arXiv:1207.0603v1 [math.NT] 3 Jul 2012

Maximal product of primes whose sum is bounded

Marc Deléglise and Jean-Louis Nicolas *

May 2, 2014

To the memory of A.A. Karatsuba, on the occasion of his 75th anniversary.

Abstract

If n is a positive integer, let $h(n)$ denote the maximal value of the product $q_1q_2\dots q_j$ for all families of primes $q_1 < q_2 < \dots < q_j$ such that $q_1 + q_2 + \dots + q_j \le n$. We shall give some properties of this function h and describe an algorithm able to compute $h(n)$ for any n up to 10^{35}.

1 Introduction

1.1 Function $h(n)$

If $n \ge 2$ is an integer, let us define $h(n)$ as the greatest product of a family of primes $q_1 < q_2 < \dots < q_j$ the sum of which does not exceed n.

Let ℓ be the additive function such that $\ell(p^\alpha) = p^\alpha$ for p prime and $\alpha \ge 1$. In other words, if the standard factorization of M into primes is $M = q_1^{\alpha_1} q_2^{\alpha_2} \dots q_j^{\alpha_j}$, we have $\ell(M) = q_1^{\alpha_1} + q_2^{\alpha_2} + \dots + q_j^{\alpha_j}$ and $\ell(1) = 0$. If μ denotes the Möbius function, $h(n)$ can also be defined as

$$h(n) = \max_{\substack{\ell(M) \le n \\ \mu(M) \ne 0}} M. \tag{1.1}$$

Note that

$$\ell(h(n)) \le n. \tag{1.2}$$

From the unicity of the factorization of $h(n)$ into primes, the maximum in (1.1) is attained in only one point. It is convenient to set

$$h(0) = h(1) = 1.$$

*Research partially supported by CNRS, Institut Camille Jordan, UMR 5208.

$(h(n))_{n\geq 1}$ is sequence A159685 of the OEIS (Online Encyclopedy of Integer Sequences). A table of the 50 first values of $h(n)$ is given at the end of the paper. A larger table may be found on the authors's web sites [2, 10].

In [9], Landau has introduced the function $g(n)$ as the maximal order of an element in the symmetric group $\mathfrak{S}_n$; he has shown that

$$g(n) = \max_{\ell(M)\leq n} M. \tag{1.3}$$

The introductions of [5] and [3] recall the main properties of Landau's function $g(n)$ which is mentionned as entry A002809 in [12]. From (1.1) and (1.3), it follows that

$$h(n) \leq g(n), \quad (n \geq 0). \tag{1.4}$$

In this article, we shall give some properties of $h(n)$ and describe an algorithm able to calculate $h(n)$ for any n up to 10^{35}.

1.2 Notation

- We denote by $\mathbb{N}$ the set of non-negative integers.
- The symbol p will always denote a prime number.
- For every arithmetic function $f\ :\mathbb{N}\to\mathbb{C}$, we define

$$\pi_f(x) = \sum_{p\leq x,\ p \text{ prime}} f(p) \tag{1.5}$$

- In particular, for $f(n) = 1$, we will note, as usual $\pi(x) = \pi_1(x)$ the number of primes up to x.
- For $f(n) = n$ we define

$$\pi_{id}(x) = \sum_{p\leq x,\ p \text{ prime}} p \tag{1.6}$$

- We denote by p_j the j-th prime and we set $\sigma_0 = 0$, $N_0 = 1$ and, for $j \geq 1$,

$$\sigma_j = \pi_{id}(p_j) = p_1 + p_2 + \ldots + p_j, \qquad N_j = p_1 p_2 \ldots p_j. \tag{1.7}$$

  In § 3, for all $j \geq 1$, we shall prove that $h(\sigma_j) = N_j$.

- If m is an integer, we denote by $m^\star$ the smallest prime p satisfying $p \geq m$ and, if $m \geq 2$, by ${}^\star m$ the largest prime p satisfying $p \leq m$.
- $P^+(m)$ (resp. $P^-(m)$) will denote the largest (resp. smallest) prime factor of $m \geq 2$. It is convenient to set $P^+(1) = -\infty$, $P^-(1) = +\infty$.

- $\omega(n)$ is the number of distinct prime factors of n and $\Omega(n)$ the number of prime factors of n, counted with multiplicity. $\mu(n)$ is Möbius's function.
- For $x > 1$, $\log_2(x) = \log\log x$.
- Li is the integral logarithm defined for $x > 1$ by

$$\mathrm{Li}(x) = \lim_{\varepsilon\to 0,\, \varepsilon>0} \int_0^{1-\varepsilon} + \int_{1+\varepsilon}^{x} \frac{\mathrm{t}}{\log t} = \gamma + \log_2 x + \sum_{n\geq 1} \frac{(\log x)^n}{nn!}$$

where γ is Euler's constant.

1.3 Functions $h_j(n)$

For $n \geq 0$, let $k = k(n)$ be the non-negative integer defined by

$$\sigma_k = \pi_{id}(p_k) \leq n < \pi_{id}(p_{k+1}) = \sigma_{k+1}. \tag{1.8}$$

It is the maximal number of prime factors of $h(n)$. For $0 \leq j \leq k = k(n)$, let us set

$$h_j(n) = \max_{\substack{\ell(M)\leq n \\ \mu(M)\neq 0,\, \omega(M)=j}} M \tag{1.9}$$

where $\omega(M)$ is the number of prime factors of M. For $n \geq 0$, we have

$$h_0(n) = 1 \tag{1.10}$$

while, for $n \geq 2$, we have

$$h_1(n) = {}^*n \geq 2. \tag{1.11}$$

Note that

$$\ell(h_j(n)) \leq n. \tag{1.12}$$

In § 6, we prove that, for all n's, the sequence $h_j(n)$ is increasing on j, so that

$$h(n) = h_k(n), \quad (n \geq 0). \tag{1.13}$$

Our proof is not that simple. A possible reason is that this increasingness relies on the properties of the whole set of primes $\mathcal{P}$. Let $\mathcal{P}'$ be a subset of $\mathcal{P}$ and $\mathcal{N}_{\mathcal{P}'}$ the set of integers whose prime factors belong to $\mathcal{P}'$. We may consider

$$h_j(n, \mathcal{P}') = \max_{\substack{M\in\mathcal{N}_{\mathcal{P}'},\, \ell(M)\leq n \\ \mu(M)\neq 0,\, \omega(M)=j}} M\ . \tag{1.14}$$

By choosing $\mathcal{P}' = \{2, 3, 11, 13, 17, 19, 23, \dots\} = \mathcal{P} \setminus \{5, 7\}$, we observe that

$$h_2(24, \mathcal{P}') = 11 \cdot 13 = 143 \quad > \quad h_3(24, \mathcal{P}') = 2 \cdot 3 \cdot 19 = 114.$$

In § 4, we give an upper bound for $h_j(n)$ which will be useful in § 6 where our proof of the increasingness of h_j is given. In (1.9), $h_j(n)$ can be considered as the solution of a problem of optimization with prime variables. The upper bound of $h_j(n)$ is obtained by relaxing some constraints so that certain variables are no longer primes, but only integers.

1.4 Elementary computation of $h(n)$ and $h_j(n)$

The naive algorithm described in [5] to compute $g(n)$ can be easily adapted to calculate $h(n)$ for $1 \le n \le N$. Note that, for the prime factors of $h(n)$, Corollary 3.1 below furnishes the upper bound

$$P^+(h(n)) \le p_{k(n)+1} + p_{k(n)+2}.$$

It also can be adapted to compute $h_j(n)$. For $r \ge j \ge 1$ and $n \ge \sigma_j$, let us define

$$h_j^{(r)}(n) = \max_{\substack{P^+(M) \le p_r,\, \ell(M) \le n \\ \mu(M) \ne 0,\, \omega(M) = j}} M \ .$$

We have the induction relation

$$h_j^{(r+1)}(n) = \max(h_j^{(r)}(n), p_{r+1} h_{j-1}^{(r)}(n - p_{r+1})).$$

Indeed, either p_{r+1} does not divide $h_j^{(r+1)}(n)$, and $h_j^{(r+1)}(n) = h_j^{(r)}(n)$ holds, or p_{r+1} divides $h_j^{(r+1)}(n)$, and $h_j^{(r+1)}(n) = p_{r+1} h_{j-1}^{(r)}(n - p_{r+1})$, which implies $n \ge p_{r+1} + \sigma_{j-1}$.

Moreover, if $p_r \ge n$, we have $h_j^{(r)}(n) = h_j(n)$, $h_r^{(r)}(n) = N_r$ and $h_1^{(r)}(n) = {}^*n$ for $n < p_r$ while, for $n \ge p_r$, $h_1^{(r)}(n) = p_r$ holds. So, we may write **algorithm 1**, which has been used to calculate the table in appendix. The merging and pruning method described in [5, §2.2] can be used to improve the running time.

In § 8, a more sophisticated algorithm to calculate $h(n)$ is given. It is based on a fast method to compute $\pi_{id}(x)$, which is explained in § 7.

2 Some lemmas

Lemma 2.1. *If $m \ge 2$ is an integer, let us denote by $m^\star$ (resp. ${}^\star m$) the smallest (resp. largest) prime p satisfying $p \ge m$ (resp. $p \le m$). Then*

$$m^\star \le \frac{11}{8} m \quad \text{and} \quad {}^*m \ge \frac{7}{10} m$$

hold.

Algorithm 1 Computation of $h_j(n)$ for $2 \le n \le nmax$ and $1 \le j \le k(n)$

```
Procedure ComputeHj(nmax)
r = 1; p = p_r; kmax = k(nmax); pmax = p_{kmax+1} + p_{kmax+2}
while p ≤ pmax do
    for n from σ_r to nmax do
        H[r, n] = N_r
    jmax = min(r − 1, kmax)
    for j from jmax by −1 to 2 do
        for n from nmax by −1 to p + σ_{j−1} do
            H[j, n] = max(H[j, n], p ∗ H[j − 1, n − p])
    for n from p to nmax do
        H[1, n] = p;
    r = r + 1; p = p_r
```

Proof. We use the result of [6]: for $x \ge 396738$, the interval $[x, x + \frac{x}{25\log^2 x}]$ contains a prime number. As 396833 is prime, we deduce that, for $p_i \ge 396833$,

$$\frac{p_{i+1}}{p_i} \le 1 + \frac{1}{25\log^2 p_i} \le 1 + \frac{1}{25\log^2 396833} < 1.00025 < \frac{11}{8} < \frac{10}{7}. \tag{2.1}$$

If m is prime, $m^\star = {}^\star m = m$ holds, while, if m is not prime, we define p_i by $p_i < m < p_{i+1}$; we have $m^\star = p_{i+1}$, ${}^\star m = p_i$,

$$\frac{m^\star}{m} \le \frac{p_{i+1}}{p_i + 1} < \frac{p_{i+1}}{p_i}, \qquad \frac{{}^\star m}{m} \ge \frac{p_i}{p_{i+1} - 1} > \frac{p_i}{p_{i+1}}$$

and, if $p_i \ge 396833$, the result follows from (2.1). Finally, it remains to check that

$$\frac{p_{i+1}}{p_i + 1} \le \frac{11}{8} \quad \text{and} \quad \frac{p_i}{p_{i+1} - 1} \ge \frac{7}{10}$$

hold for all p_i's satisfying $2 \le p_i < 396833$. □

Lemma 2.2. *Let $p < p'$ be two primes. There exists a third prime p'' satisfying*

$$p + p' \le p'' \le pp' - p + 1. \tag{2.2}$$

Proof. Let us show that $p'' = (p + p')^\star$ satisfies (2.2). By Lemma 2.1, it suffices to prove that $\frac{11}{8}(p + p') \le pp' - p + 1$, i.e:

$$pp'\left(8 - \frac{11}{p} - \frac{19}{p'} + \frac{8}{pp'}\right) \ge 0. \tag{2.3}$$

If $p \ge 3$ and $p' \ge 5$, we have $\frac{11}{p} + \frac{19}{p'} \le \frac{11}{3} + \frac{19}{5} < 8$ and (2.3) holds. Similarly, if $p = 2$ and $p' \ge 11$, the inequality $\frac{11}{p} + \frac{19}{p'} \le \frac{11}{2} + \frac{19}{11} < 8$ implies (2.3). In the three remaining cases, $p = 2$ and $p' \in \{3, 5, 7\}$, it is easy to check that $p'' = (p + p')^\star$ satisfies (2.2). □

Lemma 2.3. *Let p and p' be two prime numbers satisfying $3 \le p < p'$ and $pp' \ne 15$. There exists a prime p'' such that*

$$p + p' \le p'' \le \frac{5}{6}pp' - p. \tag{2.4}$$

Proof. The proof is similar to the one of the preceding lemma. From Lemma 2.1, to show that $p'' = (p+p')^\star$ satisfies (2.4), it suffices to show that $\frac{11}{8}(p+p') \le \frac{5}{6}pp' - p$, i.e. $33/p + 57/p' \le 20$, which evidently holds for $p \ge 3$ and $p' \ge 7$. □

Lemma 2.4. *For all $i \ge 2$, the following inequality*

$$p_i + p_{i-1} \le p_{2i-1} \tag{2.5}$$

holds. Moreover, let b be a positive integer; there exists a positive integer $i_0 = i_0(b)$ such that we have

$$p_i + p_{i-1} < p_{2i-b} \quad \textit{for } i \ge i_0(b). \tag{2.6}$$

The table below gives some values of $i_0(b)$

$b =$	1	2	3	4	5	6	7	8	9	10	12	13	18	30	3675
$i_0 =$	3	4	7	8	18	19	27	28	36	39	50	53	85	149	33127

Proof. We start from the two inequalities

$$p_i \le i(\log i + \log\log i - \alpha), \quad (\alpha = 0.9484,\ i \ge 39017), \tag{2.7}$$

$$p_i \ge i(\log i + \log\log i - 1), \quad (i \ge 2) \tag{2.8}$$

which can be found in [7]. From (2.7), it follows that

$$p_{i-1} + p_i \le (2i-1)(\log i + \log\log i - \alpha), \quad (i \ge 39018) \tag{2.9}$$

while, if $i \ge \max(2, b)$, which implies $2i - b \ge 2$ and $i \ge b$, (2.8) gives

$$p_{2i-b} \ge (2i-b)\left(\log i + \log 2 + \log\left(\frac{2i-b}{2i}\right) + \log\log i - 1\right). \tag{2.10}$$

By using the inequality $\log t \le t - 1$, we get

$$\log\left(\frac{2i-b}{2i}\right) = -\log\left(\frac{2i}{2i-b}\right) \ge -\left(\frac{2i}{2i-b} - 1\right) = -\frac{b}{2i-b}$$

and (2.10) yields

$$p_{2i-b} \ge (2i-b)(\log i + \log\log i + \log 2 - 1) - b. \tag{2.11}$$

Under the condition

$$i \geq \max(39018, b), \tag{2.12}$$

the substraction of (2.9) from (2.11) gives

$$\begin{aligned} p_{2i-b} - p_{i-1} - p_i &\geq (\log i + \log\log i + \log 2)(1-b) \\ &\quad + 2i(\log 2 - 1 + \alpha) - \log 2 - \alpha \\ &> (\log i + \log\log i + \log 2)\left[\frac{1.283\ i - 1.642}{\log i + \log\log i + \log 2} - (b-1)\right]. \end{aligned} \tag{2.13}$$

Now, the two functions $t \mapsto t/(\log t + \log\log t + \log 2)$ and $t \mapsto -1/(\log t + \log\log t + \log 2)$ are increasing for $t \geq e^2$; choosing $i_1 = 39018$ and

$$b = \left\lfloor 1 + \frac{1.283\ i_1 - 1.642}{\log i_1 + \log\log i_1 + \log 2} \right\rfloor = \lfloor 3675.52\ldots \rfloor = 3675 \tag{2.14}$$

shows that, for $i \geq i_1$, (2.12) is satisfied and that in (2.13), the bracket is positive. Therefore, (2.13) proves $p_i + p_{i-1} < p_{2i-3675}$ for $i \geq i_1 = 39018$.

To determine the entries of the table, for all i's up to 39018, we have calculated $b_i = 2i - 1 - \pi(p_i + p_{i-1})$ which is the smallest integer such that $p_{i-1} + p_i < p_{2i-b_i}$. Further, for each b in the table, we have determined $i_0(b)$ which is the smallest integer i_0 such that, for $i_0(b) \leq i \leq 39018$, $b_i \geq b$ holds.

As $i_0(1) = 3$, for all $i \geq 3$, $p_i + p_{i-1} < p_{2i-1}$ holds. So, (2.5) follows from $p_2 + p_1 = 3 + 2 = 5 = p_3$. □

Lemma 2.5. *Under Riemann hypothesis, for all $x \geq 41$ we have*

$$\left|\pi_{id}(x) - \mathrm{Li}(x^2)\right| \leq \frac{5}{24\pi} x^{3/2} \log x. \tag{2.15}$$

Proof. Let us define $r(x)$ by $\pi(x) = \mathrm{Li}(x) + r(x)$ and assume the Riemann hypothesis. Then cf. [11, (6.18)] :

$$|r(x)| = |\pi(x) - \mathrm{Li}(x)| \leq \frac{1}{8\pi}\sqrt{x}\log x \quad (\text{for } x \geq 2657). \tag{2.16}$$

Let us denote $x_0 = 2657$. Then, from (1.6), Stieltjes's integral gives :

$$\begin{aligned} \pi_{id}(x) &= \pi_{id}(x_0) + \int_{x_0^-}^{x} t\ \mathrm{d}[\pi(t)] \\ &= \pi_{id}(x_0) + \int_{x_0}^{x} t\,\mathrm{d}(\mathrm{Li}(t)) + \int_{x_0^-}^{x} t\,\mathrm{d}[r(t)] \\ &= \pi_{id}(x_0) + \mathrm{Li}(x^2) - \mathrm{Li}(x_0^2) + tr(t)|_{x_0}^{x} - \int_{x_0}^{x} r(t)\,\mathrm{d}t. \end{aligned}$$

With (2.16), it comes

$$\left|\pi_{id}(x) - \mathrm{Li}(x^2)\right| \leq \left|\pi_{id}(x_0) - \mathrm{Li}(x_0^2) - x_0 r(x_0)\right| + \frac{x^{3/2}\log x}{8\pi} + \int_{x_0}^{x} \frac{\sqrt{t}\log t}{8\pi}\,\mathrm{d}t$$

and, using $\int \sqrt{t}\log t = \frac{2}{3}t^{3/2}\left(\log t - \frac{2}{3}\right)$,

$$(2.17)\quad \left|\pi_{id}(x) - \mathrm{Li}(x^2)\right| \le \frac{5}{24\pi}x^{3/2}\log x - \frac{1}{18\pi}x^{3/2} + \left|\pi_{id}(x_0) - \mathrm{Li}(x_0^2) - x_0 r(x_0)\right| - \frac{1}{12\pi}x_0^{3/2}\log x_0 + \frac{1}{18\pi}x_0^{3/2}.$$

The computation of

$$\begin{aligned} r(x_0) &= \pi(x_0) - \mathrm{Li}(x_0) = 384 - 399.59681\ldots = -15.59681\ldots \\ \pi_{id}(x_0) - \mathrm{Li}(x_0^2) &= 464\,653 - 480610.2863\ldots = -15957.2863\ldots \end{aligned}$$

and (2.17) imply for $x \ge x_0$,

$$\left|\pi_{id}(x) - \mathrm{Li}(x^2)\right| \le \frac{5}{24\pi}x^{3/2}\log x - \frac{1}{18\pi}x^{3/2} - 740.023\ldots \le \frac{5}{24\pi}x^{3/2}\log x.$$

which proves (2.15) for $x \ge x_0 = 2657$. It remains to check (2.15) for $41 \le x \le 2657$; by setting

$$f_1(x) = \mathrm{Li}(x^2) - \frac{5}{24\pi}x^{3/2}\log x, \quad f_2(x) = \mathrm{Li}(x^2) + \frac{5}{24\pi}x^{3/2}\log x,$$

it is equivalent to check

$$(2.18)\qquad f_1(x) \le \pi_{id}(x) \le f_2(x)$$

for $41 \le x \le 2657$. One remarks that f_1 and f_2 are increasing for $x \ge 2$. Therefore, to prove (2.18), it suffices to check that for every prime p satisfying $41 \le p \le 2657$ we have $f_1(p') \le \pi_{id}(x) \le f_2(p)$ where p' is the prime following p. □

Note that, in the range [2..2657], $\pi_{id}(x) - \mathrm{Li}(x^2)$ has several changes of sign, the smallest one being for $x = 110.35\ldots$

Lemma 2.6. *Let z and u be two real numbers satisfying $z \ge 1$ and $\sqrt{z} \le u \le z$. Suppose that we have precomputed the tables* prime, piftab *and* pi. *The first two tables are indexed by the integers k, $0 \le k \le \pi(u)$, and the third one by the integers t, $0 \le t \le u$.*

- prime[k] *contains p_k ($p_0 = 1$).*
- piftab[k] *contains $\pi_f(p_k)$.*
- pi[t] *contains $\pi(t)$.*

Then the sum

$$\sum_{\sqrt{z}<q\le u,\ q\ prime} f(q)\pi_f\left(\frac{z}{q}\right) \tag{2.19}$$

may be computed in $O(\sqrt{z}/\log z)$ *time.*

Proof. For $q > \sqrt{z}$, z/q belongs to $[1, \sqrt{z})$. The number of primes in this interval is $O(\sqrt{z}/\log z)$, thus the number of values of $\pi_f(z/q)$ is $O(\sqrt{z}/\log z)$. We group the q's for which $\pi_f(z/q)$ takes the same value. Algorithm 2 carries out this computation.

Algorithm 2 : Computation of the sum (2.19) in $O\big(\sqrt{z}/\log z\big)$ time

```
S = 0;  imin = 1 + pi[⌊√z⌋]
while imin ≤ pi[u] do
    q = prime[imin]
    s = pi[z/q]
    imax = min(pi(z/prime[s]), pi[u])
    S = S + (piftab[imax] − piftab[imin − 1]) ∗ piftab[s]
    imin = imax + 1
return S
```

Let us give some words to convince of the correctness of algorithm 2 : let us note $s = \pi(z/q)$. Then p_s is the largest prime $\le z/q$. For q' prime, $q' \ge q$, we have $\pi_f(z/q') = \pi_f(p_s) = \text{piftab}[s]$ if and only if $z/q' \ge p_s$ i.e. $q' \le z/p_s$, in other terms, $\pi(q') \le \pi(z/p_s)$. Thus the largest prime q' in the range $[q..u]$ such that $\pi_f(z/q') = \pi_f(p_s)$ is p_i where $i = \min(\pi(z/p_s), \pi(u))$. □

3 First results

Proposition 3.1. *Let* j *be a positive integer and* σ_j *and* N_j *be defined by* (1.7)*. We have*

$$h(\sigma_j) = N_j.$$

Proof. It is easy to see that $h(\sigma_1) = h(2) = 2 = N_1$ and $h(\sigma_2) = h(5) = 6 = N_2$. Now, we may suppose that $j \ge 3$, i.e. $p_j \ge 5$ and we set $\rho = p_j/\log p_j$. The function $t \mapsto t/\log t$ is increasing for $t \ge e$ and, since $2/\log 2 < 5/\log 5$, we have, for $1 \le i < j$, $p_i/\log p_i < \rho$ and for $i > j$, $p_i/\log p_i > \rho$; in other words, $i - j$ and $p_i/\log p_i - \rho$ have the same sign.

Let M be a product of r distinct primes, $M = Q_1Q_2\ldots Q_r$, with $r \ge 0$. After a possible simplification by s primes ($0 \le s \le \min(j, r)$), we may write

$$\frac{M}{N_j} = \frac{p_{j_1}p_{j_2}\ldots p_{j_u}}{p_{k_1}p_{k_2}\ldots p_{k_v}}$$

with $u = r - s$, $v = j - s$ and

$$p_{k_1} < p_{k_2} < \ldots < p_{k_v} \le p_j < p_{j_1} < p_{j_2} < \ldots < p_{j_u}.$$

Let $f(M) = \ell(M) - \rho \log M$. From the definition of ℓ, the function f is additive and we have

$$f(M) - f(N_j) = \sum_{i=1}^{u} (p_{j_i} - \rho \log p_{j_i}) - \sum_{i=1}^{v} (p_{k_i} - \rho \log p_{k_i}) \ge 0 \tag{3.1}$$

since each term of the first sum is non-negative while, in the second sum, each term is non-positive.

From (1.1), since $\ell(N_j) = \sigma_j$, in order to prove that $h(\sigma_j) = N_j$, we must show that, for all squarefree number M satisfying $\ell(M) \le \sigma_j = \ell(N_j)$, we have $M \le N_j$. But, for such an M, (3.1) yields

$$f(M) = \ell(M) - \rho \log M \ge f(N_j) = \ell(N_j) - \rho \log N_j = \sigma_j - \rho \log N_j$$

whence

$$\frac{M}{N_j} \le \exp\left(\frac{\ell(M) - \sigma_j}{\rho}\right) \le 1,$$

which completes the proof of Proposition 3.1. $\square$

Proposition 3.2. *Let r and j be two positive integers and σ_j, N_j and h_j be defined by* (1.7) *and* (1.9)*. We have*

$$h_j(\sigma_{j+r} - \sigma_r) = N_{j+r}/N_r = p_{r+1} p_{r+2} \ldots p_{r+j}. \tag{3.2}$$

Moreover, if $n \ge \sigma_{j+r} - \sigma_r$ we have

$$\ell(h_j(n)) \ge \sigma_{j+r} - \sigma_r. \tag{3.3}$$

Proof. The proof is similar to the one of Proposition 3.1. Let us set

$$\rho = \frac{p_{j+r} - p_r}{\log(p_{j+r}/p_r)} \qquad \text{and} \qquad \rho' = \rho \log p_r - p_r. \tag{3.4}$$

Since, for $t \ne 1$, $(t-1)/t < \log t < t - 1$ holds, we have $p_{j+r} > \rho > p_r \ge 2$. For a squarefree number M, we consider the additive function

$$f(M) = \ell(M) - \rho \log M + \rho' \omega(M) = \sum_{p \mid M} f(P) = \sum_{p \mid M} (p - \rho \log p + \rho').$$

We will prove that f attains its minimum in $N = N_{j+r}/N_r$. From (3.4), it follows that $f(p_{j+r}) = f(p_r) = 0$ and the study of the function $t \mapsto t - \rho \log t + \rho'$ shows that

$$f(p) \begin{cases} > 0 & \text{for } p < p_r \text{ or } p > p_{j+r} \\ < 0 & \text{for } p_r < p < p_{j+r} \\ = 0 & \text{for } p = p_r \text{ or } p = p_{j+r}. \end{cases}$$

Therefore, we have

$$f(M) - f(N) = \sum_{\substack{p|M \\ p<p_r \text{ or } p>p_{j+r}}} f(p) - \sum_{\substack{p\nmid M \\ p_r<p<p_{j+r}}} f(p) \geq 0. \tag{3.5}$$

From (1.9), we have to show that, for any squarefree integer M satisfying $\ell(M) \leq \sigma_{j+r} - \sigma_r = \ell(N)$ and $\omega(M) = j = \omega(N)$, we have $M \leq N$. For such an M, (3.5) gives

$$\ell(M) - \rho \log M + \rho' \omega(M) \geq \ell(N) - \rho \log N + \rho' \omega(N)$$

yielding

$$\frac{M}{N} \leq \exp\left(\frac{\ell(M) - \ell(N)}{\rho}\right) \leq 1,$$

which, together with $\ell(N) = \sigma_{j+r} - \sigma_r$, proves (3.2).

To prove (3.3), first, from (1.9), we observe that $h_j(n) \geq N = N_{j+r}/N_r$. Setting $M = h_j(n)$ in (3.5) and noting that $\omega(M) = \omega(N) = j$, we see that

$$\ell(M) \geq \ell(N) + \rho \log \frac{M}{N} \geq \ell(N) = \sigma_{j+r} - \sigma_r$$

which proves (3.3). □

Proposition 3.3. *Let $n \geq 2$ be an integer and $p < p'$ two prime numbers which do not divide $h(n)$. Then the largest prime divisor $P^+(h(n))$ of $h(n)$ satisfies*

$$P^+(h(n)) < p + p'.$$

Proof. Let us assume that the set of prime factors of $h(n)$ not smaller than $p + p'$ is not empty and let $q \geq p + p'$ be its smallest element.

- If $q < pp'$, by setting $M = \frac{pp'}{q} h(n)$, we have by (1.2)

  $$\ell(M) = p + p' - q + \ell(h(n)) \leq \ell(h(n)) \leq n$$

  and thus, from (1.1),

  $$h(n) \geq M = \frac{pp'}{q} h(n), \tag{3.6}$$

  in contradiction with $q < pp'$.

- If $q > pp'$, i.e. $q \geq pp' + 1$, by Lemma 2.2, the interval $[p + p', q - p]$ contains a prime; thus the prime $p'' = {}^*(q - p)$ satisfies $p + p' \leq p'' \leq q - p < q$ and, from the definition of q, p'' does not divide $h(n)$. By Lemma 2.1, $p'' = {}^*(q - p) \geq \frac{7}{10}(q - p)$ holds, whence

  $$q \leq \frac{10}{7} p'' + p = \frac{pp''}{7}\left(\frac{10}{p} + \frac{7}{p''}\right).$$

We have $p \geq 2$, $p' \geq 3$ and $p'' \geq p+p' \geq 5$, so that $\frac{10}{p}+\frac{7}{p''} \leq \frac{10}{2}+\frac{7}{5} < 7$, yielding $q < pp''$. By considering $M = \frac{pp''}{q}h(n)$, as in (3.6), we get

$$h(n) \geq M = \frac{pp''}{q}h(n) > h(n),$$

a contradiction.

□

Corollary 3.1. *If $k = k(n)$ is defined by* (1.8)*, the largest prime factor of $h(n)$ satisfies*

$$P^+(h(n)) < p_{k+1} + p_{k+2}.$$

Proof. The number of prime factors of $h(n)$ does not exceed k, so that, among $p_1, p_2, \ldots, p_{k+2}$ there are certainly two prime numbers p and p' not dividing $h(n)$. By applying Proposition 3.3, we get $P^+(h(n)) < p + p' \leq p_{k+1} + p_{k+2}$. □

Proposition 3.4. *Let $n \geq 5$ be an integer, $k \geq 2$ be defined by* (1.8) *and j an integer satisfying $2 \leq j \leq k$. Let us supose that there exists two prime numbers, p, p' not dividing $h_{j-1}(n)$, and satisfying $3 \leq p < p'$ and $P^+(h_{j-1}(n)) \geq p + p'$ where $P^+(h_{j-1}(n))$ is the largest prime divisor of $h_{j-1}(n)$. Then the inequality*

$$h_j(n) > \frac{6}{5}h_{j-1}(n)$$

holds.

Proof. Let us consider two cases :

Case 1 : $pp' > 15$. Let us denote by $q \leq P^+(h_{j-1}(n))$ the smallest prime dividing $h_{j-1}(n)$ and satisfying $p + p' \leq q$.

- If $q < \dfrac{5}{6}pp'$, we set $M = \dfrac{pp'}{q}h_{j-1}(n)$; we have $\omega(M) = j$ and $\ell(M) = p + p' - q + \ell(h_{j-1}(n)) \leq \ell(h_{j-1}(n))$ so that, from (1.12), $\ell(M) \leq n$ holds and (1.9) yields

$$h_j(n) \geq M > \frac{6}{5}h_{j-1}(n) \tag{3.7}$$

  as required.

- If $q \geq \dfrac{5}{6}pp'$, we set $p'' = {}^*(q-p)$; from Lemma 2.3, $p+p' \leq p'' \leq q-p < q$ holds, and, from the definition of q, p'' does not divide $h_{j-1}(n)$.

By Lemma 2.1, we get $p'' = {}^*(q-p) \geq \frac{7}{10}(q-p)$, which implies

$$q \leq \frac{10}{7}p'' + p = \frac{pp''}{7}\left(\frac{10}{p} + \frac{7}{p''}\right).$$

But $p \geq 3$, $p' \geq 7$, $p'' \geq p + p' \geq 10$, thus $p'' \geq 11$, and $\frac{10}{p} + \frac{7}{p''} \leq \frac{10}{3} + \frac{7}{11} < \frac{35}{6}$, yielding $q < \frac{5}{6}pp''$. By setting $M = \frac{pp''}{q}h_{j-1}(n)$, as in (3.7), we get $h_j(n) \geq M > \frac{6}{5}h_{j-1}(n)$.

Case 2 : $p = 3$, $p' = 5$.

- If $P^+(h_{j-1}(n)) \leq 13$, which implies $n \leq \pi_{id}(13) = 41$, examining the table of Fig. 1 shows that, for $n \leq 41$, we have $h_j(n) \geq \frac{6}{5}h_{j-1}(n)$ with equality if and only if $h_{j-1}(n) = 5$, 35, 385 or 5005.
- If $P^+(h_{j-1}(n)) \geq 17$, and 11 does not divide $P^+(h_{j-1}(n))$, then we apply case 1 with $p = 3$, $p' = 11$, while, if 11 divides $P^+(h_{j-1}(n))$, $h_j(n) \geq \dfrac{3 \cdot 5}{11}h_{j-1}(n) > \dfrac{6}{5}h_{j-1}(n)$ holds.

□

4 Bounding $h_j(n)$

Proposition 4.1. *Let $j \geq 1$ and $n \geq \sigma_j$ (where σ_j has been introduced in (1.7)) be two integers; we define $r \geq 0$ by*

$$\sigma_{j+r} - \sigma_r \leq n < \sigma_{j+r+1} - \sigma_{r+1} \tag{4.1}$$

and n' by

$$0 \leq n' = n - (\sigma_{j+r} - \sigma_r) < p_{j+r+1} - p_{r+1}. \tag{4.2}$$

Then we have

$$h_j(n) \leq p_{r+1}p_{r+2}\dots p_{r+j}\frac{p_{j+r+1}}{p_{j+r+1} - n'} = \frac{N_{j+r+1}}{N_r(p_{j+r+1} - n')}. \tag{4.3}$$

Proof. From its definition (1.9), $h_j(n)$ is a product of j primes. Let us denote by $A_1, A_2, \dots, A_u$ (with $0 \leq u \leq j$) its prime factors exceeding p_{j+r+1} and by $B_1, B_2, \dots B_{r+1+u}$ the primes $\leq p_{j+r+1}$ and not dividing $h_j(n)$; we have

$$h_j(n) = \frac{N_{j+r+1}A_1A_2\dots A_u}{B_1B_2\dots B_{r+1+u}} \tag{4.4}$$

(where the product $A_1A_2\dots A_u$ should be replaced by 1 when $u = 0$) and

$$2 \leq B_1 < \dots < B_{r+1+u} \leq p_{j+r+1} < p_{j+r+2} \leq A_1 < \dots < A_u. \tag{4.5}$$

Further, let us introduce $\nu = \ell(h_j(n))$; by (1.12) and (4.1), we have

$$\nu \leq n < \sigma_{j+r+1} - \sigma_{r+1} \tag{4.6}$$

and it follows from Proposition 3.2, (3.3), that

$$\nu = \ell(h_j(n)) \geq \sigma_{j+r} - \sigma_r. \tag{4.7}$$

Moreover, (4.4) implies

$$\nu = \ell(h_j(n) = \sigma_{j+r+1} - \sigma_r + \sum_{i=1}^{u}(A_i - B_{r+1+i}) - \sum_{i=1}^{r}(B_i - p_i) - B_{r+1}. \tag{4.8}$$

Now, we consider the optimization problem (where $\nu, r, u, A_1, A_2, \ldots, A_u$ are fixed)

$$\mathcal{M} = \max_{\underline{Z}\in\mathcal{D}} \frac{A_1 A_2 \ldots A_u}{f(\underline{Z})} \tag{4.9}$$

where $\mathcal{D}$ is a subset of $\mathbb{N}^{r+1+u}$, $\underline{Z} = (Z_1, Z_2, \ldots, Z_{r+1+u})$,

$$f(\underline{Z}) = Z_1 Z_2 \ldots Z_{r+1+u}$$

and the set $\mathcal{D}$ is defined by

$$Z_i \geq p_i, \quad (1 \leq i \leq r+1), \tag{4.10}$$

$$Z_i < Z_{r+1}, \quad (1 \leq i \leq r), \tag{4.11}$$

$$Z_{r+1} < Z_{r+1+i} \leq A_i, \quad (1 \leq i \leq u) \tag{4.12}$$

and

$$U(\underline{Z}) - R(\underline{Z}) - Z_{r+1} + \sigma_{j+r+1} - \sigma_r = \nu \tag{4.13}$$

with

$$U(\underline{Z}) = \sum_{i=1}^{u}(A_i - Z_{r+1+i}), \qquad R(\underline{Z}) = \sum_{i=1}^{r}(Z_i - p_i). \tag{4.14}$$

Note that, from (4.5) and (4.8), $\underline{B} \in \mathcal{D}$ so that (4.4) implies

$$\frac{h_j(n)}{N_{j+r+1}} = \frac{A_1 A_2 \ldots A_u}{f(\underline{B})} \leq \mathcal{M} = \max_{\underline{Z}\in\mathcal{D}} \frac{A_1 A_2 \ldots A_u}{f(\underline{Z})}. \tag{4.15}$$

If $\underline{Z} \in \mathcal{D}$, from (4.10), (4.11) and (4.12), it follows that

$$2 = p_1 \leq Z_i \leq A_u, \qquad 1 \leq i \leq r+1+u$$

so that $f(\underline{Z})$ does not vanish on $\mathcal{D}$ and $\mathcal{D}$ is finite. Therefore, the maximum $\mathcal{M}$ defined by (4.9) is finite; let $\underline{C}$ be a point in $\mathcal{D}$ where the maximum $\mathcal{M}$ is attained. We shall prove that

$$U(\underline{C}) = R(\underline{C}) = 0. \tag{4.16}$$

For that, first we claim that one of the two numbers $U(\underline{C}), R(\underline{C})$ vanishes. Indeed, assume that $U(\underline{C}) \neq 0$ and $R(\underline{C}) \neq 0$. From (4.10), we have $R(\underline{C}) > 0$; thus there exists i_0, $1 \leq i_0 \leq r$, such that

$$C_{i_0} \geq p_{i_0} + 1 > p_{i_0}. \tag{4.17}$$

Similarly, from (4.12), we have $U(\underline{C}) > 0$, and there exists i_1, $1 \leq i_1 \leq u$ such that

$$A_{i_1} > A_{i_1} - 1 \geq C_{r+1+i_1}. \tag{4.18}$$

Let us define $\underline{C}' \in \mathbb{N}^{r+1+u}$ by

$$C'_{i_0} = C_{i_0} - 1,\ C'_{r+1+i_1} = C_{r+1+i_1} + 1,\ C'_i = C_i \quad \text{for } i \neq i_0, r+1+i_1.$$

To prove that $\underline{C}' \in \mathcal{D}$, we have to check that $C'_{i_0} \geq p_{i_0}$ (which follows from (4.17)), that $C'_{i_0} < C_{r+1}$ (which follows from $C'_{i_0} = C_{i_0} - 1 < C_{r+1} - 1$), that $C_{r+1} < C'_{r+1+i_1}$ (since $C_{r+1} < C_{r+1+i_1}$ and $C_{r+1+i_1} = C'_{r+1+i_1} - 1$), that $C'_{r+1+i_1} \leq A_{i_1}$ (which follows from (4.18)) and that $U(\underline{C}') - R(\underline{C}') = U(\underline{C}) - R(\underline{C})$ (which follows from $U(\underline{C}') = U(\underline{C}) - 1$ and $R(\underline{C}') = R(\underline{C}) - 1$). Further, we have

$$\begin{aligned} \frac{f(\underline{C}')}{f(\underline{C})} &= \frac{C'_{i_0} C'_{r+1+i_1}}{C_{i_0} C_{r+1+i_1}} = \frac{(C_{i_0} - 1)(C_{r+1+i_1} + 1)}{C_{i_0} C_{r+1+i_1}} \\ &= 1 - \frac{(C_{r+1+i_1} - C_{i_0} + 1)}{C_{i_0} C_{r+1+i_1}} < 1 \end{aligned} \tag{4.19}$$

since, from the definition of $\mathcal{D}$ (cf. (4.11) and (4.12)), $C_{i_0} < C_{r+1} < C_{r+1+i_1}$ holds. But (4.19) contradicts the fact that the maximum in (4.9) is attained in $\underline{C}$.

Let us show now that it is impossible to have simultaneously $U(\underline{C}) > 0$ and $R(\underline{C}) = 0$; indeed, let us assume that $U(\underline{C}) \geq 1$ and $R(\underline{C}) = 0$ (which implies $r = 0$ or $C_i = p_i$ for $1 \leq i \leq r$). We define i_1 as in (4.18). Since $\underline{C} \in \mathcal{D}$, we get from (4.13)

$$C_{r+1} = \sigma_{j+r+1} - \sigma_r - \nu + U(\underline{C}) = (\sigma_{j+r+1} - \sigma_{r+1} - \nu) + p_{r+1} + U(\underline{C})$$

which, by (4.6) and $U(\underline{C}) \geq 1$, yield

$$C_{r+1} > p_{r+1} + U(\underline{C}) \geq p_{r+1} + 1. \tag{4.20}$$

We define $\underline{C}' \in \mathbb{N}^{r+1+u}$ by

$$C'_{r+1} = C_{r+1} - 1,\ C'_{r+1+i_1} = C_{r+1+i_1} + 1,\ C'_i = C_i \ \text{ for } i \neq r+1, r+1+i_1.$$

To prove that $\underline{C}' \in \mathcal{D}$, we have to check that $C'_i \geq p_i$ for $1 \leq i \leq r+1$ (which follows from $C'_i = C_i = p_i$ if $i \leq r$ and from (4.20) if $i = r+1$), that $C'_i < C'_{r+1}$ for $1 \leq i \leq r$ (which follows from $C'_i = C_i = p_i \leq p_r$ and from $C'_{r+1} = C_{r+1} - 1 \geq p_{r+1}$, via(4.20)), that $C'_{r+1} < C'_{r+1+i}$ for $1 \leq i \leq u$ (which follows from $C'_{r+1} < C_{r+1}$ and $C'_{r+1+i} \geq C_{r+1+i}$), that $C'_{r+1+i_1} \leq A_{i_1}$ (which follows from (4.18)) and that $U(\underline{C}') - C'_{r+1} = U(\underline{C}) - C_{r+1}$) (which is easy). As in (4.19), we have $f(\underline{C}') < f(\underline{C})$, contradicting the fact that the maximum in (4.9) is attained in $\underline{C}$.

To prove (4.16), it remains to show that we cannot have $R(\underline{C}) > 0$ and $U(\underline{C}) = 0$. Let us suppose that $R(\underline{C}) \geq 1$ and $U(\underline{C}) = 0$, which implies $u = 0$ or, for $1 \leq i \leq u$,

$$C_{r+1+i} = A_i \geq p_{j+r+2} \geq p_{j+r+1} + 2, \tag{4.21}$$

with the help of (4.5). From (4.13) and (4.7), this time we get

$$\begin{aligned} C_{r+1} = \sigma_{j+r+1} - \sigma_r - \nu - R(\underline{C}) &\leq \sigma_{j+r+1} - \sigma_r - (\sigma_{j+r} - \sigma_r) - R(\underline{C}) \\ &= p_{j+r+1} - R(\underline{C}) \leq p_{j+r+1} - 1. \end{aligned} \tag{4.22}$$

Here we choose i_0 as in (4.17) and set

$$C'_{i_0} = C_{i_0} - 1,\ C'_{r+1} = C_{r+1} + 1,\ C'_i = C_i \ \text{ for } i \neq i_0, r+1.$$

To prove that $\underline{C}' \in \mathcal{D}$, we have to check that $C'_{i_0} \geq p_{i_0}$ (which follows from (4.17)), that $C'_{r+1} \geq p_{r+1}$ (which follows from $C'_{r+1} = C_{r+1} + 1$ and $C_{r+1} \geq p_{r+1}$), that, for $1 \leq i \leq r$, $C'_i < C'_{r+1}$ (which follows from $C'_i \leq C_i$ and $C'_{r+1} > C_{r+1}$), that, for $1 \leq i \leq u$, $C'_{r+1} < C_{r+1+i_1} = A_{i_1}$ (which follows from (4.21) and (4.22)) and that $R(\underline{C}')+C'_{r+1} = R(\underline{C})+C_{r+1}$ (which is easy). As precedingly in (4.19), we observe that $f(\underline{C}') < f(\underline{C})$, contradicting the fact that the minimum is attained in $\underline{C}$.

In conclusion, we have proved (4.16) so that $C_i = p_i$ for $1 \leq i \leq r$ and $C_{r+1+i} = A_i$ for $1 \leq i \leq u$. Moreover, (4.6) yields $\nu \leq n$, and, from (4.13) and (4.2), we get

$$C_{r+1} = \sigma_{j+r+1} - \sigma_r - \nu \geq \sigma_{j+r+1} - \sigma_r - n = p_{j+r+1} - n'.$$

Therefore, the maximum $\mathcal{M}$ in (4.9) satisfies

$$\mathcal{M} = \frac{A_1 A_2 \ldots A_u}{p_1 p_2 \ldots p_r C_{r+1} A_1 A_2 \ldots A_u} \leq \frac{1}{p_1 p_2 \ldots p_r (p_{j+r+1} - n')}$$

which, via (4.15), proves (4.3). □

Proposition 4.2. *With the notation of Proposition 4.1 , we have*

$$h_j(n) \geq \frac{N_{j+r+1}}{N_r(p_{j+r+1}-n')^\star} = \frac{N_{j+r+1}}{qN_r} \tag{4.23}$$

where $q = (p_{j+r+1}-n')^\star$ is the smallest prime satisfying $q \geq p_{j+r+1}-n'$.

Proof. From (4.2), we have $p_{r+1} < p_{j+r+1}-n' \leq p_{j+r+1}$ which implies $p_{r+1} \leq q \leq p_{j+r+1}$ so that $M = N_{j+r+1}/(qN_r)$ is an integer with exactly j prime factors. Further, by (4.2), we have

$$\ell(M) = \sigma_{j+r+1} - \sigma_r - q \leq \sigma_{j+r+1} - \sigma_r - (p_{j+r+1}-n') = n$$

and, by (1.9), $h_j(n) \geq M$ holds. □

Corollary 4.1. *We keep the notation of Proposition 4.1; if $q = p_{j+r+1}-n'$ is prime then*

$$h_j(n) = h_j(\sigma_{j+r+1}-\sigma_r-q) = \frac{N_{j+r+1}}{qN_r}\cdot \tag{4.24}$$

Proof. Corollary 4.1 follows from Propositions 4.1 and 4.2. □

5 A parity phenomenon

Proposition 5.1. *Let $k \geq 2$ be an integer and a be an even number satisfying $4 \leq a < p_{k+1}$ and h_k defined by (1.9). We have*

$$h_k(\sigma_{k+1}-a) = h_k(\sigma_{k+1}-a-1). \tag{5.1}$$

Proof. Since $n \mapsto h_k(n)$ in non-decreasing, we have

$$h_k(\sigma_{k+1}-a) \geq h_k(\sigma_{k+1}-a-1). \tag{5.2}$$

Let us set $n = \sigma_{k+1}-a$ and note that n satisfies $\sigma_k < n < \sigma_{k+1}$ so that, from (1.8), $k = k(n) = k(n-1)$. Let M be a positive squarefree integer such that $\ell(M) \leq n$ and $\omega(M) = k$. Such a number M is even; if not, we would have $\ell(M) \geq 3+5+\ldots+p_{k+1} = \sigma_{k+1}-2$ in contradiction with $\ell(M) \leq n = \sigma_{k+1}-a \leq \sigma_{k+1}-4$. Therefore, $\ell(M)$ is the sum of 2 and $k-1$ odd numbers, so that $\ell(M) \equiv \sigma_k \equiv \sigma_{k+1}+1 \equiv \sigma_{k+1}-a-1 \pmod 2$. So, $\ell(M)$ cannot be equal to $\sigma_{k+1}-a$ and $\ell(M) \leq \sigma_{k+1}-a-1$ holds. Thus, from (1.9), we get $h_k(\sigma_{k+1}-a) \leq h_k(\sigma_{k+1}-a-1)$, which, with (5.2), proves (5.1). □

Proposition 5.2. *Let k be an integer, $k \geq 2$, and q a prime number satisfying $3 \leq q \leq p_k$. By setting $m = \sigma_{k+1}-q-1$, we have*

$$h_{k-1}(m) = h_{k-1}(m-1) = h_{k-1}(\sigma_{k+1}-q-2) = \frac{N_{k+1}}{2q}\cdot \tag{5.3}$$

Proof. From the table of Figure 1, we have $h_1(6) = h_1(5) = 5$, $h_2(11) = h_2(10) = 21$ and $h_2(13) = h_2(12) = 35$ so that the proposition is true for $k = 2$, $q = 3$ and for $k = 3$ and $q = 3$ or 5. So, from now on, we assume $k \geq 4$. Corollary 4.1 with $j = k-1$, $r = 1$ implies $h_{k-1}(m-1) = N_{k+1}/(2q)$ and, since $n \mapsto h_{k-1}(n)$ is non-decreasing, it follows that

$$h_{k-1}(m) \geq h_{k-1}(m-1) = \frac{N_{k+1}}{2q}. \tag{5.4}$$

Let M be a positive squarefree integer satisfying $\ell(M) \leq m$ and $\omega(M) = k-1$. In view of (5.4) and (1.9), to prove that $h_{k-1}(m) = N_{k+1}/(2q)$, it suffices to show that

$$M \leq \frac{N_{k+1}}{2q}. \tag{5.5}$$

If M is odd, $\ell(M)$ is the sum of $k-1$ odd numbers, which implies

$$\ell(M) \equiv \sigma_k \equiv \sigma_{k+1} - q = m+1 \pmod 2.$$

So, $\ell(M)$ cannot be equal to m; since, by (1.9), $\ell(M) \leq m$ holds, we should have $\ell(M) \leq m-1$; therefore, from (1.9), we get

$$M \leq h_{k-1}(m-1) = \frac{N_{k+1}}{2q}. \tag{5.6}$$

If M is even, we have $\omega(M/2) = k-2$ and $\ell(M/2) \leq m-2$, so that

$$M \leq 2h_{k-2}(m-2). \tag{5.7}$$

- If $q \geq 11$, since we have assumed $k \geq 4$, i.e. $p_{k+1} \geq 11$, we have

  $$\sigma_k - 5 \leq m-2 = \sigma_{k+1} - q - 3 \leq \sigma_{k+1} - 14 < \sigma_{k+1} - 10.$$

  By Proposition 4.1 with $j = k-2$, $r = 2$, $n = m-2$, $n' = n-(\sigma_k - \sigma_2) = p_{k+1} - q + 2$, we get

  $$h_{k-2}(m-2) \leq \frac{N_{k+1}}{6(q-2)}$$

  which, by (5.7), gives

  $$M \leq \frac{N_{k+1}}{3(q-2)} = \frac{N_{k+1}}{3q}\frac{q}{q-2} \leq \frac{N_{k+1}}{3q}\frac{11}{9} < \frac{N_{k+1}}{2q}. \tag{5.8}$$

- If $q \in \{3, 5, 7\}$, since $k \geq 4$ and $p_{k+2} \geq p_6 = 13$, we have

  $$\sigma_{k+1} - 10 \leq m-2 = \sigma_{k+1} - q - 3 \leq \sigma_{k+1} - 6 < \sigma_{k+2} - 17.$$

and Proposition 4.1 with $j = k-2$, $r = 3$, $n = m-2$ and $n' = n - (\sigma_{k+1} - 10) = 7 - q$ yields

$$(5.9) \qquad h_{k-2}(m-2) \leq \frac{N_{k+2}}{30(p_{k+2}+q-7)}\cdot$$

Using $q \leq 7$, $k \geq 4$ and $p_{k+2} \geq p_6 = 13$ gives

$$\begin{aligned}\frac{N_{k+2}}{30(p_{k+2}+q-7)} &= \frac{N_{k+1}}{30}\frac{p_{k+2}}{p_{k+2}+q-7} \leq \frac{N_{k+1}}{30}\frac{13}{q+6}\\ &= \frac{13N_{k+1}}{30q}\frac{q}{q+6} \leq \frac{13N_{k+1}}{30q}\frac{7}{13} < \frac{N_{k+1}}{4q}\end{aligned}$$

which, together with (5.7) and (5.9), proves

$$(5.10) \qquad M < \frac{N_{k+1}}{2q}\cdot$$

Inequalities (5.6), (5.8) and (5.10) prove that (5.5) holds, which, with (5.4), completes the proof of (5.3).

□

Proposition 5.3. *Let k be a positive integer and $m = \sigma_{k+1} - 1$; we have*

$$(5.11) \qquad h_k(m) = h_k(m-1) = h_k(\sigma_{k+1}-2) = \frac{N_{k+1}}{2}\cdot$$

Proof. It is the same proof than for Proposition 5.2. By Proposition 3.2 with $j = k$ and $r = 1$, we have

$$h_k(m) \geq h_k(m-1) = \frac{N_{k+1}}{2}\cdot$$

Further, let M be a positive integer satisfying $\ell(M) \leq m$ and $\omega(M) = k$. If M is odd, by the parity phenomenon, we have $\ell(M) \equiv m-1 \pmod 2$ so that $\ell(M) \leq m-1$ and $M \leq h_k(m-1) = \frac{N_{k+1}}{2}\cdot$. If M is even, we have $M \leq 2h_{k-1}(\sigma_{k+1}-3)$ and, if $k \geq 2$, i.e. $p_{k+2} \geq 7$, Proposition 4.1 with $j = k-1$, $r = 2$, $n = \sigma_{k+1} - 3$, $n' = 2$, yields

$$M \leq 2\frac{N_{k+2}}{6(p_{k+2}-2)} = \frac{N_{k+1}}{3}\frac{p_{k+2}}{p_{k+2}-2} \leq \frac{7N_{k+1}}{15} < \frac{N_{k+1}}{2}\cdot$$

If $k = 1$, it is easy to check that (5.11) still holds. □

6 The increasingness of $h_j(n)$ on j

Theorem 6.1. *Let $n \geq 2$ be an integer and $k = k(n)$ be defined by (1.8); for j satisfying $1 \leq j \leq k$, we have*

$$h_{j-1}(n) \leq \frac{5}{6} h_j(n) \tag{6.1}$$

and (6.1) is an equality if and only if $j = k(n) \geq 2$ and $n = \sigma_{j+1} - 4$ or $n = \sigma_{j+1} - 5$.

Proof. If $j = 1$, it follows from (1.10) and (1.11) that $h_0(n) = 1$, $h_1(n) = {}^*n \geq 2$ and $h_0(n)/h_1(n) \leq 1/2 < 5/6$, which proves (6.1). So, from now on, we assume $j \geq 2$.

The sequence $(\sigma_{j+r} - \sigma_r)_{r \geq 0}$ is increasing and goes to infinity. So, we may define $r_j \geq 0$ and n'_j by

$$\sigma_{j+r_j} - \sigma_{r_j} \leq n < \sigma_{j+r_j+1} - \sigma_{r_j+1} \tag{6.2}$$

and

$$n'_j = n - (\sigma_{j+r_j} - \sigma_{r_j}). \tag{6.3}$$

We shall consider four cases : $r_j \leq j-4$, $r_j \geq j+3$, $j-3 \leq r_j \leq j+2$ and $j \geq 25$, $j-3 \leq r_j \leq j+2$ and $j \leq 24$.

First case : $r_j \leq j - 4$

From (2.5) and our hypothesis $j \geq r_j + 4$, we deduce

$$p_{r_j+1} + p_{r_j+2} \leq p_{2r_j+3} < p_{j+r_j} < p_{j+r_j+1} \tag{6.4}$$

and

$$p_{r_j+2} + p_{r_j+3} \leq p_{2r_j+5} < p_{j+r_j+2}. \tag{6.5}$$

Let us set From (6.2), we get

$$0 \leq n'_j = n - (\sigma_{j+r_j} - \sigma_{r_j}) < p_{j+r_j+1} - p_{r_j+1} \tag{6.6}$$

and applying Proposition 4.2 yield

$$h_j(n) \geq \frac{N_{j+r_j+1}}{qN_{r_j}} = \frac{N_{j+r_j+1}}{(p_{j+r_j+1} - n'_j)^* \; N_{r_j}}. \tag{6.7}$$

In view of bounding $h_{j-1}(n)$, we have to determine r_{j-1} such that

$$\sigma_{j-1+r_{j-1}} - \sigma_{r_{j-1}} \leq n < \sigma_{j+r_{j-1}} - \sigma_{r_{j-1}+1}. \tag{6.8}$$

We shall distinguish two sub cases.

Sub case one, $r_{j-1} = r_j + 1$

Let us asume that

$$\sigma_{j+r_j} - \sigma_{r_j} \le n < \sigma_{j+r_j+1} - \sigma_{r_j+2}. \tag{6.9}$$

i.e. from (6.3),

$$0 \le n'_j = n - (\sigma_{j+r_j} - \sigma_{r_j}) < p_{j+r_j+1} - p_{r_j+1} - p_{r_j+2}. \tag{6.10}$$

Note that, from (6.4), the right hand side of (6.10) is positive. Then, we have $r_{j-1} = r_j + 1$ since, from (6.9),

$$\begin{aligned} \sigma_{(j-1)+(r_j+1)} - \sigma_{r_j+1} &= \sigma_{j+r_j} - \sigma_{r_j+1} < \sigma_{j+r_j} - \sigma_{r_j} \le n \\ &< \sigma_{j+r_j+1} - \sigma_{r_j+2} = \sigma_{(j-1)+(r_j+1)+1} - \sigma_{(r_j+1)+1} \end{aligned}$$

holds. Via (6.3), this implies that

$$n'_{j-1} \overset{def}{=\!=} n - (\sigma_{j-1+r_{j-1}} - \sigma_{r_{j-1}}) = n - \sigma_{j+r_j} + \sigma_{r_j+1} = n'_j + p_{r_j+1}.$$

Applying Proposition 4.1 and noting that $j - 1 + r_{j-1} = j + r_j$ yield

$$h_{j-1}(n) \le \frac{N_{j-1+r_{j-1}+1}}{N_{r_{j-1}}(p_{j-1+r_{j-1}+1} - n'_{j-1})} = \frac{N_{j+r_j+1}}{N_{r_j+1}(p_{j+r_j+1} - n'_j - p_{r_j+1})}.$$

By using (6.7), we get

$$\frac{h_{j-1}(n)}{h_j(n)} \le \frac{(p_{j+r_j+1} - n'_j)^\star}{p_{r_j+1}(p_{j+r_j+1} - n'_j - p_{r_j+1})}. \tag{6.11}$$

From (6.6), we have $p_{j+r_j+1} - n'_j > p_{r_j+1} \ge p_1 = 2$, so that we may apply Lemma 2.1 which, with the help of (6.11) and (6.10), yields

$$\frac{h_{j-1}(n)}{h_j(n)} \le \frac{11}{8p_{r_j+1}} \left(1 + \frac{p_{r_j+1}}{p_{j+r_j+1} - n'_j - p_{r_j+1}}\right) < \frac{11}{8}\left(\frac{1}{p_{r_j+1}} + \frac{1}{p_{r_j+2}}\right).$$

If $r_j \ge 1$, $\frac{11}{8}\left(\frac{1}{p_{r_j+1}} + \frac{1}{p_{r_j+2}}\right) \le \frac{11}{8}\left(\frac{1}{3} + \frac{1}{5}\right) < \frac{5}{6}$, which proves (6.1).

It remains to consider the case $r_j = 0$, which, from (6.9) and (1.8), implies $\sigma_j \le n < \sigma_{j+1}$ and $k(n) = j$.

(6.10) becomes $0 \le n'_j = n - \sigma_j < p_{j+1} - 5$ and, by setting $a = p_{j+1} - n'_j$, we get

$$5 < a = p_{j+1} - n'_j = p_{j+1} + \sigma_j - n = \sigma_{j+1} - n < p_{j+1} \tag{6.12}$$

while (6.11) yields

$$\frac{h_{j-1}(n)}{h_j(n)} \le \frac{a^\star}{2(a-2)}. \tag{6.13}$$

By Lemma 2.1, $a^\star \le \frac{11}{8}a$ holds, and, for $a \ge 12$, $\frac{11}{16}\frac{a}{a-2} \le \frac{11}{16}\frac{12}{10} < \frac{5}{6}$, which, via (6.13), proves (6.1).

Since, from (6.12), $a > 5$, it remains to study the cases $6 \le a \le 11$. If $a = 7, 9, 10, 11$, it is easy to check that $\frac{a^\star}{2(a-2)} < \frac{5}{6}$.

If $a = 6$ or $a = 8$, by Proposition 5.1, (6.12) and (4.23) we have

$$h_j(n) = h_j(\sigma_{j+1} - a) = h_j(\sigma_{j+1} - a - 1) = h_j(n-1) \ge \frac{N_{j+1}}{(a+1)^*}$$

while, by Proposition 5.2, since $a-1$ is prime, we get

$$h_{j-1}(n) = h_{j-1}(n-1) = \frac{N_{j+1}}{2(a-1)}$$

yielding

$$\frac{h_{j-1}(n)}{h_j(n)} = \frac{h_{j-1}(n-1)}{h_j(n-1)} \le \frac{(a+1)^\star}{2(a-1)} = \begin{cases} 7/10 & \text{if } a = 6 \\ 11/14 & \text{if } a = 8 \end{cases}$$

and, in both cases, $\frac{h_{j-1}(n)}{h_j(n)} < \frac{5}{6}$ holds, which proves (6.1).

Sub case two, $r_{j-1} = r_j + 2$

Now, we asume that (6.2) holds but not (6.9); thus we have

$$\sigma_{j+r_j+1} - \sigma_{r_j+2} \le n < \sigma_{j+r_j+1} - \sigma_{r_j+1} \tag{6.14}$$

and, from (6.5),

$$p_{j+r_j+1} - p_{r_j+1} - p_{r_j+2} \le n'_j < p_{j+r_j+1} - p_{r_j+1}. \tag{6.15}$$

Here, we get $r_{j-1} = r_j + 2$, since we have

$$\sigma_{j-1+r_j+2} - \sigma_{r_j+2} \le n < \sigma_{j+r_j+1} - \sigma_{r_j+1} < \sigma_{j+r_j+2} - \sigma_{r_j+3}$$

by observing that, from (6.5),

$$\sigma_{j+r_j+2} - \sigma_{r_j+3} - (\sigma_{j+r_j+1} - \sigma_{r_j+1}) = p_{j+r_j+2} - p_{r_j+2} - p_{r_j+3} > 0$$

holds. Now, we have

$$\begin{aligned} 0 \le \quad n'_{j-1} &= n - (\sigma_{j-1+r_{j-1}} - \sigma_{r_{j-1}}) = n - \sigma_{j+r_j+1} + \sigma_{r_j+2} \\ &= n'_j + \sigma_{j+r_j} - \sigma_{r_j} - \sigma_{j+r_j+1} + \sigma_{r_j+2} \qquad \text{from (6.6)} \\ &= n'_j - p_{j+r_j+1} + p_{r_j+1} + p_{r_j+2} < p_{r_j+2} \qquad \text{from (6.15)} \end{aligned} \tag{6.16}$$

and applying Proposition 4.1 gives

$$h_{j-1}(n) \leq \frac{N_{j-1+r_{j-1}+1}}{N_{r_{j-1}}(p_{j-1+r_{j-1}+1} - n'_{j-1})} = \frac{N_{j+r_j+2}}{N_{r_j+2}(p_{j+r_j+2} - n'_{j-1})}$$

while Proposition 4.2 yields

$$h_j(n) \geq \frac{N_{j+r_j+1}}{N_{r_j}(p_{j+r_j+1} - n'_j)^\star}.$$

We set $a = p_{j+r_j+1} - n'_j$ and $\Delta = p_{j+r_j+2} - p_{r_j+1} - p_{r_j+2}$ so that (6.16) allows to write $p_{j+r_j+2} - n'_{j-1} = \Delta + a$, and we have

$$\frac{h_{j-1}(n)}{h_j(n)} \leq \frac{p_{j+r_j+2}}{p_{r_j+1}p_{r_j+2}} \frac{a^\star}{\Delta + a}. \tag{6.17}$$

(6.16) can be rewritten as

$$2 = p_1 \leq p_{r_j+1} < a = p_{j+r_j+1} - n'_j \leq p_{r_j+1} + p_{r_j+2}. \tag{6.18}$$

Lemma 2.1 implies $a^\star \leq 11a/8$ and, by (6.4), $\Delta > 0$ holds, so that the homographic function $t \mapsto t/(\Delta + t)$ is increasing. From (6.18), we thus have

$$\frac{a}{\Delta + a} \leq \frac{p_{r_j+1} + p_{r_j+2}}{\Delta + p_{r_j+1} + p_{r_j+2}} = \frac{p_{r_j+1} + p_{r_j+2}}{p_{j+r_j+2}}.$$

Therefore, we get

$$\frac{h_{j-1}(n)}{h_j(n)} \leq \frac{11}{8} \frac{p_{r_j+1} + p_{r_j+2}}{p_{r_j+1}p_{r_j+2}} = \frac{11}{8}\left(\frac{1}{p_{r_j+1}} + \frac{1}{p_{r_j+2}}\right)$$

which is smaller than $\dfrac{5}{6}$ if $r_j \geq 1$.

It remains to consider the case $r_j = 0, r_{j-1} = 2$. Formula (6.17) becomes

$$\frac{h_{j-1}(n)}{h_j(n)} \leq \frac{p_{j+2}}{6} \frac{a^\star}{p_{j+2} - 5 + a} \tag{6.19}$$

while (6.18) via (6.3) becomes

$$2 < a = p_{j+1} - n'_j = \sigma_{j+1} - n \leq 5. \tag{6.20}$$

Since $j \geq 2$ holds, note that (6.20) implies $\sigma_j \leq \sigma_{j+1} - 5 \leq n < \sigma_{j+1} - 2$, which shows from (1.8) that $k(n) = j$.

- If $a = 5$, since $n = \sigma_{j+1} - a = \sigma_{j+1} - 5$, by Corollary 4.1 with $r = r_j = 0$ and $q = 5$, we get $h_j(n) = N_{j+1}/5$, while Proposition 3.2 gives

  $$h_{j-1}(\sigma_{j+1} - 5) = h_{j-1}(\sigma_{j+1} - \sigma_2) = \frac{N_{j+1}}{N_2} = \frac{N_{j+1}}{6}.$$

  Therefore, $h_{j-1}(\sigma_{j+1} - 5)/h_j(\sigma_{j+1} - 5) = 5/6$.

- If $a = 4$, by Proposition 5.1, we get $h_j(\sigma_{j+1} - 4) = h_j(\sigma_{j+1} - 5)$ and, by Proposition 5.2, $h_{j-1}(\sigma_{j+1} - 4) = h_{j-1}(\sigma_{j+1} - 5)$ so that

$$\frac{h_{j-1}(\sigma_{j+1} - 4)}{h_j(\sigma_{j+1} - 4)} = \frac{h_{j-1}(\sigma_{j+1} - 5)}{h_j(\sigma_{j+1} - 5)} = \frac{5}{6}.$$

- If $a = 3$, Formula (6.19) becomes

$$\frac{h_{j-1}(n)}{h_j(n)} \leq \frac{p_{j+2}}{2(p_{j+2} - 2)} \leq \frac{7}{10} < \frac{5}{6}$$

  since $p_{j+2} \geq p_4 = 7$.

Second case : $r_j \geq j + 3$

From (6.2), we deduce $n \geq \sigma_{j+r_j} - \sigma_{r_j+1} = \sigma_{(j-1)+(r_j+1)} - \sigma_{(r_j+1)}$ and Proposition 3.2, (3.3), implies

$$\ell(h_{j-1}(n)) \geq \sigma_{j+r_j} - \sigma_{r_j+1}. \tag{6.21}$$

Let us now show that

$$q \stackrel{def}{=\!=} P^+(h_{j-1}(n)) \geq p_{j+r_j}. \tag{6.22}$$

Indeed, if $q \leq p_{j+r_j-1}$ holds, since $h_{j-1}(n)$ has $j-1$ prime factors, we should have

$$\begin{aligned} \ell(h_{j-1}(n)) &\leq p_{j+r_j-1} + p_{j+r_j-2} + \ldots + p_{r_j+1} \\ &< p_{j+r_j} + p_{j+r_j-1} + \ldots + p_{r_j+2} = \sigma_{j+r_j} - \sigma_{r_j+1} \end{aligned}$$

which would contradict (6.21).

Further, among the $j+1$ primes $p_2 = 3, p_3, \ldots, p_{j+2}$, there are certainly two primes p and p' not dividing $h_{j-1}(n)$ and satisfying $3 \leq p < p' \leq p_{j+2}$. By Lemma 2.4, (2.5), and (6.22), we get

$$p + p' \leq p_{j+1} + p_{j+2} \leq p_{2j+3} \leq p_{j+r_j} \leq q = P^+(h_{j-1}(n)) \tag{6.23}$$

and, applying Proposition 3.4 proves $\dfrac{h_{j-1}(n)}{h_j(n)} < \dfrac{5}{6}$.

Third case : $j - 3 \leq r_j \leq j + 2$ **and** $j \geq 25$

The proof is the same than for the second case; only, in (6.23), instead of (2.5), we use (2.6) with $b = 7$, $i = j + 2 \geq 27$:

$$p + p' \leq p_{j+1} + p_{j+2} < p_{2j-3} \leq p_{j+r_j} \leq q = P^+(h_{j-1}(n)).$$

Fourth case : $j-3 \leq r_j \leq j+2$ **and** $j \leq 24$

Here, we have $r_j \leq j+2 \leq 26$ and, from (6.2), we get

$$n < \sigma_{j+r_j+1} \leq \sigma_{51} = 5350.$$

So, for $k \leq 50$, $\sigma_k \leq n < \sigma_{k+1}$ and $1 \leq j \leq k$, we have computed $h_j(n)$ with the algorithm described in Section 1.4 and we have checked that, for $j \geq 2$,

$$\frac{h_{j-1}(n)}{h_j(n)} \leq \frac{5}{6}$$

always holds, with equality if and only if $j = k(n)$ and $n = \sigma_{j+1} - 4$ or $n = \sigma_{j+1} - 5$. $\square$

Corollary 6.1. *For all non-negative integer $n \geq 2$, we have*

$$h(n) = h_k(n) \tag{6.24}$$

where $k = k(n)$ is defined by (1.8)*.*

Proof. From (1.1) and (1.9) we have

$$h(n) = \max_{0 \leq j \leq k(n)} h_j(n)$$

and Theorem 6.1 yields $h_0(n) < h_1(n) < \ldots < h_{k(n)}(n)$. $\square$

7 Computation of $\pi_f(x)$

Let f be an arithmetic function, i.e a function defined on positive integers. The simplest way to compute $\pi_f(x)$ defined in (1.5) is to generate the primes up to x by Eratosthenes's sieve, which is too expansive for large values of x.

Definition 7.1. *An arithmetic function f is said to be completely multiplicative if $f(ab) = f(a)f(b)$ for all a and b. If $f \neq 0$, this implies $f(1) = 1$.*

Following ideas of the german astronomer Meissel, Lagarias, Miller and Odlyzko gave in [8] an algorithm that computes $\pi(x)$ with a cost $O\left(\frac{x^{2/3}}{\log x}\right)$. In this work they also remark that their algorithm allows to compute $\pi_f(x)$ for every completely multiplicative arithmetic function f.

This method has been improved in [4] to compute $\pi(x)$ with a cost $O\left(\frac{x^{2/3}}{\log^2 x}\right)$, provided that all the arithmetic operations on integers are of constant cost $O(1)$, not depending on the size of the operands. We show here that this improved algorithm may be used to compute $\pi_f(x)$ whith a cost which is still $O\left(\frac{x^{2/3}}{\log^2 x}\right)$, for a large subset of the set of completely multiplicative arithmetic functions. More precisely we have the proposition :

Proposition 7.1. *Let f be a completely multiplicative arithmetic function with integer values. Let F be the summatory function of f,*

$$F(x) = \sum_{n \le x} f(n). \tag{7.1}$$

We suppose that all the ordinary arithmetic operations about integers are of constant cost $O(1)$, and that

1. *Each value $f(n)$ may be computed in time $O(1)$, not depending of the size of n.*

2. *There is an algorithm computing*

$$S_0(y, x) = \sum_{1 \le n \le y} \mu(n) f(n) F\left(\frac{x}{n}\right) \tag{7.2}$$

   in time $O\Big(x^{2/3}/\log^2 x\Big)$.

Then, there is an algorithm computing $\pi_f(x) = \sum_{\substack{p \le x \\ p \text{ prime}}} f(p)$ in time $O\left(\dfrac{x^{2/3}}{\log^2 x}\right)$.

When $F(u)$ can be computed in $O(1)$ time, the second hypothesis is satisfied.

Remarks :

1. The second hypothesis may seem strange. Let us give a few words of explanation.

   - Our computation of $\pi_f(x)$ begins by choosing $y = O\big(x^{1/3+\varepsilon}\big)$. Then we compute $S_0 = S_0(y, x)$ (this is the *contribution of ordinary leaves* defined in lemma 5.2, equation (9) in [4] and in lemma 7.2, equation (7.14) in this article). Function F does not appear elsewhere in the algorithm. S_0 being computed, the total cost of the other computations is $O\Big(x^{2/3}/\log^2 x\Big)$. Condition (2) ensures that our algorithm computes $\pi_f(x)$ in time $O\Big(x^{2/3}/\log^2 x\Big)$.
   - In many cases, $F(u)$ can be computed in time $O(1)$, then the sum (7.2) can be computed in time $O(y)$, by precomputing the Möbius function, so that the second hypothesis is satisfied.

2. In Proposition 7.1 we restrict ourseves to the case of integer valued functions. The case of real valued functions is more delicate because of truncation errors. In [1], Bach and al. have elaborated an algorithm to compute $\pi_f(x)$ where $f(n) = 1/n$, and
   $x = 1\,801\,241\,484\,456\,448\,000 = 1.8\ldots \times 10^{18}$.

Algorithm for $\pi_f(x)$

We will describe very briefly our algorithm to compute $\pi_f(x)$, using notations and formulas which, when replacing f by 1, reduce to the correponding ones contained in [4].

For $b \in \mathbb{N}$, let us define $\Phi(x, b)$ as the sum of the $f(n)$, for the $n's \in [1,\ x]$ that subsist after sieving this interval by all primes $p_1, p_2, \ldots, p_b$,

(7.3)
$$\Phi(x,0) = \sum_{1 \le n \le x} f(n) = F(x) \quad \text{and, for } b \ge 1, \quad \Phi(x,b) = \sum_{\substack{1 \le n \le x \\ P^-(n) > p_b}} f(n)$$

For $k \geq 1$ and $b \geq 1$, let us set

$$P_k(x,b) = \sum_{\substack{1 \le n \le x \\ \Omega(n)=k,\ P^-(n) > p_b}} f(n). \tag{7.4}$$

so that, from (7.3) and (7.4) we get, for $x \geq 1$,

$$\Phi(x,b) = 1 + P_1(x,b) + P_2(x,b) + \cdots \tag{7.5}$$

From now on, we choose $y \in \mathbb{R}$

$$x^{1/3} \le y \le \sqrt{x} \quad \text{and set} \quad a = \pi(y). \tag{7.6}$$

We will precise later the best choice for y, which is closed to $x^{1/3}$.

Since $y \geq x^{1/3}$ equation (7.4) yields $P_k(x, a) = 0$ for $k \geq 3$ and (7.5) becomes

$$\Phi(x,a) = 1 + P_1(x,a) + P_2(x,a).$$

Since $P_1(x,a) = \sum_{p_a < p \le x} f(p) = \sum_{y < p \le x} f(p) = \pi_f(x) - \pi_f(y)$,

$$\pi_f(x) = \Phi(x,a) + \pi_f(y) - 1 - P_2(x,a). \tag{7.7}$$

Replacing f by 1 (and π_f by π), formula (7.7) is formula (4) in [4].

7.1 Initialization of the computation: the 2 basis tables

After fixing y, by using Eratosthenes's sieve, we precompute the table of primes up to y, and the table of the values $\pi_f(u)$ for $1 \leq u \leq y$. The cost of these initializations is $O(y \log_2 y)$.

7.2 Computation of $P_2(x, a)$

Definition (7.4) and the complete multiplicativity of f give

$$P_2(x,a) = \sum_{\substack{y<p\le q\le x\\ pq\le x}} f(pq) = \sum_{\substack{y<p\le q\le x\\ pq\le x}} f(p)f(q)$$

where p and q are primes. The $p's$ figuring in this sum satisfy $p \le \dfrac{x}{q} \le \dfrac{x}{y}$ and we get

$$P_2(x,a) = \sum_{y<p\le x/y} f(p) \sum_{p\le q\le x/p} f(q).$$

We remark that, for $p > \sqrt{x}$, the sum on q vanishes. Since, by (7.6), $\sqrt{x} \le \dfrac{x}{y}$, we have

$$P_2(x,a) = \sum_{y<p\le\sqrt{x}} f(p) \sum_{p\le q\le x/p} f(q) = \sum_{y<p\le\sqrt{x}} f(p)\left(\pi_f\left(\frac{x}{p}\right) - \pi_f(p-1)\right)$$

or

$$P_2(x,a) = \sum_{y<p\le\sqrt{x}} f(p)\pi_f\left(\frac{x}{p}\right) - \sum_{y<p\le\sqrt{x}} f(p)\pi_f(p-1). \tag{7.8}$$

In the above formula, the values of p are bounded above by $\sqrt{x}$ which is larger than y. Thus we cannot find these primes p, nor the values $\pi_f(p-1)$ in the precomputed tables (cf. §7.1), and we generate them using a sieve of $[1, \sqrt{x}]$, which we call the *auxilliary sieve*. The values of x/p lie in the interval $[1, x/y]$. So we will get the values $\pi_f(x/p)$ by an other sieve, the *main sieve*. Let us note that the respective sizes of the sieve intervals, $\sqrt{x}$ and x/y, are too large to allow a sieve in one pass. Thus the two sieves will be done by blocks of size y that must be synchronized.

- **Initialization: Computation of ϖ, the largest prime $\le \sqrt{x}$ and of $\pi_f(\varpi)$.** By Eratosthenes'sieve we compute the largest prime ϖ not exceeding $\sqrt{x}$ and calculate $\pi_f(\varpi)$. The auxilliary sieve is then initialised by putting in the sieve-table the primes p of the block $[\sqrt{x} - y + 1, \sqrt{x}]$. The main sieve is initialized by sieving the first block $[A, B] = \left[\sqrt{x},\ \sqrt{x} + y - 1\right]$. The cost of this phase is $O\Big(x^{1/2}\log_2 x\Big)$.

- **Computing $P_2(x, a)$.** We use formula (7.8), getting in decreasing order the primes $p \in]y, \sqrt{x}]$ and the $f(p)\pi_f(p-1)$ from the auxillary sieve, and getting the values $\pi_f(x/p)$ from the main sieve whose successive blocks will cover in ascending order the interval $\left[\sqrt{x},\ x/y\right)$.

  We initialize a variable p with the value ϖ, a variable T with the value $\pi_f(\varpi)$ and a variable P_2 with the value 0. Then, while $p > y$, we repeat :

– substract $f(p)$ from T. Thus the new value of T is $\pi_f(p-1)$.
– If $x/p > B$, while $x/p > B$ we replace the block $[A, B]$ by the next block $[A+y, B+y]$ and we sieve it. When $x/p \in [A, B]$ we get $\pi_f(x/p)$ in the main sieve table and we add $f(p)\pi_f(x/p) - f(p)T$ to P_2.
– Using the auxilliary sieve, replace p by its predecessor.

The final value of the variable P_2 is $P_2(x, a)$. The first step is negligible in cost, compared to the second. Thus the computation of $P_2(x, a)$ is of total cost $O\left(\frac{x}{y}\log_2 y\right)$.

7.3 Computation of $\Phi(x, a)$

The following lemma is proved as lemma 5.1 in [4].

Lemma 7.1. *For every* $u \geq 0$, *and for* $b \geq 1$,

$$\Phi(u, 0) = F(u) \tag{7.9}$$

$$\Phi(u, b) = \Phi(u, b-1) - f(p_b)\Phi\left(\frac{u}{p_b}, b-1\right) \tag{7.10}$$

This relation gives an obvious method for computing $\Phi(x, a)$. Starting from the tree with the only node $\Phi(x, a)$, and applying repeatedly (7.10) we get a tree whose all nodes, except the root node, are labelled by a formula of the form

$$\mu(n)f(n)\Phi\left(\frac{x}{n}, b\right) \tag{7.11}$$

where $b \leq a-1$ and $n = 1$ or n is a squarefree integer with prime factors $q \in \{p_{b+1}, \ldots, p_a\}$.

If we repeat this expansions until all the leaves of the resulting tree are labelled by formulas $\mu(n)f(n)\Phi\left(\frac{x}{n}, 0\right)$, using (7.9) we get the formula :

$$\Phi(x, a) = \sum_{\substack{1 \leq n \leq x \\ P^+(n) \leq y}} \mu(n)f(n)\Phi\left(\frac{x}{n}, 0\right) = \sum_{\substack{1 \leq n \leq x \\ P^+(n) \leq y}} \mu(n)f(n)F\left(\frac{x}{n}\right) \tag{7.12}$$

which, when $f = 1$ is formula $\Phi(x, a) = \sum_{\substack{1 \leq n \leq x \\ P^+(n) \leq y}} \mu(n)\left\lfloor\frac{x}{n}\right\rfloor$ (cf. [4, p. 237]).

The number of terms in (7.12) is much too large. In order to get a sum with fewer terms we replace the trivial rule

Rule 1 : *Expand* (7.11) *using* (7.10) *if* $b > 0$, which leads to (7.12) by the new rule

$$\textbf{Rule 2 :}\quad \textit{Expand node}(7.11)\ \textit{only if } b>0 \textit{ and } n\le y.$$

Expanding the computation tree whith rule 2 instead of rule 1 we get

Lemma 7.2. *We have*

$$\Phi(x,b) = S_0 + S, \tag{7.13}$$

where S_0 is the contribution of ordinary leaves

$$S_0 = S_0(y,x) = \sum_{1\le n\le y} \mu(n) f(n) \Phi\left(\frac{x}{n},0\right) = \sum_{1\le n\le y} \mu(n) f(n) F\left(\frac{x}{n}\right) \tag{7.14}$$

and S, the contribution of special leaves*, is*

$$S = \sum_{\frac{n}{P^-(n)}\le y<n} \mu(n) f(n) \Phi\left(\frac{x}{n}, \pi(P^-(n))-1\right)\cdot \tag{7.15}$$

This lemma corresponds to lemma (5.2) in [4].

7.3.1 Computation of S_0

In the general case, the computation of S_0 is done with a cost $O\big(x^{2/3}/\log^2 x\big)$ thanks to the condition 2 in proposition 7.1.

In the case we will consider later in this work, the computation of $\pi_{id}(x)$, $f(n)=n$, thus $F(u) = \dfrac{[u][u+1]}{2}$ is computed in $O(1)$ time and the computation of $S_0(x,y)$ is of cost $O(y) = o\Big(x^{2/3}/\log^2 x\Big)$.

7.3.2 Computation of S

In the sum (7.15), let us set $n=mp$ with $p=P^-(n)$. Grouping together all the n's according to the value of p, we get

$$S = -\sum_{p\le y} f(p) \sum_{\substack{P^-(m)>p\\ m\le y<mp}} \mu(m) f(m) \Phi\left(\frac{x}{mp}, \pi(p)-1\right)\cdot \tag{7.16}$$

The computation of S from (7.16) is the complicate part of the algorithm. In the following paragraph we show that it is relatively simple to get a cost $O\big(x^{2/3+\varepsilon}\big)$.

7.3.3 How to compute S in $O\left(x^{2/3+\varepsilon}\right)$

In this section, we explain a first method to get $\pi_f(x)$, rather simple to implement, and whose running time is $O(x^{2/3+\varepsilon})$. We take $y = x^{1/3}$. Since $mp > y$ all the values $u = x/mp$ appearing in (7.16) are less than $x^{2/3}$. We sieve the interval $\left[1, \dfrac{x}{y}\right)$ successively by all primes $p \le y$. After the sieve by p, from the definition (7.3) of Φ, for all the m's such that $m \le y < mp$, we get in the sieve table the value $\Phi\left(\dfrac{x}{mp}, \pi(p) - 1\right)$, and we add to S the value $f(p)\mu(m)f(m)\Phi\left(\dfrac{x}{mp}, \pi(p) - 1\right)\cdot$

But, if we proceed in the naive way, after sieving by each p, we will update the sieve table, putting in the case of index u the sum of $f(n)$ for the n's, $n \le u$ that are still in the table. This is excluded because, for each p this would cost $O(x/y)$ operations, and the total cost of these updatings would be $\gg \pi(y)(x/y) = x/\log x$. As explained in [8] (the 7 last-lines p. 545 and the first half of p. 546) we use an auxiliary data structure such that, for a price of $O(\log x)$ time in place of $O(1)$ for each access, we don't need to update the sieve table after each sieve. To be a little more precise let us say that this structure is a labelled binary tree. There is a leave for each index i of the table sieve, this leave is labelled by the value $f(i)$, and each interior node is labelled by the sum of labels of its two sons. Proceeding in this way the cost of the sieve is $O\left(\dfrac{x}{y}\log x \log_2 x\right)$, while the cost of retrieving the values $f(p)\mu(m)f(m)\Phi\left(\dfrac{x}{mp}, \pi(p) - 1\right)$ is $O(\pi(y) y \log x)$. Both costs are $O\left(x^{2/3+\varepsilon}\right)$ with our choice $y = x^{1/3}$.

7.3.4 Faster computation of S

In this section, we explain how to carry out the computation of $\pi_f(x)$ in $O\left(\dfrac{x^{2/3}}{\log^2 x}\right)$. We take $y = x^{1/3}(\log x)^3 \log_2 x$. To speed up the computation of S we partition (7.16) in 3 subsums $S = S_1 + S_2 + S_3$,

$$
\begin{aligned}
S_1 &= -\sum_{x^{\frac{1}{3}} < p \le y} f(p) \sum_{\substack{P^-(m) > p \\ m \le y < mp}} \mu(m) f(m) \Phi\left(\frac{x}{mp}, \pi(p) - 1\right) \\
S_2 &= -\sum_{x^{\frac{1}{4}} < p \le x^{\frac{1}{3}}} f(p) \sum_{\substack{P^-(m) > p \\ m \le y < mp}} \mu(m) f(m) \Phi\left(\frac{x}{mp}, \pi(p) - 1\right)
\end{aligned}
$$

$$S_3 = -\sum_{p \le x^{\frac{1}{4}}} f(p) \sum_{\substack{P^-(m)>p \\ m \le y < mp}} \mu(m) f(m) \Phi\left(\frac{x}{mp}, \pi(p) - 1\right)$$

We will show that S_1 is quickly computed in $O(y)$ time. S_3 will be computed by sieve, as explained in §7.3.3, but faster because the number of values for p is reduced from $\pi(y)$ to $\pi(x^{1/4})$. The main part of the computation will be the computation of S_2.

As in [4], we first observe that the m's involved in S_1 and S_2 are all prime and therefore :

$$S_1 = \sum_{x^{\frac{1}{3}} < p \le y} f(p) \sum_{p < q \le y} f(q) \Phi\left(\frac{x}{pq}, \pi(p) - 1\right) \tag{7.17}$$

$$S_2 = \sum_{x^{\frac{1}{4}} < p \le x^{\frac{1}{3}}} f(p) \sum_{p < q \le y} f(q) \Phi\left(\frac{x}{pq}, \pi(p) - 1\right) \tag{7.18}$$

Computing S_1 As in [4] we remark that, in (7.17), we have $\frac{x}{pq} < x^{1/3} < p$. Thus, all the values $\Phi\left(\frac{x}{pq}, \pi(p) - 1\right)$ are equal to 1. Therefore

$$S_1 = \sum_{x^{\frac{1}{3}} < p \le y} f(p) \sum_{p < q \le y} f(q) = \sum_{x^{\frac{1}{3}} < p \le y} f(p) \left(\pi_f(y) - \pi_f(p)\right).$$

This value is computed in $O(y)$ additions, using the precomputed table of the values $\pi_f(u)$ for $1 \le u \le y$.

Computing S_3 For each $p \le x^{1/4}$ we precompute the list of all the squarefree $m \le y$ whose least factor is p.

We sieve the interval $\left[1, \frac{x}{y}\right]$ successively by all the primes up to $x^{1/4}$. As soon as we have sieved by p, using the precomputded lists of squarefree whose least prime factor is a prime $q > p$ we sum the

$$f(p) \sum_{\substack{P^-(m)>p \\ m \le y < mp}} \mu(m) f(m) \Phi\left(\frac{x}{mp}, \pi(p) - 1\right)$$

for all squarefree $m \in [y/p, y)$ such that $P^-(m) > p$. This computation is done by blocks, using the auxiliary structure, as explained at the end of § 7.3.3.

Thus the cost of sieving is $O\left(\frac{x}{y}\log x \log_2 x\right)$. The number of values of p is $\pi\left(x^{1/4}\right)$ and the number of values of m is less than y, thus the cost of retrieving the values $f(p)\mu(m)f(m)\Phi\left(\frac{x}{mp}, \pi(p)-1\right)$ is $O\left(\pi\left(x^{1/4}\right)\times \log x \times y\right)$ Thus computing S_3 is of cost $O\left(\frac{x}{y}\log x \log_2 x + yx^{1/4}\right)$

Computing S_2 We split the sum (7.18) in two parts depending on $q > x/p^2$ or $q \le x/p^2$. It gives

$$S_2 = U + V$$

with

$$U = \sum_{x^{\frac{1}{4}} < p \le x^{\frac{1}{3}}} f(p) \sum_{\substack{p<q\le y\\ q>x/p^2}} f(q)\Phi\left(\frac{x}{pq}, \pi(p-1\right)$$

and

$$V = \sum_{x^{\frac{1}{4}} < p \le x^{\frac{1}{3}}} f(p) \sum_{\substack{p<q\le y\\ q\le x/p^2}} f(q)\Phi\left(\frac{x}{pq}, \pi(p-1\right)$$

Computing U With $y < \sqrt{x}$ (cf. (7.6)), the condition $q > x/p^2$ implies $p^2 > x/q \ge x/y \ge x^{1/2}$. Thus,

$$U = \sum_{\sqrt{x/y} < p \le x^{1/3}} f(p) \sum_{\substack{p<q\le y\\ q>x/p^2}} f(q)\Phi\left(\frac{x}{pq}, \pi(p-1\right)$$

From $x/p^2 < q$ we deduce $x/pq < p$ and $\Phi(x/pq, \pi(p)-1) = 1$, and we have $x/p^2 \ge p$ so that

$$U = \sum_{\sqrt{x/y} < p \le x^{1/3}} f(p) \sum_{\substack{p<q\le y\\ q>x/p^2}} f(q) = \sum_{\sqrt{x/y} < p \le x^{1/3}} f(p) \left(\pi_f(y) - \pi_f\left(\frac{x}{p^2}\right)\right).$$

Since $x/p^2 < q \le y$ the sum U is calculated in $O(y)$ operations with the table of values of $\pi_f(u)$.

Computing V For each term involved in V we have $p \le \frac{x}{pq} < x^{1/2} < p^2$. Hence, by (7.3), $\Phi(x/(pq), \pi(p)-1)$ is the sum of $f(n)$ for n satisfying $n \le x/pq$ and $P^-(n) \ge p$. These n's are $n = 1$ and all the primes n satisfying $p - 1 < n \le x/(pq)$. Thus

$$\Phi\left(\frac{x}{pq}, \pi(p)-1\right) = 1 + \pi_f\left(\frac{x}{pq}\right) - \pi_f(p-1)$$

And we write

$$V = V_1 + V_2$$

with

$$\begin{aligned} V_1 &= \sum_{x^{1/4} \leq p < x^{1/3}} f(p) \sum_{p<q\leq \min(\frac{x}{p^2}, y)} f(q)(1 - \pi_f(p-1)) \\ V_2 &= \sum_{x^{1/4} \leq p < x^{1/3}} f(p) \sum_{p<q\leq \min(\frac{x}{p^2}, y)} f(q)\pi_f\left(\frac{x}{pq}\right) \end{aligned}$$

Computing V_1 can be achieved in $O(y)$ time once we have tabulated $\pi_f(u)$ for $u \leq y$.

Computing V_2. We first split V_2 in two parts in order to simplify the condition $q \leq \min(x/p^2, y)$:

$$V_2 = \sum_{x^{1/4}<p\leq\sqrt{\frac{x}{y}}} f(p) \sum_{p<q\leq y} f(q)\pi_f\left(\frac{x}{pq}\right) + \sum_{\sqrt{\frac{x}{y}}<p<x^{1/3}} f(p) \sum_{p<q\leq\frac{x}{p^2}} f(q)\pi_f\left(\frac{x}{pq}\right)$$

In the purpose to speed up the computation of the above two sums we now write, $V_2 = W_1 + W_2 + W_3 + W_4 + W_5$ with

$$\begin{aligned} W_1 &= \sum_{x^{1/4}<p\leq\frac{x}{y^2}} f(p) \sum_{p<q\leq y} f(q)\,\pi_f\left(\frac{x}{pq}\right) \\ W_2 &= \sum_{\frac{x}{y^2}<p\leq\sqrt{\frac{x}{y}}} f(p) \sum_{p<q\leq\sqrt{\frac{x}{p}}} f(q)\,\pi_f\left(\frac{x}{pq}\right) \\ W_3 &= \sum_{\frac{x}{y^2}<p\leq\sqrt{\frac{x}{y}}} f(p) \sum_{\sqrt{\frac{x}{p}}<q\leq y} f(q)\,\pi_f\left(\frac{x}{pq}\right) \\ W_4 &= \sum_{\sqrt{\frac{x}{y}}<p\leq x^{1/3}} f(p) \sum_{p<q\leq\sqrt{\frac{x}{p}}} f(q)\,\pi_f\left(\frac{x}{pq}\right) \\ W_5 &= \sum_{\sqrt{\frac{x}{y}}<p\leq x^{1/3}} f(p) \sum_{\sqrt{\frac{x}{p}}<q\leq\frac{x}{p^2}} f(q)\,\pi_f\left(\frac{x}{pq}\right) \end{aligned}$$

Computing W_1 and W_2 These two quantities need values of $\pi_f(x/pq)$ with $x^{1/3} < x/pq < x^{1/2}$. These are computed with a sieve of the interval $[1, \sqrt{x}]$. The sieving is done by blocks, and, for each block, we sum $f(p)f(q)\pi_f(x/pq)$ for the pairs (p, q) satisfying the conditions of the sum W_1 or W_2 and such that x/pq lies in the block.

The cost of this computation is the sum of three terms :

- The cost of the above sieve on $\left[1,\,\sqrt{x}\right]$ is $O\big(\sqrt{x}\log_2 x\big)$.
- The cost of adding the terms of the sum W_1, $O\left(\dfrac{x}{y\log^2 x}\right)$.
- The cost of adding the terms of the sum W_2, $O\left(\dfrac{x^{3/4}}{y^{1/4}\log^2 x}\right)$.

Computing W_3 and W_5 For W_3, for each p we apply lemma 2.6 with $z = x/p$ and $u = y$. Thus, for each value of p, the sum on q costs $O\Big(\pi(\sqrt{x/p})\Big)$, and the total cost of the computation of W_3 is

$$O\left(\sum_{\frac{x}{y^2}<p\le\frac{x}{y}} \pi\left(\sqrt{\frac{x}{p}}\right)\right).$$

For W_5, for each p we apply lemma 2.6 with $z = x/p$ and $u = x/p^2$. Thus, for each value of p, the sum on q costs $O\Big(\pi(\sqrt{x/p})\Big)$, and the total cost of the computation of W_5 is

$$O\left(\sum_{\sqrt{\frac{x}{y}}<p\le x^{1/3}} \pi\left(\sqrt{\frac{x}{p}}\right)\right).$$

Thus the costs of computing W_3 and W_5 add to

$$O\left(\sum_{\frac{x}{y^2}<p\le x^{1/3}} \pi\left(\sqrt{\frac{x}{p}}\right)\right). \quad = \quad O\left(\sum_{\frac{x}{y^2}<p\le x^{1/3}} \left(\frac{\sqrt{\frac{x}{p}}}{\log\frac{x}{p}}\right)\right) = O\left(\frac{x^{2/3}}{\log^2 x}\right).$$

Computing W_4 We simply sum over (p, q). There would be no advantage to proceed as for W_3 since most of the values $\pi_f(x/pq)$ are distinct. The cost is

$$O\left(\sum_{\sqrt{\frac{x}{y}}<p\le x^{1/3}} \pi\left(\sqrt{\frac{x}{p}}\right)\right) = O\left(\frac{x^{2/3}}{\log^2 x}\right).$$

As in [4], section 8, we then see that, since $y = x^{1/3}\log^3 x\log_2 x$, the total cost of the computation of $\pi_f(x)$ is $O\big(x^{2/3}/\log^2 x\big)$.

8 The algorithm to calculate $h(n)$

8.1 The function $G(p_k, m)$

The function $G(p_k,m)$ has been introduced and studied in [5].

Definition 8.1. *Let p_k be the k-th prime, for some $k \geq 3$ and m an integer satisfying $0 \leq m \leq p_{k+1} - 3$. We define*

$$G(p_k, m) = \max \frac{Q_1 Q_2 \dots Q_s}{q_1 q_2 \dots q_s} \tag{8.1}$$

where the maximum is taken over the primes $Q_1, Q_2, \dots, Q_s, q_1, q_2, \dots, q_s$ ($s \geq 0$) satisfying

$$3 \leq q_s < q_{s-1} < \dots < q_1 \leq p_k < p_{k+1} \leq Q_1 < Q_2 < \dots < Q_s \tag{8.2}$$

and

$$\sum_{i=1}^{s} (Q_i - q_i) \leq m. \tag{8.3}$$

The additive function ℓ (cf. §1.1) can easily be extended to fractions by setting

$$\ell(M/N) = \ell(M) - \ell(N)$$

when M and N are coprime or are both squarefree. Therefore, the inequality (8.3) implies

$$\ell(G(p_k, m)) \leq m. \tag{8.4}$$

8.1.1 Properties of $G(p_k, m)$

Obviously, $G(p_k, m)$ is non-decreasing on m and $G(p_k, 2m+1) = G(p_k, 2m)$. The maximum in (8.1) is unique (from the unicity of the standard factorization into primes). For small m's, we have

$$0 \leq m < p_{k+1} - p_k \implies G(p_k, m) = 1. \tag{8.5}$$

From Proposition 8 of [5], we have

$$\frac{p_{k+1}}{(p_{k+1} - m)^\star} \leq G(p_k, m) \leq \frac{p_{k+1}}{p_{k+1} - m}. \tag{8.6}$$

Note that if $p_{k+1} - m$ is prime, then (8.6) yields the exact value of $G(p_k, m)$.

8.1.2 Computation of $G(p_k, m)$

In [5, §8], two algorithms are given to calculate $G(p_k, m)$.

The first one is a combinatorial algorithm. In its first step, the primes allowed to divide the denominator of $G(p_k, m)$ are determined. From (8.2) and (8.3), they are all the primes in the range $[(p_{k+1} - m), p_k]$, say $P_1 < P_2 < \ldots < P_K$. Similarly, the primes authorized to divide the numerator are all the primes $P_{K+1} < P_{K+2} < \ldots < P_R$ in $[p_{k+1}, p_k + m]$. By setting $\mathcal{P}' = \{P_1, P_2, \ldots, P_R\}$, from the definition (1.14), we get

$$G(p_k, m) = \frac{1}{P_1 P_2 \ldots P_K} h_K(P_1 + P_2 + \ldots + P_K + m, \mathcal{P}')$$

and $h_j(n, \mathcal{P}')$ can be computed by induction on j in a way similar to that exposed in §1.4. In [5, §8], one can find the details and also some tricks to improve the running time of this combinatorial algorithm which, however, remains rather slow when m is large.

The second algorithm, which is more sophisticated, is based on the following remark : if $G(p_k, m) = \dfrac{Q_1 Q_2 \ldots Q_s}{q_1 q_2 \ldots q_s}$ and m is large, the least prime factor q_s of the denominator is close to $p_{k+1} - m$ while all the other primes $Q_1, \ldots, Q_s, q_1, \ldots, q_{s-1}$ are close to p_k.

More precisely, the following proposition (which is Proposition 10 of [5]) says that if $p_{k+1} - m + \delta$ is prime for some small δ and if $G(p_{k+1}, \delta)$ is not too small, then the computation of $G(p_k, m)$ is reduced to the computation of $G(p_{k+1}, m')$ for few small values of m', which can be done by the above combinatorial algorithm.

Proposition 8.1. *We want to compute $G(p_k, m)$ as defined in (8.1) with p_k odd and $p_{k+1} - p_k \le m \le p_{k+1} - 3$. We assume that we know some even non-negative integer δ satisfying*

(8.7) $$p_{k+1} - m + \delta \quad \textit{is prime,}$$

(8.8) $$G(p_{k+1}, \delta) \ge 1 + \frac{\delta}{p_{k+1}}$$

and

(8.9) $$\delta < \frac{2m}{9} < \frac{2p_{k+1}}{9}\cdot$$

If $\delta = 0$, we know from (8.6) that $G(p_k, m) = \dfrac{p_{k+1}}{p_{k+1} - m}\cdot$ If $\delta > 0$, we have

(8.10) $$G(p_k, m) = \max_{\substack{q \ \text{prime} \\ p_{k+1} - m \le q \le \widehat{q}}} \frac{p_{k+1}}{q} \, G(p_{k+1}, m - p_{k+1} + q),$$

where $\widehat{q}$ is defined by

$$\widehat{q} = \frac{p_{k+1}p_{k+2}(p_{k+1}-m+\delta)}{(p_{k+1}+\delta)(p_{k+1}-3\delta/2)} \le p_{k+2}-m+\frac{3\delta}{2}. \tag{8.11}$$

How to compute $G(p_k, m)$? The combinatorial algorithm should be tried if m is small, but it is quadratic in m and has no chance to terminate if m is larger than, say, 10^6. We have no guarantee that the conditions of Prop. 8.1 are satisfied. However in all our numerical applications, we have found $\delta < 1000$ in (8.7) (see [5, §9.2]), so that, by (8.10) and (8.11), we have

$$m - p_{k+1} + q \le m - p_{k+1} + \widehat{q} \le p_{k+2} - p_{k+1} + \frac{3\delta}{2}$$

and, in (8.10), $G(p_{k+1,}m - p_{k+1} + q)$ can be easily calculated by the combinatorial algorithm.

8.2 Description of the algorithm to compute $h(n)$

To compute $h(n)$, the first step is to determine p_k and σ_k defined by (1.8). This step is explained in § 8.2.1 and will furnish also p_{k+1} and $n' = n - p_k$.

8.2.1 Computation of p_k and σ_k

1. Compute $x = \sqrt{\mathrm{Li}^{-1}(n)}$, so that $\mathrm{Li}(x^2) = n$ and $x \sim \sqrt{n\log n}$.
2. Using Prop. 7.1, we compute $\pi_{id}(x)$ in time $O\Big(x^{2/3}/\log^2 x\Big) = O\Big(n^{1/3}/(\log n)^{-5/3}\Big)$.
3. To get σ_k, we have to add (if $\pi_{id}(x) < n$) or to subtract (if $\pi_{id}(x) > n$) to $\pi_{id}(x)$ the primes between x and p_k, calculated by sieving. In practice, this step is very short. But we are able to estimate it only under Riemann's hypothesis. By lemma 2.5, we have

   $$\mathrm{Li}(p_k^2) - \frac{5}{24\pi}p_k^{3/2}\log p_k < \sigma_k \le n < \sigma_{k+1} < \mathrm{Li}(p_{k+1}^2) + \frac{5}{24\pi}p_{k+1}^{3/2}\log p_{k+1}$$

   which implies

   $$n = \mathrm{Li}(p_k^2) + O\Big(p_k^{3/2}\log p_k\Big) \sim \mathrm{Li}(p_k^2) \sim \frac{p_k^2}{2\log p_k}.$$

   Therefore, we get $\log n \sim 2\log p_k$, $p_k \sim \sqrt{n\log n}$ and

   $$\mathrm{Li}(p_k^2) = n + O\Big(n^{3/4}(\log n)^{7/4}\Big). \tag{8.12}$$

   Further, since $x \sim p_k \sim \sqrt{n\log n}$, we have

   $$\left|n - \mathrm{Li}(p_k^2)\right| = \left|\mathrm{Li}(x^2) - \mathrm{Li}(p_k^2)\right| \sim \frac{\left|x^2 - p_k^2\right|}{2\log x} \sim 2\,|x - p_k|\sqrt{\frac{n}{\log n}},$$

   so that, from (8.12), $|\pi(x) - \pi(p_k)| \le |x - p_k| = O\Big(n^{1/4}(\log n)^{9/4}\Big)$.

8.2.2 Computation of $h(n)/N_k$

By Corollary 6.1, we have $h(n) = h_k(n)$. Let us set $n' = n - \sigma_k$. If $n' = p_{k+1} - 1$ or $n' = p_{k+1} - 2$, Proposition 5.3 yields $h(n) = N_{k+1}/2$. So, we may suppose $n' \le \sigma_{k+1} - 3$. From the definition (8.1) of function G, we have

$$h(n) = h_k(n) = N_k G(p_k, n') \tag{8.13}$$

and we compute $G(p_k, n')$ as explained in §8.1.2. In practice, the computation of $G(p_k, n)$ is very fast. However, as explained in §8.1.2, we have no estimation of the running time.

Below, are listed some values of $\dfrac{h(n)}{N_k} = G(p_k, n')$ together with p_k, $n' = n - \sigma_k$, $e = e(n)$ the largest integer such that $h(n - e) = h(n))$ and, if the algorithm of Proposition 8.1 is used, δ and Q, the number of primes used in the sum (8.10).

$$n = 10^{12},\ p_k = 5477081,\ n' = 4935150,\ e = 0,\ \delta = 18, Q = 1,$$

$$G(p_k, n') = \frac{29998525822277}{2968309525031} = \frac{5477089 \times 5477093}{5477081 \times 541951}.$$

$$n = 10^{35},\ p_k = 2898434150644708999,\ n' = 1886081812111845520,\ e = 16$$

$$\delta = 134,\ Q = 5, \quad G(p_k, n') = \frac{2898434150644709023}{1012352338532863519}.$$

The values of $h(10^a)$ for $a \le 35$ and of $h = 2^b$ for $b \le 116$ can be found on the authors's web sites [2, 10], together with the Maple or Sage programs computing $h(n)/N_k$.

9 An open question

Given n and $j < k(n)$, how to compute $h_j(n)$? We have not succeeded in solving this problem when n is too large to use the naive algorithm described in 1.4. The case $j = 2$ is already not that simple.

A first step is certainly to calculate $r = r(n, j)$ defined by (4.1), which can be done by the method of § 8.2.1. If we are lucky enough that $q = p_{j+r+1} - n' = p_{j+r+1} - (n - \sigma_{j+r} + \sigma_r)$ is prime, then the value $h_j(n) = \frac{N_{j+r+1}}{qN_r}$ is given by (4.24).

In the general case, by setting $n' = n - (\sigma_{j+r} - \sigma_r)$, one may think that $h'_j(n) = \dfrac{N_{j+r}}{N_r} G(p_{j+r}, n')$ has a good chance to be the value of $h_j(n)$. However, there are exceptions.

References

[1] E. Bach, D. Klyve, and J. P. Sorenson. Computing prime harmonic sums. *Math. Comp.*, 78(268):2283–2305, 2009.

[2] M. Deléglise's web page. `http://math.univ-lyon1.fr/~deleglis/calculs.html`.

[3] M. Deléglise and J.-L. Nicolas. Le plus grand facteur premier de la fonction de Landau. *The Ramanujan Journal*, 27:109–145, 2012. 10.1007/s11139-011-9293-2.

[4] M. Deléglise and J. Rivat. Computing $\pi(x)$: the Meissel, Lehmer, Lagarias, Miller, Odlyzko method. *Math. Comp.*, 65(213):235–245, 1996.

[5] M. Deléglise, J.-L. Nicolas, and P. Zimmermann. Landau's function for one million billions. *J. Théor. Nombres Bordeaux*, 20(3):625–671, 2008.

[6] P. Dusart. Estimates of some functions over primes without r. h., to be published. *http://arxiv.org/abs/1002.0442v1*, 2010.

[7] P. Dusart. The kth prime is greater than $k(\ln k + \ln\ln k - 1)$ for $k \geq 2$. *Math. Comp.*, 68(225):411–415, 1999.

[8] J. C. Lagarias, V. S. Miller, and A. M. Odlyzko. Computing $\pi(x)$: the Meissel-Lehmer method. *Math. Comp.*, 44(170):537–560, 1985.

[9] E. Landau. *Über die Maximalordnung der Permutationen gegebenen Grades*. Handbuch der Lehre von der Verteilung der Primzahlen, I, 2nd ed., 1953. Archiv. der Math. und Phys., Sér 3, 5 (1903), 92-103.

[10] J.-L. Nicolas's web page. `http://math.univ-lyon1.fr/~nicolas/`.

[11] L. Schoenfeld. Sharper bounds for the Chebyshev functions $\theta(x)$ and $\psi(x)$. II *Math. Comp.,* Vol. **30** no. 134, pp. 337–360, 1976.

[12] N. J. A. Sloane and S. Plouffe. *The encyclopedia of integer sequences.* Academic Press Inc., San Diego, CA, 1995. Online edition available at `http://www.research.att.com/~njas/sequences/`.

Marc Deléglise, `deleglis@math.univ-lyon1.fr`.
Jean-Louis Nicolas, `jlnicola@in2p3.fr`.
Université de Lyon, CNRS,
Institut Camille Jordan, Mathématiques, Université Claude Bernard (Lyon 1),
21 Avenue Claude Bernard, F-69622 Villeurbanne cédex, France.

$j =$	1	2	3	4	5	6
$n = 2$	2					
3	3					
4	3					
5	5	6				
6	5	6				
7	7	10				
8	7	15				
9	7	15				
10	7	21	30			
11	11	21	30			
12	11	35	42			
13	13	35	42			
14	13	35	70			
15	13	35	105			
16	13	55	105			
17	17	55	105	210		
18	17	77	110	210		
19	19	77	165	210		
20	19	91	165	210		
21	19	91	231	330		
22	19	91	231	330		
23	23	91	385	462		
24	23	143	385	462		
25	23	143	455	770		
26	23	143	455	1155		
27	23	143	455	1155		
28	23	187	455	1365	2310	
29	29	187	715	1365	2310	
30	29	221	715	1365	2730	
31	31	221	1001	1430	2730	
32	31	247	1001	2145	2730	
33	31	247	1001	2145	2730	
34	31	253	1001	3003	4290	
35	31	253	1309	3003	4290	
36	31	323	1309	5005	6006	
37	37	323	1547	5005	6006	
38	37	323	1547	5005	10010	
39	37	323	1729	5005	15015	
40	37	391	1729	6545	15015	
41	41	391	2431	6545	15015	30030
42	41	437	2431	7735	15015	30030
43	43	437	2717	7735	19635	30030
44	43	437	2717	8645	19635	30030
45	43	437	2717	8645	23205	39270
46	43	493	2717	12155	23205	39270
47	47	493	3553	12155	25935	46410
48	47	551	3553	17017	25935	46410
49	47	551	4199	17017	36465	51870
50	47	589	4199	19019	36465	51870

Figure 1: Table of $h_j(n)$.