, J. Théor. Nombres Bordeaux **8** (1996), 215–242.

[13] H. L. Montgomery & R. C. Vaughan, *Multiplicative Number Theory: I Classical Theory*, Cambridge University Press, 2006.

[14] I. M. Pervushin, *Les formules pour la détermination approximative des nombres premiers, de leur summe et de leur différence d'aprés le numéro de ces nombres*, Nachrichten der physiko-mathematischen Gesellschaft der Kaiserlichen Universität zu Kasan, (2) **4** (1894), 94–96.

[15] I. M. Pervushin, *Les formules pour la détermination approximative des nombres premiers, de leur summe et de leur différence d'aprés le numéro de ces nombres*, Verhandlungen des ersten internationalen Mathematiker-Kongresses in Zürich vom 9 bis 11 August 1897, Leipzig, Teubner, 1898.

[16] G. Robin, *Estimation de la fonction de Tchebychef θ sur le k-ième nombre premier et grandes valeurs de la fonction $\omega(n)$, nombre de diviseurs premiers de n*, Acta Arithmetica **42** (1983), 367–389.

[17] G. Robin, *Permanence de relations de récurrence dans certains développements asymptotiques*, Publ. Inst. Math. (Beograd) (N.S.) **43(57)** (1988), 17–25.

[18] B. Rosser, *The n-th prime is greater than $n \log n$*, Proc. London Math. Soc. (2) **45** (1939), 21–44.
[19] B. Rosser & L. Schoenfeld, *Sharper bounds for the Chebyshev Functions $\theta(x)$ and $\psi(x)$*, Math. Comp. **29** (1975), 243–269.
[20] B. Salvy, *Fast computation of some asymptotic functional inverses*, J. Symbolic Comput. **17** (1994), 227–236.
[21] L. Schoenfeld, *Sharper Bounds for the Chebyshev Functions $\theta(x)$ and $\psi(x)$. II*, Math. Comp. **30** (1976), 337-360.
[22] G. Torelli, *Sulla totalità dei numeri primi fino a un limite assegnato*, Atti della Reale Accademia delle Scienze Fisiche e Matematiche, Napoli (2) **11** (1901), 1–222.

Facultad de Matemáticas, Universidad de Sevilla, Apdo. 1160, 41080-Sevilla, Spain
E-mail address: arias@us.es

51 Route de Laborie, Foulayronnes 47510 (Lot-et-Garonne) France
E-mail address: toulisse.jeremy@voila.fr

THE n-TH PRIME ASYMPTOTICALLY

JUAN ARIAS DE REYNA AND JÉRÉMY TOULISSE

Abstract. A new derivation of the classic asymptotic expansion of the n-th prime is presented. A fast algorithm for the computation of its terms is also given, which will be an improvement of that by Salvy (1994).

Realistic bounds for the error with $\operatorname{li}^{-1}(n)$, after having retained the first m terms, for $1 \leq m \leq 11$, are given. Finally, assuming the Riemann Hypothesis, we give estimations of the best possible r_3 such that, for $n \geq r_3$, we have $p_n > s_3(n)$ where $s_3(n)$ is the sum of the first four terms of the asymptotic expansion.

1. Introduction.

1.1. **Historical note.** Chebyshev failed to fully prove the Prime Number Theorem (PNT), but he obtained some notable approximations. For example, he proved that for every natural number n: if the limit

$$\lim_{x\to\infty} \frac{\log^n x}{x}\bigl(\pi(x) - \operatorname{li}(x)\bigr)$$

exists, then this limit must be equal to 0.

The question was decided by de la Vallée Poussin (1899) when he gave his bound on the error in the PNT: The above limits exist and equal 0.

In 1894, Pervushin, a priest in Perm, published several formulae obtained empirically about prime numbers[1]. One of these formulae gives the following approximation to the n-th prime

$$\frac{p_n}{n} = \log n + \log\log n - 1 + \frac{5}{12\log n} + \frac{1}{24\log^2 n}.$$

Cesàro then published a note [1, 1894] where he asserts that the true formula is

$$\frac{p_n}{n} = \log n + \log\log n - 1 + \frac{\log\log n - 2}{\log n} - \\ - \frac{(\log\log n)^2 - 6\log\log n + 11}{2\log^2 n} + o(\log^{-2} n).$$

Despite no mention by Cesàro in [1], the editors of his collected works added a note to [1] pointing out that certain formulae quoted by Cesàro, since they followed from the results of Chebyshev, were only established under the assumption of the existence of the implied limits. It therefore remains unsurprising that Hilbert, in the Jahrbuch[2] stated that Césaro did not prove his formula.

[1] Ivan Mikheevich Pervushin (1827-1900) (Иван Михеевич Первушин). No small achievement if we note that he had only a table of primes up to 3 000 000.

[2] *Jahrbuch über die Fortschritte der Mathematik (1868–1942)*, a forerunner for the *Zentralblatt für Mathematik*, at present digitalized at `http://www.emis.de/MATH/JFM/JFM.html`.

Landau [7, 1907] several years later was better informed: a formula, like that of Cesàro, would imply the PNT, which had yet to be proved at Cesàro's time. However, using the results of Chebyshev, Cesàro may claim that if there is some formula for p_n correct to the order $n(\log n)^{-2}$, then it must coincide with his formula.

Cipolla [3, 1902] obtained an infinite asymptotic expansion for p_n and gave a recursive formula to compute its terms. He published after the results of de la Vallée Poussin but it seems that he was unaware of these results, so that gave his proof under the same hypotheses as Cesáro. So uninformed was he that he attempted to prove some false formulae of Pervushin already *corrected* by Torelli [22]

$$p_{n+1} - p_n = \log n + \log\log n + \frac{\log\log n - 1}{\log n} + \mathcal{O}\Big(\frac{\log\log n}{\log n}\Big)^2$$

with an *impeccable proof* that if such a formula exists, then it must be this formula. (Such a formula would refute the twin prime conjecture, and today the above formula is known to be false.)

In his *Handbuch* Landau [8, § 57] obtained by means of the procedure of Cesàro, some approximative formulae for p_n, and explained that the method could give further terms. He also mentioned some recursive formulae without giving any clue for their derivation.

We may say that Pervushin was the first to deal with a formula for p_n, albeit that he gave only the first few terms. Cesàro then proved that in the case such a formula exists, it must be one from which he would be able to derive several terms. Cipolla found a method to write all the terms of the expansion if there is one. Landau saw that the results of de la Vallée Poussin imply that the expansion certainly exists.

The algorithm given by Cipolla is not very convenient for the computation of the terms of the expansion. He iteratively computes the derivative of some polynomials appearing in the expansion but computes the constant terms as determinants of increasing order. Robin [17, 1988] considers the problem of computing these and other similar expansions, leaving the problem of computing the constant terms of the polynomials as an open problem. Later Salvy [20, 1994] gives a satisfactory algorithm. This algorithm needs $\mathcal{O}(n^{7/2}\sqrt{\log n})$ coefficient operations to compute all the polynomials up to the n-th polynomial.

The asymptotic expansion of p_n also plays a role in the study of $g(n)$, which is the maximum order of any element in the symmetric group S_n. In fact, $\log g(n)$ has the same asymptotic expansion as $\sqrt{\operatorname{li}^{-1}(x)}$ [11].

There are many results giving true bounds on p_n, for example we mention $p_n \geq n\log n$ [18, 1939], and $p_n \geq n(\log n + \log\log n - 1)$ [4, 1999] both for $n \geq 2$ (with partial results given in [19], [16], [12], [4]). In [5] it is also proved that

$$p_n \leq n\Big(\log n + \log\log n - 1 + \frac{\log\log n - 2}{\log n}\Big), \qquad n \geq 688\,383.$$

1.2. **Organization of the paper.** In this paper we present a new derivation of the asymptotic expansion for p_n and obtain explicit bounds for the error.

First, it must be said that the asymptotic expansion has, in a certain sense, nothing to do with prime numbers: it is an asymptotic expansion of $\operatorname{ali}(x) := \operatorname{li}^{-1} x$ which is the inverse of the usual logarithmic integral function.

In Section 3 a proof of the existence of the expansion is given, following the path of Cesàro, since it cannot be found elsewhere, although it is frequently claimed it can be done. This Section is not needed in the rest of the paper.

In Section 4, a new formal derivation of the expansion is given. We obtain a new algorithm to compute the polynomials (Theorem 4.9). This is simpler than that given by Salvy [20]. Our algorithm allows all the polynomials up to the n-th one to be computed in $\mathcal{O}(n^2)$ coefficient operations (Theorem 4.11). It must be said that these polynomials have $\mathcal{O}(n^2)$ coefficients.

In Section 5, independently of Section 3, we prove that the formal expansion given in Section 4 is in fact the asymptotic expansion of $\operatorname{ali}(x)$ and gives realistic bounds on the error (Theorem 5.16 and 5.17).

In Section 6, the results are applied to p_n the n-th prime. Using the results of de la Vallée Poussin it can be shown that the asymptotic expansion of $\operatorname{ali}(n)$ is also an asymptotic expansion for p_n.

By assuming the Riemann Hypothesis, we found (Theorem 6.2) that

$$|p_n - \operatorname{ali}(n)| \leq \frac{1}{\pi}\sqrt{n}\,(\log n)^{\frac{5}{2}}, \qquad n \geq 11.$$

This bound of p_n is better than all the bounds cited above.

We end the paper by motivating why the above bounds have not been extended to further terms of the asymptotic expansion (Theorem 6.4).

Notations: With a certain hesitation we have introduced the notation $\operatorname{ali}(x)$ to denote the inverse function of $\operatorname{li}(x)$.

In Section 5, where explicit bounds are sought, it has been useful to denote by θ a real or complex number of absolute value $|\theta| \leq 1$, which will not always be the same, and depends on all parameters or variables in the corresponding equation.

Acknowledgement: The authors would like to thank Jan van de Lune (Hallum, The Netherlands) for his linguistic assistance in preparing the paper, and his interest in our results.

2. The inverse function of the logarithmic integral.

Usually $\operatorname{li}(x)$ is defined for real x as the principal value of the integral

$$\operatorname{li}(x) = \text{p. v.} \int_0^x \frac{dt}{\log t}.$$

It may be extended to an analytic function over the region $\Omega = \mathbb{C} \smallsetminus (-\infty, 1]$, which is the complex plane with a cut along the real axis $x \leq 1$. The main branch of the logarithm is defined in Ω and does not vanish there. Therefore, $\operatorname{li}(z)$ may be defined in Ω by

$$\operatorname{li}(z) = \operatorname{li}(2) + \int_2^z \frac{dt}{\log t}, \qquad z \in \Omega \tag{1}$$

where we integrate, for example, along the segment from 2 to z.

For real $x > 1$, the function $\operatorname{li}(x)$ is increasing and maps the interval $(1, +\infty)$ onto $(-\infty, +\infty)$, so that we may define the inverse function $\operatorname{ali}\colon \mathbb{R} \to (1, +\infty)$ by

$$\operatorname{li}(\operatorname{ali}(x)) = x. \tag{2}$$

The function $\operatorname{li}(x)$ is analytic on Ω, so that $\operatorname{ali}(x)$ is real analytic. It is clear that we have the following rules of differentiation

$$\frac{d}{dx}\operatorname{li}(x)=\frac{1}{\log x},\qquad \frac{d}{dx}\operatorname{ali}(x)=\log\operatorname{ali}(x). \tag{3}$$

It is well known that the function $\operatorname{li}(x)$ has an asymptotic expansion:

Theorem 2.1. *For each integer $N\ge 0$*

$$\operatorname{li}(x)=\frac{x}{\log x}\Big(1+\sum_{k=1}^{N}\frac{k!}{\log^k x}+\mathcal{O}\Big(\frac{1}{\log^{N+1}x}\Big)\Big),\qquad (x\to+\infty). \tag{4}$$

This may be proved by repeated integration by parts (see [13, p. 190–192]).

3. Asymptotic expansion of $\operatorname{ali}(x)$.

In this section, we prove the following

Theorem 3.1. *For each integer $N\ge 0$*

$$\frac{\operatorname{ali}(e^x)}{xe^x}=1+\sum_{n=1}^{N}\frac{P_{n-1}(\log x)}{x^n}+\mathcal{O}\Big(\frac{\log^{N+1}x}{x^{N+1}}\Big),\qquad (x\to+\infty) \tag{5}$$

where the $P_{n-1}(z)$ are polynomials of degree $\le n$.

In the case of $N=0$ the sum must be understood as equal to 0.

The theorem says that, for each N, there exists an $x_N>1$ and a constant C_N such that

$$\Big|\frac{\operatorname{ali}(e^x)}{xe^x}-1-\sum_{n=1}^{N}\frac{P_{n-1}(\log x)}{x^n}\Big|\le C_N\frac{\log^{N+1}x}{x^{N+1}},\qquad (x>x_N).$$

In the course of the proof we will make repeated use of the following

Lemma 3.2. *Let $f(x)$ be a function defined on a neighbourhood of $x=0$ such that*

$$f(x)=a_1x+\cdots+a_Nx^N+\mathcal{O}(x^{N+1}),\qquad (x\to 0) \tag{6}$$

where the a_k are given constants. Assume that $g(x)$ satisfies

$$g(x)=\sum_{n=1}^{N}\frac{p_n(\log x)}{x^n}+\mathcal{O}\Big(\frac{\log^{N+1}x}{x^{N+1}}\Big),\qquad (x\to+\infty) \tag{7}$$

where the $p_n(z)$ are polynomials of degree $\le n$. Then there exist polynomials $q_k(z)$ of degree $\le k$ such that

$$f(g(x))=\sum_{k=1}^{N}\frac{q_k(\log x)}{x^k}+\mathcal{O}\Big(\frac{\log^{N+1}x}{x^{N+1}}\Big),\qquad (x\to+\infty). \tag{8}$$

Proof. It is clear that, for each $1\le n\le N$ we have $p_n(\log x)x^{-n}=\mathcal{O}((\log x/x)^n)$. Therefore, $g(x)=\mathcal{O}(\log x/x)$ (this is true even when $N=0$ and there is no p_n). It follows that $\lim_{x\to+\infty}g(x)=0$ and by substitution in (6), we obtain

$$f(g(x))=\sum_{n=1}^{N}a_ng(x)^n+\mathcal{O}\Big(\frac{\log^{N+1}x}{x^{N+1}}\Big). \tag{9}$$

By expanding the powers $g(x)^n$ by (7) it is easy to obtain an expression of the form

$$g(x)^n = \sum_{k=1}^{N} \frac{p_{n,k}(\log x)}{x^k} + \mathcal{O}\Big(\frac{\log^{N+1} x}{x^{N+1}}\Big), \qquad (x \to +\infty) \tag{10}$$

where each $p_{n,k}(z)$ is a polynomial of degree $\leq k$. By substituting these values in equation (9) and collecting terms with the same power of x, (8) is obtained. □

We will prove Theorem 3.1 by induction. The following theorem yields the first step of this induction.

Theorem 3.3.

$$\frac{\operatorname{ali}(x)}{x \log x} = 1 + \mathcal{O}\Big(\frac{\log\log x}{\log x}\Big), \qquad (x \to +\infty). \tag{11}$$

Proof. From (4) with $N = 0$ we have $\frac{\operatorname{li}(y)\log y}{y} = 1 + \mathcal{O}(\log^{-1} y)$ for $y \to \infty$. Since $\lim_{x\to+\infty} \operatorname{ali}(x) = +\infty$ we may substitute $y = \operatorname{ali}(x)$ and obtain

$$\frac{x \log \operatorname{ali}(x)}{\operatorname{ali}(x)} = 1 + \mathcal{O}\Big(\frac{1}{\log \operatorname{ali}(x)}\Big). \tag{12}$$

By taking logarithms

$$\log x - \log \operatorname{ali}(x) + \log\log \operatorname{ali}(x) = \mathcal{O}\Big(\frac{1}{\log \operatorname{ali}(x)}\Big)$$

we obtain

$$\frac{\log x}{\log \operatorname{ali}(x)} = 1 + \mathcal{O}\Big(\frac{\log\log \operatorname{ali}(x)}{\log \operatorname{ali}(x)}\Big) \tag{13}$$

and it follows that

$$\lim_{x\to+\infty} \frac{\log x}{\log \operatorname{ali}(x)} = 1. \tag{14}$$

By taking log in (13)

$$\log\log x - \log\log \operatorname{ali}(x) = \mathcal{O}\Big(\frac{\log\log \operatorname{ali}(x)}{\log \operatorname{ali}(x)}\Big)$$

we obtain

$$\frac{\log\log x}{\log\log \operatorname{ali}(x)} = 1 + \mathcal{O}\Big(\frac{1}{\log \operatorname{ali}(x)}\Big)$$

so that

$$\lim_{x\to+\infty} \frac{\log\log x}{\log\log \operatorname{ali}(x)} = 1. \tag{15}$$

In view of (14) and (15), we may write (12) and (13) in the form

$$\frac{x \log \operatorname{ali}(x)}{\operatorname{ali}(x)} = 1 + \mathcal{O}\Big(\frac{1}{\log x}\Big), \qquad \frac{\log x}{\log \operatorname{ali}(x)} = 1 + \mathcal{O}\Big(\frac{\log\log x}{\log x}\Big) \tag{16}$$

and by multiplying these two, we obtain

$$\frac{x \log x}{\operatorname{ali}(x)} = 1 + \mathcal{O}\Big(\frac{\log\log x}{\log x}\Big)$$

from which (11) can easily be deduced. □

Proof of Theorem 3.1. We proceed by induction. For $N = 0$, our theorem is simply Theorem 3.3 with e^x instead of x.

Hence we assume (5) and try to prove the case $N + 1$.

Our objective will be obtained by starting from the expansion of $\mathrm{li}(y)$. By (4)

$$\mathrm{li}(y) = \frac{y}{\log y}\Big(1 + \sum_{k=1}^{N+1} \frac{k!}{\log^k y} + \mathcal{O}\Big(\frac{1}{\log^{N+2} y}\Big)\Big).$$

By substituting $y = \mathrm{ali}(e^x)$ and applying (14) we obtain

$$\frac{e^x \log \mathrm{ali}(e^x)}{\mathrm{ali}(e^x)} = 1 + \sum_{k=1}^{N+1} \frac{k!}{(\log \mathrm{ali}(e^x))^k} + \mathcal{O}\Big(\frac{1}{x^{N+2}}\Big). \tag{17}$$

From our induction hypothesis, the expansion of $\log \mathrm{ali}(x)$ and $(\log \mathrm{ali}(x))^{-k}$ is now sought.

By taking the log of (5) we obtain

$$\log \mathrm{ali}(e^x) = x + \log x + \log\Big\{1 + \sum_{n=1}^{N} \frac{P_{n-1}(\log x)}{x^n} + \mathcal{O}\Big(\frac{\log^{N+1} x}{x^{N+1}}\Big)\Big\}.$$

Lemma 3.2 may be applied with

$$\log(1+X) = X - \frac{X^2}{2} + \frac{X^3}{3} - \cdots (-1)^{N+1}\frac{X^N}{N} + \mathcal{O}(X^{N+1})$$

to obtain

$$\log \mathrm{ali}(e^x) = x + \log x + \sum_{n=1}^{N} \frac{Q_{n+1}(\log x)}{x^n} + \mathcal{O}\Big(\frac{\log^{N+1} x}{x^{N+1}}\Big).$$

The reason why we have written Q_{n+1} instead of $Q_n(x)$ is revealed below. The above may be written as

$$\log \mathrm{ali}(e^x) = x\Big\{1 + \frac{\log x}{x} + \sum_{n=1}^{N} \frac{Q_{n+1}(\log x)}{x^{n+1}} + \mathcal{O}\Big(\frac{\log^{N+1} x}{x^{N+2}}\Big)\Big\}$$

or

$$\log \mathrm{ali}(e^x) = x\Big\{1 + \sum_{n=1}^{N+1} \frac{Q_n(\log x)}{x^n} + \mathcal{O}\Big(\frac{\log^{N+2} x}{x^{N+2}}\Big)\Big\} \tag{18}$$

where the $Q_n(z)$ are polynomials of degree $\le n$. Observe that knowing the expansion of $\mathrm{ali}(e^x)$ up to $(\log x/x)^{N+1}$ has enabled us to obtain $\log \mathrm{ali}(e^x)$ up to $(\log x/x)^{N+2}$; this will be of great importance in what follows.

From (18), for all natural numbers n,

$$\frac{1}{(\log \mathrm{ali}(e^x))^n} = \frac{1}{x^n}\Big\{1 + \sum_{k=1}^{N+1} \frac{Q_k(\log x)}{x^k} + \mathcal{O}\Big(\frac{\log^{N+2} x}{x^{N+2}}\Big)\Big\}^{-n}.$$

By applying Lemma 3.2 with

$$(1+x)^{-n} - 1 = \sum_{r=1}^{N+1} \binom{-n}{r} x^r + \mathcal{O}(x^{-N-2})$$

we obtain

$$\frac{1}{\{\log \operatorname{ali}(e^x)\}^n} = \frac{1}{x^n}\Big(1 + \sum_{k=1}^{N+1} \frac{V_{n,k}(\log x)}{x^k} + \mathcal{O}\Big(\frac{\log^{N+2} x}{x^{N+2}}\Big)\Big) \tag{19}$$

where the $V_{n,k}(z)$ are polynomials of degree $\le k$.

By substituting these values of $\{\log \operatorname{ali}(e^x)\}^{-n}$ in (17), we obtain

$$\frac{e^x \log \operatorname{ali}(e^x)}{\operatorname{ali}(e^x)} = 1 + \sum_{k=1}^{N+1} \frac{U_k(\log x)}{x^k} + \mathcal{O}\Big(\frac{\log^{N+2} x}{x^{N+2}}\Big).$$

Hence again from (19) with $n = 1$

$$\frac{e^x}{\operatorname{ali}(e^x)} = \frac{1}{x}\Big\{1 + \sum_{k=1}^{N+1} \frac{V_{1,k}(\log x)}{x^k} + \mathcal{O}\Big(\frac{\log^{N+2} x}{x^{N+2}}\Big)\Big\}\times$$
$$\times\Big\{1 + \sum_{k=1}^{N+1} \frac{U_k(\log x)}{x^k} + \mathcal{O}\Big(\frac{\log^{N+2} x}{x^{N+2}}\Big)\Big\}$$

from which we derive that there exist polynomials $W_k(z)$ of degree $\le k$ such that

$$\frac{xe^x}{\operatorname{ali}(e^x)} = 1 + \sum_{k=1}^{N+1} \frac{W_k(\log x)}{x^k} + \mathcal{O}\Big(\frac{\log^{N+2} x}{x^{N+2}}\Big).$$

Another application of Lemma 3.2 yields

$$\frac{\operatorname{ali}(e^x)}{xe^x} = 1 + \sum_{k=1}^{N+1} \frac{T_k(\log x)}{x^k} + \mathcal{O}\Big(\frac{\log^{N+2} x}{x^{N+2}}\Big)$$

with polynomials $T_k(z)$ of degree $\le k$. Therefore, we have an asymptotic expansion of type (5) with $N + 1$ instead of N. The usual argument of uniqueness of the asymptotic expansion applies here so that $T_k(z) = P_k(z)$ for $1 \le k \le N$. □

4. Formal Asymptotic expansion.

First we give some motivation. We have seen that the asymptotic expansion of $\operatorname{ali}(e^x)$ is

$$\operatorname{ali}(e^x) = xe^x V(x, \log x), \quad \text{where} \quad V(x,y) := 1 + \sum_{n=1}^{\infty} \frac{P_{n-1}(y)}{x^n}$$

and differentiation yields

$$e^x \log \operatorname{ali}(e^x) = (e^x + xe^x)V + xe^x V_x + e^x V_y.$$

Here $\log \operatorname{ali}(e^x) = \log\big(xe^x V(x, \log x)\big) = y + x + \log V$, so that

$$y + x + \log V = V + xV + xV_x + V_y$$

which we write as

$$V = 1 + \frac{y}{x} - \frac{1}{x}V - V_x - \frac{1}{x}V_y + \frac{1}{x}\log V. \tag{20}$$

This ends our motivation for considering this equation.

Consider now the ring A of the formal power series of the type

$$\sum_{n=0}^{\infty} \frac{q_n(y)}{x^n}$$

where the $q_n(y)$ are polynomials with complex coefficients of degree less than or equal to n. In particular $q_0(y)$ is a constant.

It is clear that A, with the obvious operations, is a ring. The elements with $q_0 = 0$ form a maximal ideal I. An element $1 + u$ with $q_0 = 1$ is invertible, with inverse $1 - u + u^2 - \cdots$. It follows that if $a \notin I$, then a is also invertible. Hence I is the unique maximal ideal and A is a local ring. If $a \in A$ is a non-vanishing element, then there exists a least natural number n with $q_n(y) \neq 0$. We define $\deg(a) = n$ in this case, with $\deg(0) = \infty$.

As is usual in local rings, (see [9]) we may define a topology induced by the norm $\|a\| = 2^{-\deg(a)}$, which, with the associated metric, induces a complete metric space. Indeed A is isomorphic to $\mathbb{C}[[X, Y]]$, by means of the application that sends $X \mapsto x^{-1}$, $Y \mapsto yx^{-1}$.

Given $a \in A$ with $a = \sum_{n=0}^{\infty} \frac{q_n(y)}{x^n}$, we define two derivates

$$a_x = -\sum_{n=1}^{\infty} \frac{nq_n(y)}{x^{n+1}} \quad \text{and} \quad a_y = \sum_{n=1}^{\infty} \frac{q_n'(y)}{x^n}.$$

Finally the set $U \subset A$ of elements with $q_0 = 1$ form a multiplicative subgroup of A^* (the group of invertible elements of A). For $1 + u \in U$, we define

$$\log(1+u) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{u^k}{k}$$

which is a series that is easily shown to converge since $u^k \in I^k$.

We are now ready to prove the following

Theorem 4.1. *The equation* (20) *has one and only one solution in the ring A.*

Proof. For $V \in U$, we define $T(V)$ as

$$T(V) := 1 + \frac{y}{x} - \frac{1}{x}V - V_x - \frac{1}{x}V_y + \frac{1}{x}\log V.$$

It is clear that $T(V) \in U$. We may apply Banach's fixed-point theorem. Indeed, we have $\deg(T(V) - T(W)) \leq 1 + \deg(V - W)$, so that $\|T(V) - T(W)\| \leq \frac{1}{2}\|V - W\|$.

By Banach's theorem there is a unique solution to $V = T(V)$. We may obtain this solution as the limit of the sequence $T^n(1)$. In fact, since $\deg(T(V) - T(W)) \leq 1 + \deg(V - W)$, in each iteration we obtain one further term of the expansion. In this way, it is easy to prove that the solution is

$$V = 1 + \frac{y-1}{x} + \frac{y-2}{x^2} + \cdots$$

However, we are going to find more direct methods to compute the terms of the expansion. □

Definition 4.2. Let V be the unique solution to equation (20). Since it is in A, it has the form

$$V(x, y) = 1 + \sum_{n=1}^{\infty} \frac{P_{n-1}(y)}{x^n} \tag{21}$$

where for $n \geq 0$, $P_{n-1}(y)$ is a polynomial of degree $\leq n$.

In the following sections, we prove that V yields the asymptotic expansion of $\mathrm{ali}(e^x)$. For this proof the following property is crucial.

Theorem 4.3. *For $N \geq 1$, let*

$$W(x,y) = W_N(x,y) := 1 + \sum_{k=1}^{N} \frac{P_{n-1}(y)}{x^n}. \tag{22}$$

Then

$$W - 1 - \frac{y}{x} + \frac{1}{x}W + W_x + \frac{1}{x}W_y - \frac{1}{x}\log W = -\frac{P_N(y)}{x^{N+1}} + \frac{u(y)}{x^{N+2}} + \frac{v(y)}{x^{N+3}} + \cdots \tag{23}$$

Proof. By the definition of the P_n we know that $\deg(V - W) \geq N+1$. Therefore, $\deg(V - T(W)) \geq N+2$. That is

$$V - T(W) = \frac{u_0(y)}{x^{N+2}} + \frac{v_0(y)}{x^{N+3}} + \cdots$$

where u_0, v_0 are polynomials. We also have

$$V = W + \sum_{n=N+1}^{\infty} \frac{P_{n-1}(y)}{x^n}$$

so that

$$W - T(W) = \frac{u_0(y)}{x^{N+2}} + \frac{v_0(y)}{x^{N+3}} + \cdots - \sum_{n=N+1}^{\infty} \frac{P_{n-1}}{x^n}.$$

That is,

$$W - T(W) = -\frac{P_N(y)}{x^{N+1}} + \frac{u(y)}{x^{N+2}} + \frac{v(y)}{x^{N+3}} + \cdots$$

for certain polynomials u, v, … □

In the sequel V will denote the unique solution to (20). The element $\log V$ belongs to A, so that there are polynomials $Q_n(y)$ of degree less than or equal to n such that

$$\log V = \sum_{n=1}^{\infty} \frac{Q_n(y)}{x^n}. \tag{24}$$

From equation (20), we may obtain $\log V$ in terms of V and its derivatives. It is easy to obtain from this expression the following relation

$$Q_n(y) = P_n(y) - (n-1)P_{n-1}(y) + P'_{n-1}(y), \qquad (n \geq 1). \tag{25}$$

Theorem 4.4. *The polynomials $P_n(y)$ that appear in the unique solution* (21) *to equation* (20) *may be computed by the following recurrence relations:*

$$P_0 = y - 1, \quad \textit{and for } n \geq 1$$

$$P_n = nP_{n-1} - P'_{n-1} + \frac{1}{n}\sum_{k=1}^{n-1} k\big\{(k-1)P_{k-1} - P_k - P'_{k-1}\big\}P_{n-k-1}. \tag{26}$$

Proof. By differentiating (24) with respect to x, we obtain

$$\Bigl(\sum_{n=1}^\infty \frac{nQ_n(y)}{x^{n+1}}\Bigr)\Bigl(1+\sum_{n=1}^\infty \frac{P_{n-1}(y)}{x^n}\Bigr) = \sum_{n=1}^\infty \frac{nP_{n-1}(y)}{x^{n+1}}.$$

By equating the coefficients of x^{-n-1}, we obtain

$$nP_{n-1} = nQ_n + \sum_{k=1}^{n-1} kQ_kP_{n-k-1}, \qquad (n\ge 2). \tag{27}$$

Now we substitute the values of the Q_n given in (25)

$$nP_{n-1} = nP_n - n(n-1)P_{n-1} + nP'_{n-1} + \sum_{k=1}^{n-1} kQ_kP_{n-k-1}$$

so that

$$nP_n = n^2P_{n-1} - nP'_{n-1} - \sum_{k=1}^{n-1} k\bigl\{P_k-(k-1)P_{k-1}+P'_{k-1}\bigr\}P_{n-k-1}.$$

□

From this expression it is very easy to compute the first terms of the expansions

$$V = 1+\frac{y-1}{x}+\frac{y-2}{x^2}-\frac{y^2-6y+11}{2x^3}+\frac{2y^3-21y^2+84y-131}{6x^4}- \\ -\frac{3y^4-46y^3+294y^2-954y+1333}{12x^5}+\cdots,$$

$$\log V = \frac{y-1}{x}-\frac{y^2-4y+5}{2x^2}+\frac{2y^3-15y^2+42y-47}{6x^3}- \\ -\frac{3y^4-34y^3+156y^2-366y+379}{12x^4}+\cdots$$

Theorem 4.5. *We have*

(a) *For $n\ge 1$, the degree of P_n is less than or equal to n.*
(b) *$n!\,P_n(y)$ has integer coefficients.*

Proof. The equation (20) may be written

$$V-1-\frac{y}{x} = \frac{1}{x}\bigl(\log V - V - xV_x - V_y\bigr).$$

Since $xV_x\in A$, it is clear that

$$xV - x - y = -1 + \sum_{n=1}^\infty \frac{P_n(y)}{x^n} \in A.$$

This implies that the degree of P_n is less than or equal to n.

We prove (b) by induction. The first few P_n satisfy this property. We define $p_k := k!P_k$ so that the recurrence relation (26) may be written as

$$p_n = n^2p_{n-1} - np'_{n-1} + \\ +(n-1)\sum_{k=1}^{n-1}\binom{n-2}{k-1}\bigl\{k(k-1)p_{k-1}-p_k-kp'_{k-1}\bigr\}p_{n-k-1}.$$

Hence, by induction, all p_n have integer coefficients. □

The most significant contribution by Cipolla is his proof of a recurrence for the coefficients $a_{n,k}$ of P_n (see (30)), which is better than the recurrence given in (26). We intend to give a slightly different proof. The result of Cipolla is equivalent to the following surprising fact: The solution V of equation (20) formally satisfies the following linear partial differential equation:

$$V = (x-1)V_y - xV_x. \tag{28}$$

This equation can easily be deduced from the following Theorem.

Theorem 4.6. *For $n \geq 1$, we have*

$$\begin{aligned} (n-1)P_{n-1}(y) &= P'_{n-1}(y) - P'_n(y), \qquad (n \geq 1) \\ (n-1)Q_{n-1}(y) &= Q'_{n-1}(y) - Q'_n(y), \qquad (n \geq 2). \end{aligned} \tag{29}$$

Proof. We will proceed by induction. For $n \leq 3$ it can be verified that these equalities are satisfied.

We now assume that (29) is satisfied for $n \leq N$, and we will show that these equations are true for $n = N+1$.

By differentiating (24) with respect to y we get

$$\Big(\sum_{n=1}^{\infty} \frac{Q'_n(y)}{x^n}\Big)\Big(1 + \sum_{n=1}^{\infty} \frac{P_{n-1}(y)}{x^n}\Big) = \sum_{n=1}^{\infty} \frac{P'_{n-1}(y)}{x^n}$$

so that by equating the coefficients of x^{-N-1} and of x^{-N} we obtain

$$Q'_{N+1} = P'_N - \sum_{k=0}^{N-1} P_k Q'_{N-k}, \qquad Q'_N = P'_{N-1} - \sum_{k=0}^{N-2} P_k Q'_{N-k-1}.$$

Subtracting these equations we get

$$Q'_{N+1} - Q'_N = P'_N - P'_{N-1} - \sum_{k=0}^{N-2} P_k (Q'_{N-k} - Q'_{N-k-1}) - P_{N-1}$$

and by the induction hypothesis this is equal to

$$-(N-1)P_{N-1} + \sum_{k=0}^{N-2} P_k \cdot (N-k-1)Q_{N-k-1} - P_{N-1} =$$

$$= -NP_{N-1} + \sum_{k=1}^{N-1} kQ_k P_{N-k-1}.$$

By (27) this is equal to $NP_{N-1} - NQ_N$ so that we obtain

$$Q'_{N+1} - Q'_N = -NQ_N.$$

This is the second equation of (29) for $n = N+1$. In order to achieve the result for the first equation, observe that from (25) we get

$$\begin{aligned} NQ_N &= NP_N - N(N-1)P_{N-1} + NP'_{N-1} \\ -Q'_N &= -P'_N + (N-1)P'_{N-1} - P''_{N-1} \\ Q'_{N+1} &= P'_{N+1} - NP'_N + P''_N. \end{aligned}$$

By adding these equations we obtain

$$0 = NP_N - P'_N + P'_{N+1} + N\{P'_{N-1} - P'_N - (N-1)P_{N-1}\} - \\ - \{P''_{N-1} - P''_N - (N-1)P'_{N-1}\} = NP_N - P'_N + P'_{N+1}$$

which is the first equation of (29) for $n = N+1$. □

We define the coefficients $a_{n,k}$ implicitly by

$$P_n(y) = \frac{(-1)^{n+1}}{n!}\Big(a_{n,0}y^n - a_{n,1}y^{n-1} + \cdots (-1)^n a_{n,n}\Big) = \\ = \frac{(-1)^{n+1}}{n!}\sum_{k=0}^{n}(-1)^k\, a_{n,k}\, y^{n-k}, \qquad (n \geq 1). \tag{30}$$

Analogously, Q_n is of a degree less than or equal to n, and we define the coefficients $b_{n,k}$ implicitly by

$$Q_n(y) = \frac{(-1)^{n+1}}{n!}\sum_{k=0}^{n}(-1)^k\, b_{n,k}\, y^{n-k}, \qquad (n \geq 1). \tag{31}$$

Remark 4.7. $P_0(y)$ has degree 1, which is not given by (30). However, we can extend the definition of $a(n,k)$ in such a way that, for $n \geq 1$ we have $a(n,k) = 0$ for $k < 0$ or $k > n$. Then a formula such as (30) also holds for $n = 0$ if we add up the values from $k = -1$ to $k = n$ and define $a(0,0) = 1$, $a(0,-1) = 1$ and $a(0,k) = 0$ for other values of k

$$P_n(y) = \frac{(-1)^{n+1}}{n!}\sum_{k=-1}^{n}(-1)^k\, a_{n,k}\, y^{n-k}, \qquad (n \geq 0). \tag{30 bis}$$

Note that $Q_0(y)$ remains undefined.

Theorem 4.8. *For $1 \leq n$ and $0 \leq k < n$, we have (when defined)*

$$a_{n,k} = na_{n-1,k-1} + \frac{n(n-1)}{n-k}a_{n-1,k}, \quad b_{n,k} = nb_{n-1,k-1} + \frac{n(n-1)}{n-k}b_{n-1,k}. \tag{32}$$

For $1 \leq n$ and $0 \leq k \leq n$, we have

$$a_{n,k} = b_{n,k} + (n-k+1)a_{n,k-1}. \tag{33}$$

For $n \geq 1$, we have

$$b_{n,n} = na_{n-1,n-1} + \sum_{k=1}^{n-1}\binom{n-1}{k}\, k\, b_{k,k}\, a_{n-k-1,n-k-1}. \tag{34}$$

Proof. (32) is obtained by equating the coefficients of y^{n-k-1} in the first equation in (29). In this way, we obtain

$$(n-1)\frac{(-1)^{n+k}}{(n-1)!}a_{n-1,k} = \\ = (n-k)\frac{(-1)^{n+k+1}}{(n-1)!}a_{n-1,k-1} - (n-k)\frac{(-1)^{n+k+1}}{n!}a_{n,k}.$$

If $n \neq k$, then the equation for $a_{n,k}$ in (32) is obtained. The other equation in $b_{n,k}$ is obtained analogously from the second equation in (29).

To prove (33), observe that by (25), $Q_n = P_n - (n-1)P_{n-1} + P'_{n-1}$, and from (29) it follows that

$$Q_n = P_n + P'_n, \qquad (n \geq 1). \tag{35}$$

Now by equating the coefficient of y^{n-k} in both members of this equality we obtain (33).

Finally (34) follows from (27). Recall that $-\frac{a_{n,n}}{n!}$ and $-\frac{b_{n,n}}{n!}$ are respectively the values of $P_n(0)$, and $Q_n(0)$. Hence, by setting $y = 0$ in (27), we obtain (34) through multiplication by $(n-1)!$ and the reordering of the terms. □

The main problem now is that equations (32) do not allow us to compute the coefficients $a_{n,n}$. Cipolla gives an algorithm to simultaneously compute the coefficients $a_{n,k}$ and $b_{n,k}$ based on Theorem 4.8. In the procedure of Cipolla, these key coefficients $a_{n,n}$ are recursively computed using all the previous coefficients. We prefer a method that computes $A_n := a_{n,n}$ and $B_n := a_{n,n-1}$ separately and then compute the remaining coefficients by using (32).

Theorem 4.9. *In order to compute the numbers $a_{n,k}$, we may first compute the sequences $A_n := a_{n,n}$ and $B_n := a_{n,n-1}$ by the recursions*

$$A_0 = 1, \quad A_1 = 2, \quad B_0 = 1, \quad B_1 = 1, \tag{36}$$

$$B_n = nB_{n-1} + n(n-1)A_{n-1} \tag{37}$$

$$A_n = n^2 A_{n-1} + nB_{n-1} - (n-1)\sum_{k=1}^{n-1}\binom{n-2}{k-1}\big\{k(k-1)A_{k-1} - A_k + kB_{k-1}\big\}A_{n-k-1}. \tag{38}$$

After this one we may obtain $a(n,k) := a_{n,k}$. Setting

$$a(0,0) = 1, \quad a(0,-1) = 1, \quad a(1,0) = 1, \quad a(1,1) = 2$$

and all other $a(0,k)$ and $a(1,k) = 0$. Then, for $n \geq 2$, put

$$a(n,n) = A_n,$$

$$a(n,k) = na(n-1,k-1) + \frac{n(n-1)}{n-k}a(n-1,k), \quad (0 \leq k < n) \tag{39}$$

where $a(n,k) = 0$ for $k < 0$ or $k > n$.

Finally, we may obtain the $b(n,k) := b_{n,k}$ from

$$b(n,k) = a(n,k) - (n-k+1)a(n,k-1). \tag{40}$$

Proof. The constant term of P_n is $-\frac{A_n}{n!}$ and the coefficient of y in P_n is $\frac{B_n}{n!}$, so that equation (37) follows from the first equation in (29) taking it with $y = 0$.

In the same way, (38) follows from (26), and (39) is the first equation in (32).

Equation (40) for the $b(n,k)$ follows easily from (35). □

The array of coefficients $a(n,k)$ for $0 \leq n, k \leq 7$, reads

1	0	0	0	0	0	0	0
1	2	0	0	0	0	0	0
1	6	11	0	0	0	0	0
2	21	84	131	0	0	0	0
6	92	588	1908	2666	0	0	0
24	490	4380	22020	62860	81534	0	0
120	3084	35790	246480	1075020	2823180	3478014	0
720	22428	322224	2838570	16775640	66811920	165838848	196993194

and the $b(n,k)$ for $1 \le n \le 7$ and $0 \le k \le 7$ are

1	1	0	0	0	0	0	0
1	4	5	0	0	0	0	0
2	15	42	47	0	0	0	0
6	68	312	732	758	0	0	0
24	370	2420	8880	18820	18674	0	0
120	2364	20370	103320	335580	673140	654834	0
720	17388	187656	1227450	5421360	16485000	32215008	31154346

Theorem 4.10. (a) *The coefficients $b(n,k)$ are integers.*
(b) *$a(n,k) \ge 0$ and $b(n,k) \ge 0$.*
(c) *$a(n,k-1) \le a(n,k)$ for $1 \le k \le n$.*
(d) *For $n \ge 1$, $a(n,0) = (n-1)!$.*

Proof. (a) We have proved in Theorem 4.5 that the numbers $a(n,k)$ are integers, so that from (40), the coefficients $b(n,k)$ are also integers.

(b) We proceed by induction on n. Assuming that we have proved that $a(m,k)$ and $b(m,k)$ are positive for $m < n$, it follows from (32) that $a(n,k)$ and $b(n,k)$ are positive for $0 \le k < n$. Then (34) implies that $b(n,n) \ge 0$, and (33) with $k = n$ proves that $a(n,n) \ge 0$.

(c) This is a simple consequence of (33).

(d) The equation follows from (39) by induction. □

Theorem 4.11. *By means of the rule in Theorem 4.9, one may compute all coefficients $a_{n,k}$ of the polynomials $P_n(y)$ for $1 \le n \le N$ in $\mathcal{O}(N^2)$ coefficient operations.*

Proof. We count the operations needed, following the indications in Theorem 4.9, to compute every $a_{n,k}$ for $0 \le n \le N$ and $0 \le k \le n$.

First we must compute the numbers $\binom{m}{j}$ for $0 \le m \le N-2$. Using the scheme of the usual triangle, we need to carry out $\sum_{k=1}^{N-3} k$ additions, which involves $(N-2)(N-3)/2$ operations.

The numbers B_n must now be computed for $2 \le n \le N$ by means of the formula

$$B_n = n * (B_{n-1} + (n-1) * A_{n-1}).$$

Each B_n requires 4 operations, therefore a total of $4(N-1)$ operations are needed. We compute the A_n for $2 \le n \le N$ using the formula

$$A_n = n * n * A_{n-1} + n * B_{n-1} - (n-1) * \\ * \sum_{k=1}^{n-1} \binom{n-2}{k-1} * \{k * (k-1) * A_{k-1} - A_k + k * B_{k-1}\} * A_{n-k-1}.$$

Hence A_n requires $7+\sum_{k=1}^{n-1} 8 = 8n-1$ operations. All A_n together take $\sum_{n=2}^{N}(8n-1) = 4N^2+3N-7$ operations. These numbers are the $a_{n,n}$. The $a_{0,k}$ and $a_{1,k}$ require no operations. Finally we compute for $0 \le k < n$

$$a_{n,k} = n * \bigl\{a_{n-1,k-1} + (n-1) * a_{n-1,k}/(n-k)\bigr\}.$$

Therefore, $a_{n,k}$ takes 6 operations. For each n, every $a_{n,k}$ for $1 \le k < n$ takes $6(n-1)$ operations. And each $a_{n,k}$ for $2 \le n \le N$ takes $\sum_{n=2}^{N} 6(n-1) = 3N(N-1)$.

The total cost in number of operations is therefore

$$\frac{(N-2)(N-3)}{2} + 4(N-1) + 4N^2 + 3N - 7 + 3N(N-1) = \\ = \frac{1}{2}(15N^2+3N-16).$$

□

5. Bounds for the asymptotic expansion.

5.1. **The sequence (a_n).** First we define a sequence of numbers as the coefficients of a formal expansion in A.

Lemma 5.1. *There exists a sequence of integers (a_n) such that*

$$\log\Bigl(1 - \sum_{n=1}^{\infty} \frac{n!}{x^n}\Bigr)^{-1} = \sum_{n=1}^{\infty} \frac{a_n}{n}\frac{1}{x^n}. \tag{41}$$

The coefficients may be computed by the recursion

$$a_1 = 1, \quad a_n = n!\cdot n + \sum_{k=1}^{n-1} k!\, a_{n-k}. \tag{42}$$

Proof. It is clear that $u = 1 - \sum n! x^{-n} \in U \subset A$, so that $u^{-1} \in U$ and $\log u^{-1}$ are well defined. To obtain the recursion we differentiate (41) to obtain

$$-\sum_{n=1}^{\infty} \frac{n\cdot n!}{x^{n+1}} = -\Bigl(\sum_{n=1}^{\infty}\frac{a_n}{x^{n+1}}\Bigr)\Bigl(1-\sum_{n=1}^{\infty}\frac{n!}{x^n}\Bigr).$$

Equation (42) is obtained by equating the coefficients of x^{-n-1}. The recurrence (42) proves that a_n is a natural number for each $n \ge 1$. □

The first terms of the sequence $(a_n)_{n=1}^{\infty}$ are

$$1, 5, 25, 137, 841, 5825, 45529, 399713, 3911785, 42302225, \dots$$

Lemma 5.2. *For each natural number n we have*

$$a_n \le 2n\cdot n!. \tag{43}$$

Proof. We may verify this property for a_1, a_2, a_3 and a_4 directly. For $n \ge 4$ we proceed by induction. Assume the inequality for a_k with $k < n$, so that by (42)

$$1 \le \frac{a_n}{n!\cdot n} \le 1 + \sum_{k=1}^{n-1} \frac{a_{n-k}}{(n-k)!\cdot(n-k)}\frac{n-k}{n}\binom{n}{k}^{-1} \le \\ \le 1 + 2\Bigl(\frac{1}{n} + \sum_{k=2}^{n-2}\binom{n}{k}^{-1} + \frac{1}{n^2}\Bigr) \le 1 + 2\Bigl(\frac{1}{n}+\frac{1}{n^2}+(n-3)\frac{2}{n(n-1)}\Bigr).$$

For $n \geq 4$, it is easy to see that this is ≤ 2. □

Lemma 5.3. *For each natural number N there is a positive constant c_N such that*

$$x\Big(1-\sum_{n=1}^{N}\frac{n!}{x^n}\Big)\geq 1,\qquad x\geq c_N. \tag{44}$$

Proof. It is clear that the left-hand side of (44) is increasing and tends to $+\infty$ when $x\to+\infty$, from which the existence of c_N is clear.

The value of c_N may be determined as the solution of the equation

$$x\Big(1-\sum_{n=1}^{N}\frac{n!}{x^n}\Big)=1,\qquad x>1. \tag{45}$$

In this way we found the following values.

c_1	2	c_6	4.15213	c_{11}	5.61664	c_{20}	8.70335
c_2	2.73205	c_7	4.43119	c_{12}	5.93649	c_{30}	12.34925
c_3	3.20701	c_8	4.71412	c_{13}	6.26449	c_{40}	16.03475
c_4	3.56383	c_9	5.00517	c_{14}	6.59947	c_{50}	19.72833
c_5	3.86841	c_{10}	5.30597	c_{15}	6.94035	c_{60}	23.42351

□

Remark 5.4. Notice that for $x\geq c_N$ the sum in (44) is positive and less than 1.

Proposition 5.5. *For each natural number N there exists $d_N>0$ such that, for $x\in\mathbb{C}$ with $|x|\geq d_N$, there exists θ with $|\theta|\leq 1$ such that*

$$\log\Big(1-\sum_{n=1}^{N}\frac{n!}{x^n}\Big)^{-1}=\sum_{n=1}^{N}\frac{a_n}{n}\frac{1}{x^n}+\theta\frac{a_{N+1}}{N+1}\frac{1}{x^{N+1}},\qquad |x|>d_N. \tag{46}$$

Proof. By comparing the expansions (41) and

$$\log\Big(1-\sum_{n=1}^{N}\frac{n!}{x^n}\Big)^{-1}=\sum_{n=1}^{N}\frac{a_n}{n}\frac{1}{x^n}+\sum_{n=N+1}^{\infty}\frac{b_n}{n}\frac{1}{x^n} \tag{47}$$

it is clear that $\frac{b_{N+1}}{N+1}+(N+1)!=\frac{a_{N+1}}{N+1}$, so that $b_{N+1}<a_{N+1}$.

The above expansion is convergent for all sufficiently large $|x|$, so that

$$\sum_{n=N+1}^{\infty}\frac{b_n}{n}\frac{1}{x^n}=\frac{b_{N+1}}{N+1}\frac{1}{x^{N+1}}g_N(x)$$

where $\lim_{x\to\infty}g_N(x)=1$. Hence there exist sufficiently large d_N such that

$$|b_{N+1}g_N(x)|<a_{N+1},\qquad |x|>d_N.$$

This ends the proof of the existence of d_N.

We have

$$\Big(\log\Big(1-\sum_{n=1}^{N}\frac{n!}{x^n}\Big)^{-1}-\sum_{n=1}^{N}\frac{a_n}{n}\frac{1}{x^n}\Big)\frac{(N+1)x^{N+1}}{a_{N+1}}=$$

$$=\frac{(N+1)}{a_{N+1}}\sum_{n=N+1}^{\infty}\frac{b_n}{n}\frac{1}{x^{n-N-1}}.$$

Since all a_n and b_n are positive, this is a decreasing function for $x \to +\infty$, and the lowest value of d_N will be the unique solution of

$$\Bigl(\log\Bigl(1 - \sum_{n=1}^{N} \frac{n!}{x^n}\Bigr)^{-1} - \sum_{n=1}^{N} \frac{a_n}{n}\frac{1}{x^n}\Bigr)\frac{(N+1)x^{N+1}}{a_{N+1}} = 1.$$

We obtain the following table of values

d_1	1.03922	d_6	4.54145	d_{11}	5.73661	d_{20}	8.73298
d_2	2.38568	d_7	4.75734	d_{12}	6.03061	d_{30}	12.37349
d_3	3.33232	d_8	4.97336	d_{13}	6.33969	d_{40}	16.05983
d_4	3.92171	d_9	5.20626	d_{14}	6.66091	d_{50}	19.75448
d_5	4.28707	d_{10}	5.46090	d_{15}	6.99175	d_{60}	23.45053

□

Remark 5.6. The numbers d_N in Lemma 5.5 are very similar to the numbers c_N of Lemma 5.3. This is no more than an experimental observation, but since the c_N numbers are easy to compute and d_N are somewhat elusive, it has been useful to start from c_N as an approximation to d_N in order to compute d_N.

5.2. **Some inequalities.**

Lemma 5.7. *For $u \ge 2$ we have* $\log \operatorname{ali}(u) \le 2\log u$. *For $u \ge e^2$ we have* $\operatorname{ali}(u) \le 2u\log u$.

Proof. The first inequality is equivalent to $\operatorname{ali}(u) \le u^2$. Since $\operatorname{li}(x)$ is strictly increasing, the inequality is equivalent to $u \le \operatorname{li}(u^2)$.

For $u > 2$ we have $\operatorname{li}(u) > \operatorname{li}(2) = 1.04516\dots$ so that

$$\operatorname{li}(u^2) = \operatorname{li}(u) + \int_u^{u^2} \frac{dt}{\log t} > 1 + \frac{u^2 - u}{\log u^2}.$$

Hence, our inequality follows from $\frac{u^2-u}{\log u^2} > u - 1$, that is from $u > 2\log u$. However, this last inequality is certainly true for $u > 2$.

The second inequality is equivalent to $u \le \operatorname{li}(2u\log u)$ and has a similar easy proof. □

Lemma 5.8. *For all integers $n \ge 1$ we have*

$$\int_{e^{f_n}}^{u} \frac{(\log\log t)^n}{\log^{n+1} t}\,dt \le 4u\frac{(\log\log u)^n}{\log^{n+1} u}, \qquad (u \ge e^{f_n}) \tag{48}$$

where $f_n = 4(n+1)/3$.

Proof. Notice that $f_n > 1$. For $t \ge e$ the function $\log\log t$ is positive and increasing so that

$$\int_{e^{f_n}}^{u} \frac{(\log\log t)^n}{\log^{n+1} t}\,dt \le (\log\log u)^n \int_{e^{f_n}}^{u} \frac{dt}{\log^{n+1} t}.$$

It remains to be shown that

$$\int_{e^{f_n}}^{u} \frac{dt}{\log^{n+1} t} \le \frac{4u}{\log^{n+1} u}, \qquad (u \ge e^{f_n}).$$

Replacing u by e^x this is equivalent to

$$\int_{f_n}^{x} \frac{e^t}{t^{n+1}}\,dt \le \frac{4e^x}{x^{n+1}}, \qquad (x \ge f_n).$$

For the function

$$G(x) := \frac{4e^x}{x^{n+1}} - \int_{f_n}^x \frac{e^t}{t^{n+1}}\,dt$$

we have

$$G'(x) = \frac{e^x}{x^{n+1}}\Bigl(4 - \frac{4(n+1)}{x} - 1\Bigr)$$

so that for $x > 4(n+1)/3$ we obtain $G'(x) > 0$. Since $G(f_n) > 0$ we have $G(x) > 0$ for all $x > f_n$. □

Theorem 5.9. *The polynomials $P_n(y)$ defined in* (21) *satisy the inequalities*

$$|P_n(y)| \le 3\cdot n!\, y^n, \qquad y \ge 2, \quad n \ge 1 \tag{49}$$

and $|P_0(y)| \le y$ for $y \ge 2$.

Proof. Since $P_0(y) = y - 1$, the second assertion is trivial.

Given $r > 0$, for each polynomial $P(x) = \sum_{n=0}^N a_n x^n$ we define

$$\|P\| = \sum_{n=0}^N |a_n| r^n.$$

It is easy to show that

$$\|P + Q\| \le \|P\| + \|Q\|, \qquad \|PQ\| \le \|P\|\cdot\|Q\|$$

and that for the derivative of a polynomial of degree $\le N$

$$\|P'\| = \sum_{n=0}^N n|a_n|r^{n-1} \le \frac{N}{r}\sum_{n=0}^N |a_n| r^n = \frac{N}{r}\|P\|.$$

For $y \ge r$ we have the inequality

$$|P(y)| = \Bigl|\sum_{n=0}^N a_n y^n\Bigr| \le \sum_{n=0}^N |a_n| y^n \le \sum_{n=0}^N |a_n| r^n (y/r)^n \le (y/r)^N \|P\|.$$

Hence, our Theorem follows if it can be shown that for $n \ge 1$ we have $\|P_n\| \le 3\cdot 2^n\, n!$ (for $r = 2$).

Define $S_n := \|P_n\|$. By (30) we have $S_n = -P_n(-2)$, and it can be shown that $S_n \le 3\cdot 2^n\, n!$ for $0 \le n \le 15$.

For $n > 15$ it follows from (26) and the aforementioned properties of $\|P\|$ that

$$S_n \le nS_{n-1} + \frac{n}{2}S_{n-1} + \frac{1}{n}\sum_{k=1}^{n-1} k\Bigl((k-1)S_{k-1} + S_k + \frac{k}{2}S_{k-1}\Bigr)S_{n-k-1}.$$

It follows that $S_n \le T_n$ where $T_n := S_n \le 3\cdot 3^n\, n!$ for $0 \le n \le 15$, and that for $n > 15$

$$T_n := \frac{3n}{2}T_{n-1} + \frac{1}{n}\sum_{k=1}^{n-1}\Bigl(kT_k + \frac{k(3k-2)}{2}T_{k-1}\Bigr)T_{n-k-1}.$$

Now we proceed by induction. For $n > 15$ and assuming that we have proved $T_k \le 3 \cdot 2^k\, k!$ for $k < n$, we obtain

$$T_n \le \frac{9n}{2} 2^{n-1}(n-1)! + \\ + \frac{9}{n}\sum_{k=1}^{n-1}\Bigl(k 2^k k! + \frac{k(3k-2)}{2} 2^{k-1}(k-1)!\Bigr) 2^{n-k-1}(n-k-1)!.$$

Hence

$$\frac{T_n}{3\cdot 2^n\, n!} \le \frac{3}{4} + \frac{3}{n}\sum_{k=1}^{n-1}\Bigl(\frac{k\cdot k!(n-k-1)!}{2\cdot n!} + \frac{(3k-2)k!(n-k-1)!}{8\cdot n!}\Bigr) \le \\ \le \frac{3}{4} + \frac{3}{8n^2}\sum_{k=1}^{n-1}\frac{7k-2}{\binom{n-1}{k}} \le \frac{3}{4} + \frac{3(7n-9)}{8n^2} + \frac{3}{8n^2}\sum_{k=1}^{n-2}\frac{7k-2}{\binom{n-1}{k}}.$$

Therefore, by using the symmetry of the combinatorial numbers, we obtain

$$\frac{T_n}{3\cdot 2^n\, n!} \le \frac{3}{4} + \frac{3(7n-9)}{8n^2} + \frac{3}{16n^2}\sum_{k=1}^{n-2}\frac{7n-11}{\binom{n-1}{k}} \le \\ \le \frac{3}{4} + \frac{3(7n-9)}{8n^2} + \frac{3}{16n^2}\cdot(n-2)\frac{7n-11}{n-1} \le \\ \le \frac{3}{4} + \frac{3(7n-9)}{8n^2} + \frac{3}{16n^2}\cdot(7n-11) = 1 - \frac{n(4n-63)+87}{16n^2} < 1$$

for $n > 15$. □

Corollary 5.10. *We have*

$$|P_{n-1}(y)| \le n!\, y^n, \qquad n \ge 1, \quad y \ge 2. \tag{50}$$

Proof. This follows easily from the above Theorem. □

5.3. **Main inequalities.** To simplify our formulae we introduce some notation. First we set $r_n := 3\cdot n!$ so that, for $n \ge 1$, we have $|P_n(y)| \le r_n y^n$ when $y > 2$.

Let c_n and d_n be the constants introduced in Lemma 5.3 and Proposition 5.5. Let α_n be equal to $\max(e, c_n, d_n)$ and let $\beta_n \ge e$ be the solution of the equation

$$\frac{x}{\log x} = \alpha_n. \tag{51}$$

(The function $\frac{t}{\log t}$ is increasing for $t \ge e$).

Finally, define $x_n := \max(\beta_n, f_n, e^2)$, where f_n is defined in Lemma 5.8.

Proposition 5.11. *Let x be a real number such that $x \ge x_n$, and set $y := \log x$. Then*

$$y \ge 2, \qquad x \ge c_n y, \qquad x \ge d_n y, \qquad x \ge f_n.$$

Proof. Since $x \ge x_n = \max(\beta_n, f_n, e^2)$ we have $x \ge e^2$, so that $y = \log x \ge 2$.

We also have $x \ge \beta_n \ge e$. Since $\frac{t}{\log t}$ is an increasing function for $t \ge e$ we obtain $\frac{x}{\log x} \ge \frac{\beta_n}{\log \beta_n} = \alpha_n = \max(e, c_n, d_n)$. Therefore, $\frac{x}{y} \ge c_n$ and $\frac{x}{y} \ge d_n$ as required. □

We insert a table of the constants x_n.

x_1	7.38906	x_6	10.81135	x_{11}	16.00000	x_{20}	29.57923
x_2	7.38906	x_7	11.70187	x_{12}	17.33333	x_{30}	47.86556
x_3	7.38906	x_8	12.60164	x_{13}	18.66667	x_{40}	67.69154
x_4	8.29874	x_9	13.58167	x_{14}	20.00000	x_{50}	88.57644
x_5	9.77283	x_{10}	14.66667	x_{15}	21.42740	x_{60}	110.29065

For each natural number N we set

$$W_N = 1 + \sum_{n=1}^{N} \frac{P_{n-1}(y)}{x^n} \tag{52}$$

and frequently we write $W := W_N$ when N is fixed.

Proposition 5.12. *For $N \geq 1$ let $W = W_N$ (as in (52)). Then for $x \geq x_N$ and $y = \log x$ there exists θ with $|\theta| \leq 1$ such that*

$$W + xW + xW_x + W_y - x - y - \log W = \theta \cdot r_{N+1} \frac{y^N}{x^N}. \tag{53}$$

Proof. Denote by $T = T(x, y)$ the value of $W + xW + xW_x + W_y - x - y - \log W$. Then we have

$$T = (1+x)\sum_{n=0}^{N} \frac{P_{n-1}(y)}{x^n} - \sum_{n=1}^{N} \frac{nP_{n-1}(y)}{x^n} + \sum_{n=1}^{N} \frac{P'_{n-1}(y)}{x^n} - \\ - x - y + \log\Bigl(1 + \sum_{n=1}^{N} \frac{P_{n-1}(y)}{x^n}\Bigr)^{-1}.$$

From (24) we have the expansion

$$\log\Bigl(1 + \sum_{n=1}^{\infty} \frac{P_{n-1}(y)}{x^n}\Bigr)^{-1} = -\sum_{n=1}^{\infty} \frac{Q_n(y)}{x^n}. \tag{54}$$

From Proposition 5.11 we know that $y = \log x > 2$ and $x \geq yd_N$. From (50), for $y > 2$, we have $|P_{n-1}(y)| \leq n! y^n$ so that we have the majorant

$$\log\Bigl(1 + \sum_{n=1}^{N} \frac{P_{n-1}(y)}{x^n}\Bigr)^{-1} \ll \log\Bigl(1 - \sum_{n=1}^{N} \frac{n!}{(x/y)^n}\Bigr)^{-1} \tag{55}$$

(by considering this expression as a power series in x^{-1}, and y as a parameter). From (54) and (55), we obtain

$$\log\Bigl(1 + \sum_{n=1}^{N} \frac{P_n(y)}{x^n}\Bigr)^{-1} = -\sum_{n=1}^{N} \frac{Q_n(y)}{x^n} + S_N(x, y) \tag{56}$$

where $S_N(x, y)$ is a power series majorized by the Taylor expansion of

$$\log\Bigl(1 - \sum_{n=1}^{N} \frac{n!}{(x/y)^n}\Bigr)^{-1} - \sum_{n=1}^{N} \frac{a_n}{n} \frac{1}{(x/y)^n}$$

(compare equation (47)).

By applying Proposition 5.5 we deduce that, for $x > yd_N$, there exists θ with $|\theta| \le 1$ and

$$S_N(x,y) = \theta \frac{a_{N+1}}{N+1}\frac{y^{N+1}}{x^{N+1}}. \tag{57}$$

If we substitute (56) in the expression for T, then by Theorem 4.3, all the terms in x^{-n} with $n < N$ cancel out, and the terms in x^{-N} add up to $-P_N(y)x^{-N}$. It follows that

$$T = -\frac{P_N(y)}{x^N} + S_N(x,y). \tag{58}$$

Therefore, since $y > 2$, we have

$$|T| \le r_N \frac{y^N}{x^N} + \frac{a_{N+1}}{N+1}\frac{y^{N+1}}{x^{N+1}}$$

so that from (43) ,

$$|T| \le \frac{y^N}{x^N}\Big(3\cdot N! + 2\cdot(N+1)!\frac{\log x}{x}\Big) \le \frac{3\cdot(N+1)!y^N}{x^N} = r_{N+1}\frac{y^N}{x^N}$$

where $N \ge 1$ and $\frac{3}{2} + \frac{2\log x}{x} \le 3$ for $x \ge e^2$ are applied. □

Proposition 5.13. *For each natural number N let $u_N = e^{x_N}$. Then there exists $v_N > u_N$ such that*

$$\operatorname{li}(f_N(u)) - u = \theta\cdot 13(N+1)!\,\frac{u(\log\log u)^N}{\log^{N+1} u}, \qquad (u > v_N) \tag{59}$$

where $|\theta| \le 1$ and

$$f_N(e^x) := xe^x W_N(x,\log x) = xe^x\Big(1 + \sum_{n=1}^{N}\frac{P_{n-1}(\log x)}{x^n}\Big). \tag{60}$$

Proof. To simplify the notation, the abbreviation $W(x,y) = W_N(x,y)$ is used. Differentiating (60) we obtain

$$\frac{d}{dx}\big(\operatorname{li}(f_N(e^x)) - e^x\big) =$$
$$= \frac{1}{\log(f_N(e^x))}\Big\{e^xW + xe^xW + xe^x\Big(W_x + \frac{1}{x}W_y\Big)\Big\} - e^x.$$

Assume that $x \ge x_N$, so that $x \ge d_N\log x$ and $x \ge e^2$. We may apply (53) to obtain

$$\frac{d}{dx}\big(\operatorname{li}(f_N(e^x)) - e^x\big) = \frac{e^x}{\log(f_N(e^x))}\Big\{W + xW + xW_x + W_y\Big\} - e^x =$$
$$= \frac{e^x}{\log(f_N(e^x))}\Big\{x + \log x + \log W + \theta r_{N+1}\frac{\log^N x}{x^N}\Big\} - e^x.$$

This may be simplified to

$$\frac{d}{dx}\big(\operatorname{li}(f_N(e^x)) - e^x\big) = \frac{e^x}{\log(f_N(e^x))}\cdot\theta\frac{r_{N+1}\log^N x}{x^N}.$$

Since $x > x_N$ we have $x \geq yc_N$, so that by Lemma 5.3

$$\Big|x\Big(1+\sum_{n=1}^N \frac{P_{n-1}(\log x)}{x^n}\Big)\Big| \geq x\Big(1-\sum_{n=1}^N \frac{n!}{(x/y)^n}\Big) \geq 1$$

that is $xW_N(x,\log x) \geq 1$, so that $\log(f_N(e^x)) \geq x$. Hence, for $x \geq x_N$ (with another θ), we have

$$\frac{d}{dx}\big(\mathrm{li}(f_N(e^x)) - e^x\big) = \theta\frac{r_{N+1}\log^N x}{x^{N+1}}e^x.$$

Defining $H_N(u) := \mathrm{li}(f_N(u)) - u$ the above equation is equivalent to

$$H_N'(e^x) = \theta\frac{r_{N+1}\log^N x}{x^{N+1}}, \qquad (x \geq x_N)$$

and, since $u_N := e^{x_N}$,

$$H_N'(u) = \theta\frac{r_{N+1}(\log\log u)^N}{\log^{N+1} u}, \qquad (u \geq u_N).$$

Lemma 5.8 can be applied since $x_N \geq f_N$, so that $u \geq u_N \geq e^{f_N}$. Hence, integrating over the interval (u_N, u) we get

$$H_N(u) = H_N(u_N) + \theta\frac{4r_{N+1}\,u(\log\log u)^N}{\log^{N+1} u}, \qquad (u \geq u_N).$$

The function $(N+1)!\frac{u}{\log^{N+1}u}\cdot(\log\log u)^N$ is increasing (as product of two positive increasing functions) for $u > e^{f_n}$, so that there exists $v_N > u_N$ for which this function is greater than $H_N(u_N)$, so that

$$H_N(u) = \theta\frac{13\cdot(N+1)!\,u(\log\log u)^N}{\log^{N+1} u}, \qquad (u \geq v_N).$$

□

Remark 5.14. For the values of n appearing in our tables, the equality $u_n = v_n$ holds, since, in these cases,

$$H_n(u_n) \leq \frac{(n+1)!\,u_n(\log\log u_n)^n}{\log^{n+1} u_n}.$$

Lemma 5.15. *For any natural number N, and $u > e^{x_N}$ we have* $\log f_N(u) < 2\log u$.

Proof. First observe that the hypothesis $u > e^{x_N}$ implies (with $u = e^x$) that $x > x_N$, so that $\log x > 2$ and $x > c_N \log x$. (Proposition 5.11).

The inequality $\log f_N(u) < 2\log u$ is equivalent to $f_N(u) < u^2$, and together with $u = e^x$ it is equivalent to

$$xe^x\Big(1+\sum_{n=1}^N \frac{P_{n-1}(\log x)}{x^n}\Big) < e^{2x}.$$

From Corollary 5.10, since $x \geq e^2$ and $x \geq c_N y$, and by Remark 5.14,

$$x\Big(1+\sum_{n=1}^N \frac{P_{n-1}(\log x)}{x^n}\Big) \leq x\Big(1+\sum_{n=1}^N \frac{n!}{(x/y)^n}\Big) < x\Big(1-\sum_{n=1}^N \frac{n!}{(x/y)^n}\Big)^{-1}.$$

Hence our inequality follows from

$$x < \frac{ye^x}{x}\cdot\frac{x}{y}\Bigl(1-\sum_{n=1}^{N}\frac{n!}{(x/y)^n}\Bigr),\qquad y=\log x.$$

Since we assume that $x \geq yc_N$, the second factor is greater than 1, so that

$$\frac{ye^x}{x}\cdot\frac{x}{y}\Bigl(1-\sum_{n=1}^{N}\frac{n!}{(x/y)^n}\Bigr) > \frac{ye^x}{x}$$

Finally, it is easy to prove that $e^x\log x > x^2$ for $x > e^2$. □

The asymptotic expansion with bounds can now be proved.

Theorem 5.16. *For each integer $N \geq 1$*

$$\text{ali}(u) = f_N(u) + 26\theta(N+1)!\,u\Bigl(\frac{\log\log u}{\log u}\Bigr)^N,\qquad (u \geq v_N), \tag{61}$$

where v_N is the number defined in Proposition 5.13 .

Proof. Since $\text{li}(\text{ali}(u)) = u$, Proposition 5.13 yields, for $u > v_N$,

$$\text{li}(f_N(u)) - \text{li}(\text{ali}(u)) = \int_{\text{ali}(u)}^{f_N(u)}\frac{dt}{\log t} = 13\theta(N+1)!\,u\frac{(\log\log u)^N}{\log^{N+1}u}.$$

Since $v_N \geq u_N = e^{x_N}$, $u \geq v_N$ implies $\log u \geq 2$, hence $u \geq 2$.

From Lemma 5.7, $\log\text{ali}(u) \leq 2\log u$, for $u > 2$. Analogously, Lemma 5.15 implies that $\log f_N(u) \leq 2\log u$, for $u > e^{x_N}$. Therefore, for $u > v_N$, we have

$$\frac{|\,\text{ali}(u) - f_N(u)|}{2\log u} \leq \Bigl|\int_{\text{ali}(u)}^{f_N(u)}\frac{dt}{\log t}\Bigr|.$$

It follows that there exists θ' with $|\theta'| \leq 1$ such that

$$\text{ali}(u) - f_N(u) = \theta'(2\log u)\int_{\text{ali}(u)}^{f_N(u)}\frac{dt}{\log t}$$

and the result follows easily. □

The actual error appears to be much smaller than that given in Theorem 5.16. However, as usual with asymptotic expansions, having a true bound allows realistic bounds to be given of the remainder for specific values of N.

The true error after N terms of an asymptotic expansion, while the terms are decreasing in magnitude, is often of the size of the first omitted term. In our case, the magnitude of the term $P_N(\log x)x^{-N-1}$ depends on the polynomial $P_N(\log x)$.

Numerically, it appears that for $n \geq 3$:

$$|P_n(y)| \leq \Bigl(\frac{n}{e\log n}\Bigr)^n y^n,\qquad (y > 2\log n) \tag{62}$$

although we have not been able to prove this.

From Theorem 5.16, more realistic bounds can be obtained for the first values of N. This is done in the following Theorem.

Theorem 5.17. *For $2\le N\le 11$, we have*

$$\frac{\operatorname{ali}(e^x)}{xe^x}=1+\sum_{n=1}^N\frac{P_{n-1}(\log x)}{x^n}+\theta\cdot 20\cdot\Big(\frac{N}{e\log N}\Big)^N\cdot\frac{\log^N x}{x^{N+1}},\qquad (x>z_N)\tag{63}$$

where

$$z_2=1.50,\quad z_3=2.34,\quad z_4=3.32,\quad z_5=4.33,\quad z_6=5.36,$$
$$z_7=6.39,\quad z_8=7.43,\quad z_9=8.46,\quad z_{10}=9.50,\quad z_{11}=10.53.$$

Proof. By taking $N=10$ in Theorem 5.16, we have, for $u=e^x>e^{x_{10}}$, (recall also Remark **??**)

$$\frac{\operatorname{ali}(e^x)}{xe^x}=1+\sum_{n=1}^{10}\frac{P_{n-1}(\log x)}{x^n}+\theta R\frac{\log^{10}x}{x^{11}},\qquad (x>x_{10})$$

with $R=26\cdot 11!=1\,037\,836\,800$.

We compute the maximum[3] M_n of $|P_{n-1}(\log x)/\log^{n-1}x|$ for $x>x_{10}$, so that for any $2\le N\le 10$, we have

$$\frac{\operatorname{ali}(e^x)}{xe^x}=1+\sum_{n=1}^N\frac{P_{n-1}(\log x)}{x^n}+ \\ +\frac{\log^N x}{x^{N+1}}\Big(\sum_{n=N+1}^{10}\frac{P_{n-1}(\log x)}{\log^{n-1}x}\frac{\log^{n-N-1}x}{x^{n-N-1}}+\theta R\frac{\log^{10-N}x}{x^{10-N}}\Big)$$

so that

$$\frac{\operatorname{ali}(e^x)}{xe^x}=1+\sum_{n=1}^N\frac{P_{n-1}(\log x)}{x^n}+ \\ +\theta\frac{\log^N x}{x^{N+1}}\Big(\sum_{n=N+1}^{10}\frac{M_n\log^{n-N-1}x}{x^{n-N-1}}+R\frac{\log^{10-N}x}{x^{10-N}}\Big).$$

We determine a value $z_N'>x_{10}$ such that, for $x=z_N'$,

$$\Big(\sum_{n=N+1}^{10}\frac{M_n\log^{n-N-1}x}{x^{n-N-1}}+R\frac{\log^{10-N}x}{x^{10-N}}\Big)<20\Big(\frac{N}{e\log N}\Big)^N.$$

Since this is a decreasing function of x, we obtain for $x>z_N'$

$$\frac{\operatorname{ali}(e^x)}{xe^x}=1+\sum_{n=1}^N\frac{P_{n-1}(\log x)}{x^n}+\theta\cdot 20\Big(\frac{N}{e\log N}\Big)^N\frac{\log^N x}{x^{N+1}}.\tag{64}$$

We consider the function

$$\Big(\frac{\operatorname{ali}(e^x)}{xe^x}-1-\sum_{n=1}^N\frac{P_{n-1}(\log x)}{x^n}\Big)\frac{x^{N+1}}{\log^N x}$$

on the interval $(1.3,z_N')$, to determine the least value of z_N for which (64) is true.

[3] $M_2=1$, $M_3=1/2$, $M_4=1/3$, $M_5=0.250636$, $M_6=0.526887$, $M_7=1.300565$, $M_8=3.719653$, $M_9=12.070813$, $M_{10}=43.788782$. This last maximum would be much smaller if the maximum were taken from a point slightly greater than x_{10}.

In this way we find: $z'_2 = 32$ and then $z_2 = 1.5$; $z'_3 = 49.5$ and then $z_3 = 2.3395$; $z'_4 = 82$ and then $z_4 = 3.3114$; $z'_5 = 155$ and then $z_5 = 4.3237$.

If we take $N = 20$ in Theorem 5.16, we obtain $z'_6 = 113$, $z'_7 = 143$, $z'_8 = 187$, $z'_9 = 251$ $z'_{10} = 353$, $z'_{11} = 528$ from which $z_6 = 5.3514$, $z_7 = 6.3851$, $z_8 = 7.4208$, $z_9 = 8.4566$, $z_{10} = 9.4914$ and $z_{11} = 10.5251$ are obtained. □

Remark 5.18. We have proved (63) only for $2 \le N \le 11$, although something similar appears to be true for the general case. If (63) were true for all n, then for a large $u = e^x$ we could take $N \approx x$ terms in the expansion and in this way the error would be $\approx \frac{20}{xe^x}$, so that $\operatorname{ali}(u)$ could be computed with an error less than ≈ 20.

In fact, for several values of u, the terms of the expansion have been computed up to the point where these terms start to increase. Always the computation is terminated when $N \approx x$ and the error appears to be bounded. (For example, with $u = 10^{100}$, we compute 230 terms of the expansion, which coincides with $\log 10^{100} \approx 230.259$. The approximate value obtained for $\operatorname{ali}(u)$ has an absolute error equal to 40.94738, which can be compared with the fact that $\operatorname{ali}(10^{100})$ has 103 digits).

6. Applications to p_n.

6.1. **Asymptotic expansion of p_n.** Inequalities for the n-th prime number can be found in [18], [16], [12], [4]. In fact, from $\pi(x) = \operatorname{li}(x) + \mathcal{O}(r(x))$ we may obtain $p_n = \operatorname{ali}(n) + \mathcal{O}(r(n\log n)\log n)$, if $r(x)/x$ is sufficiently small. For example, in [12], it is noticed that from a result of Massias [10], it follows that

$$p_n = \operatorname{ali}(n) + \mathcal{O}(ne^{-c\sqrt{\log n}}) \tag{65}$$

so that the asymptotic expansion of $\operatorname{ali}(n)$ is also an asymptotic expansion for p_n, that is,

$$p_n = n\log n\Big(1 + \sum_{k=1}^{N} \frac{P_{k-1}(\log\log n)}{\log^k n}\Big) + \mathcal{O}\Big(n\Big(\frac{\log\log n}{\log n}\Big)^N\Big). \tag{66}$$

By assuming the Riemann hypothesis, Schoenfeld [21] has proved

$$|\pi(x) - \operatorname{li}(x)| < \frac{1}{8\pi}\sqrt{x}\log x, \qquad (x > 2657). \tag{67}$$

This result will be used to obtain (under RH) some precise bounds for p_n.

Lemma 6.1. *We have* $\sqrt{x}(\log x)^{\frac52} < \operatorname{ali}(x)$ *for* $x > 94$.

Proof. The inequality is equivalent to

$$\operatorname{li}(\sqrt{x}(\log x)^{\frac52}) < x.$$

Differentiating the function $f(x) := x - \operatorname{li}(\sqrt{x}(\log x)^{\frac52})$ we get

$$f'(x) = 1 - \frac{1}{\log(\sqrt{x}(\log x)^{\frac52})}\Big(\frac{\log^{5/2} x}{2\sqrt{x}} + \frac{5\log^{3/2} x}{2\sqrt{x}}\Big).$$

Hence this derivative is positive if and only if

$$\log^{3/2} x\Big(\frac{\log x}{2} + \frac52\Big) < \sqrt{x}\Big(\frac{\log x}{2} + \frac52\log\log x\Big).$$

For $x > 94$ we have $(\log x)^{\frac32} < \sqrt{x}$, so that $f'(x) > 0$ for $x > 94$.

Finally, one may verify that $f(94) > 0$. Hence $f(x) > 0$ for $x > 94$. □

Theorem 6.2. *The Riemann hypothesis is equivalent to the assertion*

$$|p_n - \operatorname{ali}(n)| < \frac{1}{\pi}\sqrt{n}\log^{\frac52} n \qquad \text{for all } n\ge 11. \tag{68}$$

Proof. First we assume the Riemann Hypothesis and prove (68). Let $r(x) := \frac{1}{8\pi}\sqrt{x}\log x$, $f(x) := \operatorname{li}(x) - r(x)$, and $g(x) := \operatorname{li}(x) + r(x)$. For $x > 1$ we have $f(x) < \operatorname{li}(x) < g(x)$, where the three functions are strictly increasing. From (67) for $x > 2657$, we also have $f(x) < \pi(x) < g(x)$.

The inverse functions satisfy $g^{-1}(y) < \operatorname{ali}(y) < f^{-1}(y)$, and if $y = n > \pi(2657) = 384$ is a natural number, then $g^{-1}(n) < p_n < f^{-1}(n)$. It follows that the distance from $\operatorname{ali}(n)$ to p_n is bounded by

$$|p_n - \operatorname{ali}(n)| \le \max\bigl(f^{-1}(n) - \operatorname{ali}(n), \operatorname{ali}(n) - g^{-1}(n)\bigr).$$

Hence, we have to bound $f^{-1}(y) - \operatorname{ali}(y)$ and $\operatorname{ali}(y) - g^{-1}(y)$.

We consider y as a parameter and set $\alpha = \operatorname{ali}(y)$, so that $\operatorname{li}(\alpha) = \operatorname{li}(\operatorname{ali}(y)) = y$.

Consider the function $u(\xi) := f(\xi) - \operatorname{li}(\alpha) = f(\xi) - y$, which is strictly increasing and satisfies

$$u(\xi) = \operatorname{li}(\xi) - r(\xi) - \operatorname{li}(\alpha) = \int_\alpha^\xi \frac{dt}{\log t} - r(\xi).$$

Therefore, $u(\alpha) = -r(\alpha) < 0$ and

$$u(f^{-1}(y)) = f(f^{-1}(y)) - \operatorname{li}(\alpha) = y - \operatorname{li}(\operatorname{ali}(y)) = 0.$$

If a point b is found where $u(b) > 0$, then $\alpha < f^{-1}(y) < b$, so that $b-\alpha > f^{-1}(y)-\alpha$ and one of the required bounds is obtained.

Therefore, we try $b = \alpha + c\sqrt{y}(\log y)^{\frac52}$ with $c < 1$. We have

$$u(\alpha + c\sqrt{y}(\log y)^{\frac52}) = \int_\alpha^{\alpha + c\sqrt{y}(\log y)^{\frac52}} \frac{dt}{\log t} - r(\alpha + c\sqrt{y}(\log y)^{\frac52}) > \\ > \frac{c\sqrt{y}(\log y)^{\frac52}}{\log(\alpha + c\sqrt{y}(\log y)^{\frac52})} - r(\alpha + c\sqrt{y}(\log y)^{\frac52}).$$

From Lemma 6.1 for $y > 94$, we have $\sqrt{y}(\log y)^{\frac52} < \operatorname{ali}(y) = \alpha$, so that

$$u(\alpha + c\sqrt{y}(\log y)^{\frac52}) > \frac{c\sqrt{y}(\log y)^{\frac52}}{\log(2\alpha)} - r(2\alpha) \\ = \frac{c\sqrt{y}(\log y)^{\frac52} - \frac{1}{8\pi}\sqrt{2\alpha}\log^2(2\alpha)}{\log(2\alpha)}.$$

We want to show that this expression is positive. For $y > 94$, we have $\alpha = \operatorname{ali}(y) < 2y\log y$ (by Lemma 5.7), so that $\alpha < 2y\log y < 4y\log y < y^2$ (for $y > 94$), which yields (with $c = 1/\pi$)

$$c\sqrt{y}(\log y)^{\frac52} - \frac{1}{8\pi}\sqrt{2\alpha}\log^2(2\alpha) > \\ > c\sqrt{y}(\log y)^{\frac52} - \frac{1}{8\pi}\sqrt{4y\log y}\log^2(4y\log y) > \\ > c\sqrt{y}(\log y)^{\frac52} - \frac{1}{8\pi}\sqrt{4y\log y}\log^2(y^2)) = 0.$$

Hence, we have proved that $f^{-1}(y) - \alpha < \frac{1}{\pi}\sqrt{y}(\log y)^{5/2}$ for $y > 94$.

To bound $\alpha - g^{-1}(y)$, we consider the function $v(\xi) := g(\xi) - \mathrm{li}(\alpha) = g(\xi) - y$. Then

$$v(\xi) = r(\xi) - \int_\xi^\alpha \frac{dt}{\log t}$$

and $v(\alpha) = r(\alpha) > 0$, $v(g^{-1}(y)) = g(g^{-1}(y)) - y = 0$. If a value b is found such that $v(b) < 0$, it will follow that $\alpha - g^{-1}(y) < \alpha - b$.

Choose $b = \alpha - c\sqrt{y}(\log y)^{5/2}$ with $c = \frac{1}{\pi}$. We claim that $v(b) < 0$. We have

$$v(b) = r(b) - \int_b^\alpha \frac{dt}{\log t} < r(b) - \frac{\alpha - b}{\log \alpha}$$

and our claim will follow from $r(b)\log\alpha - c\sqrt{y}(\log y)^{5/2} < 0$. Finally, since $b < \alpha = \mathrm{ali}(y) < 2y\log y$, and by Lemma 5.7,

$$r(b)\log\alpha < \frac{1}{8\pi}\sqrt{\mathrm{ali}(y)}(\log \mathrm{ali}(y))^2 < \frac{1}{8\pi}\sqrt{2y\log y}(2\log y)^2$$

which proves our claim.

Hence, (by assuming RH), we have proved that $|p_n - \mathrm{ali}(n)| < \frac{1}{\pi}\sqrt{n}(\log n)^{\frac{5}{2}}$ for $n > 384 > 94$. By verifying all $1 \le n \le 385$, we find that the inequality holds except for $n < 11$.

The reverse implication is simple. From $|p_n - \mathrm{ali}(n)| = \mathcal{O}(n^{\frac{1}{2}+\varepsilon})$ for any $\varepsilon > 0$, we may derive that $\pi(x) = \mathrm{li}(x) + \mathcal{O}(x^{\frac{1}{2}+\varepsilon})$. It is well known that this is equivalent to the Riemann Hypothesis. □

Remark 6.3. The inequality (68) is only proved by assuming the Riemann Hypothesis, but is stronger than those contained in [18], [16], [12], [4]. Inequality (68) gives approximately half of the digits of p_n. If our conjecture that (63) is true for all N is also assumed, then the asymptotic expansion gives about half of the digits of p_n.

6.2. **Inequalities for the n-th prime.** Let

$$s_N(n) = n\log n\Big(1 + \sum_{k=1}^N \frac{P_{k-1}(\log\log n)}{\log^k n}\Big)$$

where $s_0 = n\log n$.

Cipolla noted that for $k \ge 1$, $P_k(y) = (-1)^{k+1}\frac{y^k}{k} + \cdots$, and $P_0(y) = y - 1$. Hence, except for the first term, eventually the sign of the k-th term $P_{k-1}(\log\log n)\log^{-k} n$ becomes $(-1)^k$. The asymptotic expansion implies that there exist r_N such that

$$\begin{gathered} p_n > s_0(n), \quad n > r_0, \qquad p_n > s_1(n), \quad n > r_1, \\ p_n < s_{2N}(n), \quad n > r_{2N}, \qquad p_n > s_{2N+1}(n), \quad n > r_{2N+1}. \end{gathered} \tag{69}$$

In fact, $r_0 = 2$ is the main result in [18], $r_1 = 2$ is proved in [4] and $r_2 = 688\,383$ is proved in [5]. The value of r_N for $N \ge 3$ has not been determined. See Theorem 6.4 for an estimation of r_3 by assuming RH.

The above reasoning may give the impression that the terms of the asymptotic expansion of $\mathrm{ali}(u)$ are alternating in sign, starting from the second term. However this is not true. For example, computing the first 230 terms for $\mathrm{ali}(10^{100})$, we found only three positive terms P_0/x, P_1/x^2, and P_3/x^4. In fact, the sign of the k-th term is that of $P_{k-1}(\log\log n)$. Thus we are interested in the sign of these polynomials.

The polynomials $P_N(y)$, for $1 \le N \le 23$ of odd index, have one and only one real root, which is positive. Starting from $P_1(y)$ which vanishes at $y = 2$, these roots are

$$2, \quad 4.23415, \quad 5.83131, \quad 7.43591, \quad 9.07979, \quad 10.6881, \quad 12.2538,$$
$$13.7876, \quad 15.2977, \quad 16.79, \quad 18.2683, \quad 19.7353.$$

The polynomials P_2, P_4 and P_6 have no real roots, and all $P_8, \dots, P_{22}$ have two positive real roots. These pairs of roots are:

$$(6.4306, 8.2185), \quad (7.16158, 9.88528), \quad (7.90293, 11.4752),$$
$$(8.63359, 13.0241), \quad (9.3507, 14.5452), \quad (10.055, 16.0458),$$
$$(10.7478, 17.5307), \quad (11.4307, 19.003).$$

For example, $P_9(\log\log n)$ is positive only for $n > \exp(e^{9.07\cdots})$, which is a very big number.

The even terms at first sight appear *negative*. However $P_{10}(\log\log n)$, for example, is negative except in the interval $\exp(e^{7.16\cdots}) < n < \exp(e^{9.88\cdots})$.

In a certain sense, the inequalities (69) are the wrong inequalities. These inequalities would hold only for very large values of r_N, especially when we want a lower bound of $p_n > s_{2N+1}$ (except for the three known cases). We estimate r_3.

Theorem 6.4. *Let r_3 be the smallest number such that*

$$p_n > s_3 := n\log n + n(\log\log n - 1) + n\frac{\log\log n - 2}{\log n} - n\frac{(\log\log n)^2 - 6\log\log n + 11}{2\log^2 n}, \qquad n \ge r_3. \tag{70}$$

Then, if the Riemann Hypothesis is assumed,

$$39 \times 10^{29} < r_3 \le 39.58 \times 10^{29}.$$

Proof. By Theorem 6.2, there exists θ_1 with $|\theta_1| \le 1$ such that

$$p_n = \operatorname{ali}(n) + \theta_1 \frac{\sqrt{n}}{\pi}(\log n)^{\frac52}, \qquad n \ge 11.$$

By Theorem 5.17, with $5 \le N \le 10$ and setting $n = e^x$, $x = \log n$, and $y = \log\log n$, we have for $n > e^{z_N}$

$$\operatorname{ali}(n) = xe^x\Big(1 + \frac{y-1}{x} + \frac{y-2}{x^2} - \frac{y^2 - 6y + 11}{2x^3} + \frac{P_3(y)}{x^4} + \sum_{k=5}^{N}\frac{P_{k-1}(y)}{x^k} + \theta_2 c_N \frac{y^N}{x^{N+1}}\Big).$$

The inequality of the Theorem is obtained if

$$xe^x\Big(\frac{P_3(y)}{x^4} + \sum_{k=5}^{N}\frac{P_{k-1}(y)}{x^k} - c_N\frac{y^N}{x^{N+1}}\Big) - \frac{e^{x/2}}{\pi}x^{\frac52} > 0.$$

This is equivalent to

$$P_3(y) + \sum_{k=5}^{N} P_{k-1}(y)e^{-(k-4)y} > c_N y^N e^{-(N-3)y} + \frac{1}{\pi}e^{\frac{11}{2}y}e^{-\frac12 e^y}.$$

Since $P_3(y) \to +\infty$ and all the other terms tend to 0 as $y \to +\infty$, it is clear that the inequality is true for $y > y_0$. With $N = 10$, we find $y_0 = 4.254946453\dots$ The inequality $p_n < s_n$ is true for $n \geq 3.95702224148845656 \times 10^{30}$. This proves that $r_3 \leq 39.58 \times 10^{29}$.

In order to show that $r_3 > 39 \times 10^{29}$, we directly show that, for $n = 39 \times 10^{29}$, the opposite inequality $p_n < s_3$ is obtained.

We compute

$$s_3 = 2.87527\,18639\,02974\,79681\,42399\,35057\,89294\,02005\,87915 \times 10^{32}.$$

Now we can compute $\operatorname{ali}(n)$, for which we already have obtained a good approximation through the asymptotic expansion, and then apply the Newton method

$$\operatorname{ali}(n) = 2.87527\,18639\,02495\,21516\,14800\,14732\,45414\,39731 \times 10^{32}.$$

Therefore, from Theorem 6.2, we obtain $p_n < \operatorname{ali}(n) + \frac{1}{\pi}\sqrt{n}\log^{\frac{5}{2}} n$, so that

$$p_n < 2.87527\,18639\,02756\,97808\,39055\,05640\,30082\,86370\,11482 \times 10^{32}$$

and we can conclude that $p_n < s_3$. □

References

[1] E. Cesàro, *Sur une formule empirique de M. Pervouchine*, C. R. Math. Acad. Sci. Paris **119** (1894), 848–849. (also in [2]).

[2] E. Cesàro, *Opere Scelte*, Vol. I, Parte Seconda, Edizioni Cremonese, Roma, 1965.

[3] M. Cipolla, *La determinazione assintotica dell' n^{imo} numero primo*, Rend. Accad. Sci. Fis-Mat. Napoli (3) **8** (1902), 132–166.

[4] P. Dusart, *The k-th prime is greater than $k(\log k + \log\log k - 1)$ for $k \geq 2$*, Math. Comp. **68** (1999), 411-415.

[5] P. Dusart, *Estimates of some functions over primes without R.H.*, arXiv 1002:0442 (2010). http://front.math.ucdavis.edu/1002.0442.

[6] A. E. Ingham, *The Distribution of Prime Numbers*, Cambridge University Press, 1932.

[7] E. Landau, *Über die Multiplikation Dirichlet'scher Reihen*, Rend. Circ. Matem. Palermo, **24** (1907), 81–159.

[8] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, Teubner, Leipzig 1909. [Reprint: Chelsea 1953].

[9] S. Lang, *Algebra*, Revised 3rd edition, Springer, New York 2002.

[10] J.-P. Massias, *Ordre maximum d'un élément du groupe symétrique et applications*, Thèse de 3ème cycle, Limoges, France, 1985.

[11] J.-P. Massias, J. L. Nicolas & G. Robin, *Evaluation asymptotique de l'ordre maximum d'un élément du groupe symétrique*, Acta Arithmetica **50** (1988), 221–242

[12] J.-P. Massias & G. Robin, *Bornes effectives pour certaines fonctions concernant les nombres premiers*, J. Théor. Nombres Bordeaux **8** (1996), 215–242.

[13] H. L. Montgomery & R. C. Vaughan, *Multiplicative Number Theory: I Classical Theory*, Cambridge University Press, 2006.

[14] I. M. Pervushin, *Les formules pour la détermination approximative des nombres premiers, de leur summe et de leur différence d'aprés le numéro de ces nombres*, Nachrichten der physiko-mathematischen Gesellschaft der Kaiserlichen Universität zu Kasan, (2) **4** (1894), 94–96.

[15] I. M. Pervushin, *Les formules pour la détermination approximative des nombres premiers, de leur summe et de leur différence d'aprés le numéro de ces nombres*, Verhandlungen des ersten internationalen Mathematiker-Kongresses in Zürich vom 9 bis 11 August 1897, Leipzig, Teubner, 1898.

[16] G. Robin, *Estimation de la fonction de Tchebychef θ sur le k-ième nombre premier et grandes valeurs de la fonction $\omega(n)$, nombre de diviseurs premiers de n*, Acta Arithmetica **42** (1983), 367–389.

[17] G. Robin, *Permanence de relations de récurrence dans certains développements asymptotiques*, Publ. Inst. Math. (Beograd) (N.S.) **43(57)** (1988), 17–25.

[18] B. Rosser, *The n-th prime is greater than* $n \log n$, Proc. London Math. Soc. (2) **45** (1939), 21–44.

[19] B. Rosser & L. Schoenfeld, *Sharper bounds for the Chebyshev Functions* $\theta(x)$ *and* $\psi(x)$, Math. Comp. **29** (1975), 243–269.

[20] B. Salvy, *Fast computation of some asymptotic functional inverses*, J. Symbolic Comput. **17** (1994), 227–236.

[21] L. Schoenfeld, *Sharper Bounds for the Chebyshev Functions* $\theta(x)$ *and* $\psi(x)$. *II*, Math. Comp. **30** (1976), 337-360.

[22] G. Torelli, *Sulla totalità dei numeri primi fino a un limite assegnato*, Atti della Reale Accademia delle Scienze Fisiche e Matematiche, Napoli (2) **11** (1901), 1–222.

Facultad de Matemáticas, Universidad de Sevilla,
Apdo. 1160, 41080-Sevilla, Spain
E-mail address: arias@us.es

51 Route de Laborie, Foulayronnes 47510 (Lot-et-Garonne) France
E-mail address: toulisse.jeremy@voila.fr