Journal de Théorie des Nombres de Bordeaux

anciennement Séminaire de Théorie des Nombres de Bordeaux

Marc DELÉGLISE, Jean-Louis NICOLAS et Paul ZIMMERMANN

Landau's function for one million billions

Tome 20, n° 3 (2008), p. 625-671.

<http://jtnb.cedram.org/item?id=JTNB_2008__20_3_625_0>

Journal de Théorie des Nombres de Bordeaux **20** (2008), 625-671

Landau's function for one million billions

par Marc DELÉGLISE, Jean-Louis NICOLAS et Paul ZIMMERMANN

À Henri Cohen pour son soixantième anniversaire

Résumé. Soit $\mathfrak{S}_n$ le groupe symétrique sur n lettres et $g(n)$ l'ordre maximal d'un élément de $\mathfrak{S}_n$. Si la factorisation en nombres premiers de M est $M = q_1^{\alpha_1} q_2^{\alpha_2} \ldots q_k^{\alpha_k}$, nous définissons $\ell(M)$ comme étant $q_1^{\alpha_1} + q_2^{\alpha_2} + \ldots + q_k^{\alpha_k}$; il y a un siècle, E. Landau a montré que $g(n) = \max_{\ell(M) \le n} M$ et que, quand n tend vers l'infini, $\log g(n) \sim \sqrt{n \log(n)}$.

Il existe un algorithme élémentaire pour calculer $g(n)$ pour $1 \le n \le N$; son temps d'exécution est en $\mathcal{O}\left(N^{3/2}/\sqrt{\log N}\right)$ et la place mémoire nécessaire est en $\mathcal{O}(N)$; cela permet de calculer $g(n)$ jusqu'à, disons, un million. Nous donnons un algorithme pour calculer $g(n)$ pour n jusqu'à 10^{15}. L'idée principale est de considérer les nombres dits *ℓ-superchampions.* Des nombres similaires, les *nombres hautement composés supérieurs*, ont été introduits par S. Ramanujan pour étudier les grandes valeurs de la fonction nombre de diviseurs $\tau(n) = \sum_{d \mid n} 1$.

Abstract. Let $\mathfrak{S}_n$ denote the symmetric group with n letters, and $g(n)$ the maximal order of an element of $\mathfrak{S}_n$. If the standard factorization of M into primes is $M = q_1^{\alpha_1} q_2^{\alpha_2} \ldots q_k^{\alpha_k}$, we define $\ell(M)$ to be $q_1^{\alpha_1} + q_2^{\alpha_2} + \ldots + q_k^{\alpha_k}$; one century ago, E. Landau proved that $g(n) = \max_{\ell(M) \le n} M$ and that, when n goes to infinity, $\log g(n) \sim \sqrt{n \log(n)}$.

There exists a basic algorithm to compute $g(n)$ for $1 \le n \le N$; its running time is $\mathcal{O}\left(N^{3/2}/\sqrt{\log N}\right)$ and the needed memory is $\mathcal{O}(N)$; it allows computing $g(n)$ up to, say, one million. We describe an algorithm to calculate $g(n)$ for n up to 10^{15}. The main idea is to use the so-called *ℓ-superchampion numbers.* Similar numbers, the *superior highly composite numbers*, were introduced by S. Ramanujan to study large values of the divisor function $\tau(n) = \sum_{d \mid n} 1$.

Manuscrit reçu le 27 février 2008.

Mots clefs. Arithmetical function, symmetric group, maximal order, highly composite number.

1. Introduction

1.1. Known results about Landau's function. For $n \geq 1$, let $\mathfrak{S}_n$ denote the symmetric group with n letters. The order of a permutation of $\mathfrak{S}_n$ is the least common multiple of the lengths of its cycles. Let us call $g(n)$ the maximal order of an element of $\mathfrak{S}_n$. As far as we know, E. Landau (cf. [9]) was the first to study the function $g(n)$, which was called *Landau's function* in [21], whence the title of this paper.

If the standard factorization of M into primes is $M = q_1^{\alpha_1} q_2^{\alpha_2} \dots q_k^{\alpha_k}$, we define $\ell(M)$ to be the additive function defined by

$$\ell(M) = q_1^{\alpha_1} + q_2^{\alpha_2} + \dots + q_k^{\alpha_k}. \tag{1}$$

E. Landau proved in [9] that

$$g(n) = \max_{\ell(M) \leq n} M \tag{2}$$

which implies

$$\ell(g(n)) \leq n \tag{3}$$

and for all positive integers n, M

$$\ell(M) \leq n \Longrightarrow M \leq g(n) \quad \Longleftrightarrow \quad M > g(n) \Longrightarrow \ell(M) > n. \tag{4}$$

P. Erdős and P. Turán proved in [6] that

$$M \text{ is the order of some element of } \mathfrak{S}_n \iff \ell(M) \leq n. \tag{5}$$

E. Landau also proved in [9] that

$$\log g(n) \sim \sqrt{n \log n}, \qquad n \to \infty. \tag{6}$$

This asymptotic estimate was improved by S. M. Shah [29] and M. Szalay [30]; in [12], it is shown that

$$\log g(n) = \sqrt{\mathrm{Li}^{-1}(n)} + \mathcal{O}(\sqrt{n} \exp(-a\sqrt{\log n})) \tag{7}$$

for some $a > 0$; Li^{-1} denotes the inverse function of the integral logarithm.

The survey paper [14] of W. Miller is a nice introduction to $g(n)$; it contains elegant and simple proofs of (2), (5) and (6).

J.-P. Massias proved in [11] that for $n \geq 1$

$$\log g(n) \leq \frac{\log g(1319366)}{\sqrt{1319366 \log(1319366)}} \sqrt{n \log n} \approx 1.05313 \sqrt{n \log n}. \tag{8}$$

In [13] more accurate effective results are given, including

$$\log g(n) \geq \sqrt{n \log n}, \qquad n \geq 906 \tag{9}$$

and

$$\log g(n) \leq \sqrt{n \log n} \left(1 + \frac{\log \log n - 0.975}{2 \log n}\right), \qquad n \geq 4. \tag{10}$$

Let $P^+(g(n))$ denote the greatest prime factor of $g(n)$. In [8], J. Grantham proved

$$P^+(g(n)) \leq 1.328\sqrt{n \log n}, \qquad n \geq 5. \tag{11}$$

Some other functions similar to $g(n)$ were studied in [7], [10], [22], [30] and [31].

1.2. Computing Landau's function. A table of Landau's function up to 300 is given at the end of [18]. It has been computed with the algorithm described and used in [19] to compute $g(n)$ up to 8000. By using similar algorithms, a table up to 32000 is given in [15], and a table up to 500000 is mentioned in [8]. The algorithm given in [19] will be referred in this paper as the basic algorithm. We shall recall it in Section 2. It can be used to compute $g(n)$ for n up to, say, one million, maybe a little more. It cannot compute $g(n)$ without calculating simultaneously $g(n')$ for $1 \leq n' \leq n$.

If we look at a table of $g(n)$ for $31000 \leq n \leq 31999$ (such a table can be easily built by using the Maple procedure given in Section 2), we observe three parts among the prime divisors of $g(n)$. More precisely, let us set

$$g(n) = \prod_p p^{\alpha_p}, \qquad g^{(1)}(n) = \prod_{p \leq 17} p^{\alpha_p},$$
$$g^{(2)}(n) = \prod_{19 \leq p \leq 509} p^{\alpha_p}, \qquad g^{(3)}(n) = \prod_{p > 509} p^{\alpha_p};$$

the middle part $g^{(2)}(n)$ is constant (and equal to $\prod_{19 \leq p \leq 509} p$) for all n between 31000 and 31999, while the first part $g^{(1)}(n)$ takes only 18 values, and the third part $g^{(3)}(n)$ takes 92 values.

So, if n' is in the neighbourhood of n, $g(n')/g(n)$ is a fraction which is the product of a *prefix* (made of small primes) and a *suffix* (made of large primes).

The aim of this article is to make precise this remark to get an algorithm able to compute $g(n)$ for some fixed n up to 10^{15}.

1.3. The new algorithm. Let $\tau(n) = \sum_{d \mid n} 1$ be the divisor function. To study *highly composite numbers* (that is the n's such that $m < n$ implies $\tau(m) < \tau(n)$), S. Ramanujan (cf. [24, 25, 20]) has introduced the *superior highly composite numbers* which maximize $\tau(n)/n^{\varepsilon}$ for some $\varepsilon > 0$. This definition can be extended to function ℓ: N is said to be ℓ-superchampion if it minimizes $\ell(N) - \rho \log(N)$ for some $\rho > 0$. These numbers will be defined (cf. (4.1)) and discussed in Section 4: they are easy to compute and have the property that, if $n = \ell(N)$, then $g(n) = N$.

If N minimizes $\ell(N) - \rho \log(N)$, we call *benefit* of an integer M the non-negative quantity $\operatorname{ben}(M) = \ell(M) - \ell(N) - \rho \log(M/N)$ (cf. (6.1)). If n is not too far from $\ell(N)$, a relatively small bound can be obtained for $\operatorname{ben} g(n)$, and this allows computing it. This notion of benefit will be discussed in Section 6.

To compute $g(n)$, the main steps of our algorithm are

1. Determine the two consecutive ℓ-superchampion numbers N and N' such that $\ell(N) \le n < \ell(N')$ and their common parameter ρ (cf. Section 5).1
2. For a guessed value B', determine a set $\mathcal{D}(B')$ of plain prefixes whose benefit is smaller than B' (cf. Section 7.1 and Section 7.2).
3. Use the set $\mathcal{D}(B')$ to compute an upper bound B such that $\operatorname{ben} g(n) \le \operatorname{ben} g(n) + n - \ell(g(n)) \le B$ (cf. Section 7.3); note that, from (3), $\ell(g(n)) \le n$ holds.
4. Determine $\mathcal{D}(B)$, a set containing the plain prefix of $g(n)$. If $B < B'$, to get $\mathcal{D}(B)$, we just have to remove from $\mathcal{D}(B')$ the elements whose benefit is bigger than B. If $B > B'$, we start again the algorithm described in Section 7.2 to get $\mathcal{D}(B')$ with a new value of B' greater than B.
5. Compute a set containing the normalized prefix of $g(n)$ (cf. Sections 7.7, 7.8 and 7.9).
6. Determine the suffix of $g(n)$ by using the function $G(p_k, m)$ introduced in Section 1.4 and discussed in Sections 8 and 9.

In the sequel of our article, “ step ” will refer to one of the above six steps, and “ the algorithm ” will refer to the algorithm sketched in Section 1.3.

On the web site of the second author, there is a Maple code of this algorithm where each instruction is explained according to the notation of this article.

If we want to calculate $g(n)$ for consecutive values $n = n_1, n = n_1 + 1, \dots, n = n_2$, most of the operations of the algorithm are similar and can be put in common; however, due to some technical questions, it is more difficult to treat this problem, and here, we shall restrict ourselves to the computation of $g(n)$ for one value of n.

To compute the first 5000 highly composite numbers, G. Robin (cf. [27]) already used a notion of benefit similar to that introduced in this article.

1.4. The function $G(p_k, m)$. In step 6, the computation of the suffix of $g(n)$ leads to the function $G(p_k,m)$, defined by

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

$$G(p_k, m) = \max \frac{Q_1 Q_2 \dots Q_s}{q_1 q_2 \cdots q_s} \tag{12}$$

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

$$3 \leq q_s < q_{s-1} < \ldots < q_1 \leq p_k < p_{k+1} \leq Q_1 < Q_2 < \ldots < Q_s \tag{13}$$

and

$$\sum_{i=1}^{s} (Q_i - q_i) \leq m. \tag{14}$$

This function $G(p_k, m)$ is interesting in itself. It satisfies

$$\ell(G(p_k, m)) \leq m. \tag{15}$$

We study it in Section 8, where a combinatorial algorithm is given to compute its value when m is not too large. For m large, a better algorithm is given in Section 9.

Let us denote by $\mu_1(n) < \mu_2(n) < \ldots$ the increasing sequence of the primes which do not divide $g(n)$, and by $P(n)$ the largest prime factor of $g(n)$. It is shown in [17] that $\lim_{n\to\infty} P(n)/\mu_1(n) = 1$. We may guess from Proposition 10 that $\mu_1(n)$ can be much smaller than $P(n)$ while $\mu_2(n)$ is closer to $P(n)$. It seems difficult to prove any result in this direction.

1.5. The running time. Though we have the feeling that the algorithm presented in this paper (and implemented in MAPLE) yields the value of $g(n)$ for all n's up to 10^{15} (and possibly for greater n's) in a reasonable time, it is not proved to do so.

Indeed, we do not know how to get an effective upper bound for the *benefit* of $g(n)$ (see Sections 6, 7.3 and 11.1) and in the second and third steps, what we do is just, for a given n, to provide such an upper bound $B = B(n)$ by an experimental way.

In the fourth step, the algorithm determines a set $\mathcal{D}(B)$ of plain prefixes (cf. Sections 7.2 and 7.3). It turns out that the number $\nu(n)$ of these prefixes is rather small and experimentally satisfies $\nu(n) = O(n^{0.3})$ (cf. (55)); but we do not know how to prove such a result, and it might exist some values of n for which $\nu(n)$ is much larger.

Let us now analyze each of the six steps described in Section 1.3.

The first step determines the greatest superchampion number N such that $\ell(N) \leq n$. Let $S(x) = \sum_{p \leq x} p$ be the sum of the primes up to x. The main part of this step is to compute $S(x)$ for x close to $\sqrt{n \log n}$. In our MAPLE program, by Eratosthenes' sieve, we have precomputed a function close to $S(x)$, the details are given in Section 5. However, a faster way exists to evaluate $S(x)$. By extending Meissel's technique to compute $\pi(x) = \sum_{p \leq x} 1$, (cf. [3]), M. Deléglise is able to compute $\sum_{p \leq x} f(p)$ where f is a multiplicative function. E. Bach (cf. [1, 2]) has considered a wider class of functions for which this method also works. By his algorithm, M.

Deléglise has computed $S(10^{18})$, and computing $S(x)$ costs $O(x^{2/3}/\log^2 x)$. We hope to implement soon this new evaluation of $S(x)$ in our first step.

The second and the fourth steps compute respectively $\mathcal{D}(B')$ and $\mathcal{D}(B)$. If B' is "well" chosen, we may hope that $\mathrm{Card}(\mathcal{D}(B'))$ is not much larger than $\nu(n) = \mathrm{Card}(\mathcal{D}(B))$. The running time of the computation of $\mathcal{D}(B')$ as explained in Section 7.2 could be larger than $\nu(n)$. For $n \approx 10^{20}$, most of the time of the computation of $g(n)$ is spent in the second and fourth steps. But any precise estimation of these steps seems unaccessible.

The running time of the third step is $O(\mathrm{Card}(\mathcal{D}(B')))$, and we may hope that it is $O(\nu(n))$.

In practice, the fifth step (finding the possible normalized prefixes) is fast. For every plain prefix $\widehat{\pi}$, Inequations (80) have at most one solution, and the cost of this step is $O(\nu(n))$.

The sixth and last step also is fast. Under the strong assumption that $\delta_1(p)$ is polynomial in $\log p$ (see (117)), for any m, the computation of $G(p,m)$ (where p is a prime satisfying $p \approx \sqrt{n \log n}$) is polynomial in $\log n$, and the number of normalized prefixes surviving the fight (cf. Section 7.9) seems to be bounded (we have no examples of more than three of them), so that (see Section 7.8) this step might be polynomial in $\log n$.

1.6. Plan of the paper. In Section 3, some mathematical lemmas are given. The various steps of the algorithm presented in Section 1.3 are explained in Sections 4-9; Section 10 presents some results while Section 11 describes five open problems.

1.7. Notation. We denote by $\mathcal{P} = \{2, 3, 5, 7, \ldots\}$ the set of primes, by $p \in \mathcal{P}$ a generic prime, by p_i the i-th prime and by $v_p(N)$ the p-adic valuation of N, that is the greatest integer α such that p^α divides N. Q_i and q_i also denote primes, except in Lemma 1 which is stated in a more general form, but which is used with Q_i and q_i primes. The integral part of a real number t is denoted by $\lfloor t \rfloor = \max_{n\in\mathbb{Z}, n\le t} n$. The additive function ℓ can be easily extended to a rational number by setting $\ell(A/B) = \ell(A) - \ell(B)$ (with A and B coprime).

2. The basic algorithm

2.1. The first version. For $j \geq 0$, let us denote by $\mathcal{S}_j$ the set of numbers having only $p_1, p_2, \ldots, p_j$ as prime divisors

$$\mathcal{S}_j = \{M; p \mid M \implies p \leq p_j\}. \tag{16}$$

We have $\mathcal{S}_0 = \{1\}$, $\mathcal{S}_1 = \{1, 2, 4, 8, 16, \ldots\}$. The algorithm described in [19] computes the functions

$$g_j(n) = \max_{M\in\mathcal{S}_j,\ \ell(M)\leq n} M \tag{17}$$

which obviously satisfy the induction relation

$$g_j(n) = \max\left[g_{j-1}(n), p_j g_{j-1}(n-p_j), \ldots, p_j^k g_{j-1}(n-p_j^k)\right] \tag{18}$$

where k is the largest integer such that $p_j^k \leq n$, and $g_0(n) = 1$ for all $n \geq 0$. Using the upper bound (11), we write the following MAPLE procedure:

Algorithm 1 The basic algorithm: this MAPLE procedure computes $g(n)$ for $0 \leq n \leq N$ and stores the results in table g.

```
gden := proc(N) local n, g, pmax, p, k, a
for n from 0 to N do
   g[n] := 1
endo;
pmax := floor(1.328 ⋆ eval(sqrt(N ⋆ log N)));
p := 2;
while p ≤ pmax do
   for n from N to p by −1 do
      for k from 1 while p^k ≤ n do
         a := p^k ⋆ g[n − p^k];
         if g[n] < a then
            g[n] := a
         end if
      endo
   endo;
   p := nextprime(p)
end while;
end;
```

The running time of this procedure is 13 hours for $N = 10^6$ on a 3 Ghz Pentium 4 with a storage of 337 Mo. To compute $g(n)$, $1 \leq n \leq N$, the theoretical running time is $\mathcal{O}\left(N^{3/2}/\sqrt{\log N}\right)$ and the needed memory is $\mathcal{O}(N)$ integers of size $\exp(O(\sqrt{N \log N}))$.

2.2. The merging and pruning algorithm. The above algorithm takes a very long time to compute $g_j(n)$ when j is small. It is better to represent $(g_j(n))_{n\geq 1}$ by a list $L_j = [[M_1, l_1], \ldots, [M_i, l_i], \ldots]$ (where $l_i = \ell(M_i)$) ordered so that $M_{i+1} > M_i$ and $l_{i+1} > l_i$. If $l_i \leq n < l_{i+1}$, then $g_j(n) = M_i$. So, $L_0 = [[1, 0]]$ and $L_1 = [[1, 0], [2, 2], [4, 4], [8, 8], \ldots]$.

To calculate L_{j+1} from L_j we construct the list of all elements $[M_i p_{j+1}^a, l_i + \ell(p_{j+1}^a)]$ for all elements $[M_i, l_i] \in L_j$ and $a \geq 0$ such that $l_i + \ell(p_{j+1}^a) \leq N$. We sort this new list with respect to the first term of the elements (merge sort is here specially recommended) to get a list

$\Lambda = [[K_1, \lambda_1], [K_2, \lambda_2], \ldots]$ with $K_1 < K_2 < \ldots$ Now, to take (18) into account, we have to prune the list Λ: if $K_r < K_s$ and $\lambda_r \geq \lambda_s$, we take off the element $[K_r, \lambda_r]$ from the list Λ. The list L_{j+1} will be the pruned list of Λ.

3. Two lemmas

Lemma 1. *Let s be a non-negative integer, and $t_1, q_1, q_2, \ldots, q_s, Q_1, Q_2, \ldots, Q_s$ be real numbers satisfying*

$$0 < t_1 \leq q_s < q_{s-1} < \ldots < q_1 < Q_1 < Q_2 < \ldots < Q_s. \tag{19}$$

If we set $S = \sum_{i=1}^{s} Q_i - q_i$, then the following inequality holds:

1. $$\frac{Q_1 Q_2 \ldots Q_s}{q_1 q_2 \ldots q_s} \leq \exp\left(\frac{S}{t_1}\right).$$

Moreover, if $s \geq 1$ and $S < Q_1$, we have

2. $$\frac{Q_1 Q_2 \ldots Q_s}{q_1 q_2 \ldots q_s} \leq \frac{Q_s}{Q_s - S} < \frac{Q_{s-1}}{Q_{s-1} - S} < \ldots < \frac{Q_1}{Q_1 - S}$$

with the first inequality in 2. strict when $s \geq 2$.

Proof. Lemma 1 is a slight improvement of Lemma 3 of [18] where, in 2., only the upper bound $Q_1/(Q_1 - S)$ was given. Point 1. is easy by applying $1 + u \leq \exp u$ to $u = Q_i/q_i - 1$. Let us prove 2. by induction. For $s = 1$, 2. is an equality. Let us assume that $s \geq 2$. Setting $S' = \sum_{i=2}^{s} Q_i - q_i = S - (Q_1 - q_1)$, we have $S' < S < Q_1 < Q_s$ and by the induction hypothesis, we get

$$\frac{Q_1 Q_2 \ldots Q_s}{q_1 q_2 \ldots q_s} = \frac{Q_1}{q_1} \frac{Q_2 \ldots Q_s}{q_2 \ldots q_s} \leq \frac{Q_1}{q_1} \frac{Q_s}{Q_s - S'}. \tag{20}$$

We shall use the following principle:

Principle 1. *If x and y add to a constant, the product xy decreases when $|y - x|$ increases.*

We have $Q_s - S' \leq Q_s - (Q_s - q_s) = q_s < q_1$, and using Principle 1, we get by increasing q_1 to Q_1 and decreasing $Q_s - S'$ to $Q_s - S$

$$q_1(Q_s - S') > Q_1(Q_s - S)$$

which, from (20), proves 2. □

Lemma 2. *Let $x > 4$ and $y = y(x)$ be defined by $\dfrac{y^2 - y}{\log y} = \dfrac{x}{\log x}\cdot$ The function y is an increasing function satisfying $y(x) > 2$ and*

1. $y(x) = \sqrt{\dfrac{x}{2}}\left(1 - \dfrac{\log 2}{2\log x} - \dfrac{(4 + \log 2)\log 2}{8\log^2 x} + \mathcal{O}\left(\dfrac{1}{\log^3 x}\right)\right)$, $x \to \infty$
2. $y(x) < \sqrt{x}$ *for* $x > 4$.

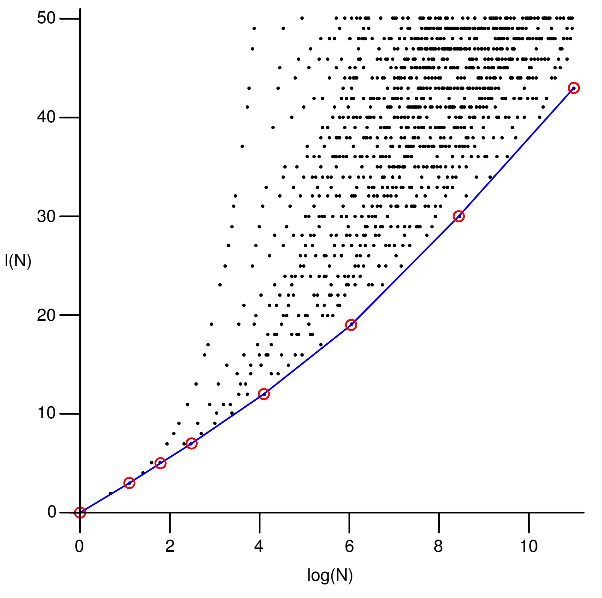

FIGURE 1. The points $(\log(N), \ell(N))$, with $\ell(N) \le 50$, for $1 \le N \le 60060$.

3. $y(x) \le \sqrt{\dfrac{x}{2}}$ *for* $x \ge 80$.

Proof. 1. and 3. are proved in [12], p. 227. Since $t \mapsto (t^2 - t)/\log t$ is increasing for $t > 1$, in order to show 2., one should prove $\dfrac{x-\sqrt{x}}{\frac{1}{2}\log x} > \dfrac{x}{\log x}$ which holds for $x > 4$. $\square$

4. The superchampion numbers

Definition 2. *An integer N is said ℓ-superchampion (or more simply superchampion) if there exists $\rho > 0$ such that, for all $M \ge 1$*

$$\ell(M) - \rho \log M \ge \ell(N) - \rho \log N. \tag{21}$$

When this is the case, we say that N is a ℓ-superchampion associated *to ρ.*

Geometrically, if we represent $\log M$ in abscissa and $\ell(M)$ in ordinate, the straight line of slope ρ going through the point $(\log M, \ell(M))$ has an intersep equal to $\ell(M) - \rho \log(M)$ and so, the superchampion numbers are the vertices of the convex envelop of all these points (see FIG. 1).

Similar numbers, the so-called *superior highly composite numbers* were first introduced by S. Ramanujan (cf. [24]). The ℓ-superchampion numbers were already used in [17, 18, 11, 12, 13, 21, 22]. The first ones are (with, in the third column, the corresponding values of ρ) shown in Fig. 2.

Lemma 3. *If N is an ℓ-superchampion, the following property holds:*

$$N = g(\ell(N)). \tag{22}$$

Proof. Indeed, let N be any positive number and $n = \ell(N)$; it follows from (4) that $N \le g(n) = g(\ell(N))$. If moreover N is a ℓ-superchampion, then, for all M such that $\ell(M) \le n = \ell(N)$, from (21), we have $\rho \log M \le \rho \log N + \ell(M) - \ell(N) \le \rho \log N$ which implies $M \le N$, and thus, from (2), (22) holds. □

Definition 3.

1. *For each prime $p \in \mathcal{P}$, let us define the sets*

$$\mathcal{E}'_p = \left\{\frac{p}{\log p}\right\}, \quad \mathcal{E}''_p = \left\{\frac{p^2-p}{\log p}, \dots, \frac{p^{i+1}-p^i}{\log p}, \dots\right\}, \quad \mathcal{E}_p = \mathcal{E}'_p \cup \mathcal{E}''_p. \tag{23}$$

2. *And we define*

$$\mathcal{E}' = \bigcup_{p\in\mathcal{P}} \mathcal{E}'_p, \quad \mathcal{E}'' = \bigcup_{p\in\mathcal{P}} \mathcal{E}''_p \quad \textit{and} \quad \mathcal{E} = \mathcal{E}' \cup \mathcal{E}''. \tag{24}$$

Remark: Note that all the elements of $\mathcal{E}_p$ are distinct at the exception, for $p = 2$, of $\dfrac{2}{\log 2} = \dfrac{2^2-2}{\log 2}$ and that, for $p \ne q$, $\mathcal{E}_p \cap \mathcal{E}_q = \emptyset$ holds.

Lemma 4. *Let ρ a real number.*

1. *If $\rho \in \mathcal{E}_p$, $\rho \ne \frac{2}{\log 2}$, there exist exactly 2 superchampion numbers associated to ρ. Let N_ρ be the smaller one and N_ρ^+ the bigger one. Then $N_\rho^+ = pN_\rho$ and*

$$N_\rho = \prod_{p/\log p < \rho} p^{\alpha_p} \quad \textit{with} \quad \alpha_p = \begin{cases} 1 & \textit{if} \quad \dfrac{p}{\log p} < \rho \le \dfrac{p^2-p}{\log p} \\[2ex] i & \textit{if} \quad \dfrac{p^i-p^{i-1}}{\log p} < \rho \le \dfrac{p^{i+1}-p^i}{\log p} \end{cases} \tag{25}$$

$$N_\rho^+ = \prod_{p/\log p \le \rho} p^{\alpha_p^+} \quad \textit{with} \quad \alpha_p^+ = \begin{cases} 1 & \textit{if} \quad \dfrac{p}{\log p} \le \rho < \dfrac{p^2-p}{\log p} \\[2ex] i & \textit{if} \quad \dfrac{p^i-p^{i-1}}{\log p} \le \rho < \dfrac{p^{i+1}-p^i}{\log p} \end{cases} \tag{26}$$

2. *If* $\rho = \dfrac{2}{\log 2} = \dfrac{2^2-2}{\log 2} \in \mathcal{E}$, *there exist* 3 *superchampion numbers associated to* ρ: N_ρ *defined by* (25) *is equal to* 3, N_ρ^+ *defined by* (26) *is equal to* 12 *and the third one is* 6.
3. *If* $\rho \notin \mathcal{E}$, *there exists a unique superchampion number* $N_\rho = N_\rho^+$ *associated to* ρ. *Its value is given by both formulas* (25) *and* (26). *Let* ρ' *and* ρ'' *be the two consecutive elements of* $\mathcal{E}$ *such that* $\rho' < \rho < \rho''$. *Then we have* $N_\rho = N_{\rho''} = N_{\rho'}^+$.
4. *Let us consider the sequence* $\rho^{(i)}$ *defined by* $\rho^{(0)} = -\infty$, $\rho^{(1)} = 3/\log 3$, $\rho^{(2)} = 2/\log 2$, $\rho^{(3)} = (2^2-2^1)/\log 2 = \rho^{(2)}$, $\rho^{(4)} = 5/\log 5$ *and such that* $\left\{\rho^{(i)}, i \geq 1\right\} = \mathcal{E}$ *and* $\rho^{(i)} > \rho^{(i-1)}$ *for* $i \geq 4$. *If* $N^{(0)} = 1$, $N^{(1)} = 3$, $N^{(2)} = 6$, $N^{(3)} = 12$, $N^{(4)} = 60$, *etc... is the increasing sequence of all superchampion numbers, it satisfies:*
   i. *For* $i \geq 0$, $N^{(i)}$ *divides* $N^{(i+1)}$ *and the quotient* $N^{(i+1)}/N^{(i)}$ *is a prime number. The number of prime factors of* $N^{(i)}$, *counting them with multiplicity, is equal to* i.
   ii. *For* $i \neq 2$, *we have* $N^{(i)} = N^+_{\rho^{(i)}} = N_{\rho^{(i+1)}}$ *where* $N^+_{\rho^{(i)}}$ *and* $N_{\rho^{(i+1)}}$ *are defined respectively in* (25) *and* (26).
   iii. *For all* $i \geq 0$, $N^{(i)}$ *is associated to* ρ *if and only if* $\rho^{(i)} \leq \rho \leq \rho^{(i+1)}$.
   iv. *If* $i \neq 1$ *(i.e.,* $N^{(i)} \neq 3$*), then* $v_p(N^{(i)})$ *is a non-increasing function of the prime* p.

Proof. We are looking for an $N = \prod p^{\alpha_p}$ which minimizes $F(N) = \ell(N) - \rho \log N$.

An arithmetic function h is said *additive* if $h(M_1M_2) = h(M_1) + h(M_2)$ when M_1 and M_2 are coprime. The functions log and ℓ are additive. Thus F is additive, and to minimize $F(N) = \sum_{p \mid N} F(p^{v_p(N)})$ we have to minimize $F(p^\alpha)$ on α for each $p \in \mathcal{P}$. We have $F(1) = 0$ and for p prime and $i \geq 1$, $F(p^i) = p^i - \rho\, i \log p$. The difference

$$F(p^{i+1}) - F(p^i) = \begin{cases} p - \rho \log p & \text{if } i = 0 \\ p^i(p-1) - \rho \log p & \text{if } i > 0 \end{cases} \tag{27}$$

is a non-decreasing function of i that tends to $+\infty$ with i. Thus if $F(p) = F(p) - F(0) = p - \rho \log p > 0$, the smallest value of $F(p^\alpha)$ is 0 obtained for $\alpha = 0$. If $F(p) \leq 0$ let i be the largest positive integer such that $F(p^i) - F(p^{i-1}) \leq 0$. Then the smallest value of $F(p^\alpha)$ is obtained on the set $\{j \leq i \mid F(p^j) = F(p^i)\}$ and the number of choices for α_p is the cardinal of this set.

This proves that we have more than one choice for the exponent α_p if and only if there exists $i \geq 0$ such that $F(p^i) = F(p^{i+1})$. Due to (27) this

N	$\ell(N)$				
1	0	$-\infty$	$<$	$\rho \le 3/\log 3$	≈ 2.73
3	3	$3/\log 3$	$\le$	$\rho \le 2/\log 2$	≈ 2.89
6	5			$\rho = (2^2-2^1)/\log 2$	≈ 2.89
12	7	$2/\log 2$	$\le$	$\rho \le 5/\log 5$	≈ 3.11
60	12	$5/\log 5$	$\le$	$\rho \le 7/\log 7$	≈ 3.60
420	19	$7/\log 7$	$\le$	$\rho \le 11/\log 11$	≈ 4.59
4620	30	$11/\log 11$	$\le$	$\rho \le 13/\log 13$	≈ 5.07
60060	43	$13/\log 13$	$\le$	$\rho \le (3^2-3^1)/\log 3$	≈ 5.46

FIGURE 2. The first ℓ-superchampion numbers.

is the case if and only if $\rho \in \mathcal{E}_p$. Moreover, the sets $\mathcal{E}_p$ being disjoint, there exists at most one p for which there is more than one choice for α_p.

If $p \ge 3$ we have $p < (p^2-p) < (p^3-p^2) < \cdots$ and there is at most one i such that $F(p^{i+1}) - F(p^i) = 0$, so there are at most two choices for α_p.

For $p = 2$ we have $2 = 2^2-2 < 2^3-2^2 < \cdots$ and for $\rho = 2/\log 2$ we have $F(1) = F(2) = F(2^2)$, so we can choose for α_2 every one of the three values $0, 1, 2$. With this value of ρ we have $F(3) = 3 - (2/\log 2)\log 3 < 0$ and $F(p) > 0$ for $p \ge 5$. Thus there are 3 superchampion numbers associated to $\rho = 2/\log 2$ which are $3, 6, 12$. This proves 1., 2., 3. and 4.; for more details, see [18].

□

Lemma 5. *Let ρ satisfy $\rho \ge 5/\log 5 \approx 3.11$. There exists a unique decreasing sequence $(x_j) = (x_j(\rho))$ such that $x_1 \ge \exp(1)$ and, for all $j \ge 2$, x_j satisfies $x_j > 1$ and*

$$\frac{x_j^j - x_j^{j-1}}{\log x_j} = \frac{x_1}{\log x_1} = \rho. \tag{28}$$

We have also

$$x_1 \ge 5 \quad \textit{and} \quad x_2 > 2. \tag{29}$$

Proof. The uniqueness of x_1 results from $\rho > \exp(1)$ and the fact that $t \mapsto t/\log t$ is an increasing bijection of $[\exp(1), +\infty[$. The uniqueness of x_j for $j \ge 2$ comes from the fact that $t \mapsto (t^j - t^{j-1})/\log t = t^{j-1}(t-1)/\log t$ is an increasing bijection of $]1, +\infty[$. The inequality $x_j > x_{j+1}$ for $j \ge 2$ comes from the increase of $j \mapsto (t^j - t^{j-1})/\log t$ for each $t > 1$.

Let us prove that $x_1 > x_2$. The definition (28) of x_2 implies

$$\frac{x_2^2 - x_2}{\log x_2} = \rho > \frac{2}{\log 2} = \frac{2^2 - 2}{\log 2} \approx 2.89 .$$

With the increase of $t \mapsto (t^2-t)/\log t$ this proves $x_2 > 2$. Thus $x_2^2 - x_2 > x_2$, and therefore

$$\frac{x_2}{\log x_2} < \frac{x_2^2 - x_2}{\log x_2} = \rho = \frac{x_1}{\log x_1}$$

which, with the increase of $t \mapsto t/\log t$ on $[\exp(1), +\infty[$ yields $x_2 > x_1$ and the decrease of (x_n). Finally $x_1/\log x_1 = \rho \geq 5/\log 5$ gives $x_1 \geq 5$. □

Proposition 1. *Let ρ be a real number satisfying $\rho \geq 5/\log 5$, N_ρ the smallest superchampion number associated to ρ and N_ρ^+ the largest superchampion number associated to ρ (cf. Lemma 4). Then, with x_j as introduced in Lemma 5, we have*

$$N_\rho = \prod_{j\geq 1} \prod_{x_{j+1} \leq p < x_j} p^j \quad \textit{and} \quad N_\rho^+ = \prod_{j\geq 1} \prod_{x_{j+1} < p \leq x_j} p^j. \tag{30}$$

Proof. Due to (25), $\alpha_p = 1$ holds if and only we have

$$\frac{p}{\log p} < \rho \leq \frac{p^2 - p}{\log p}, \tag{31}$$

and by the definition (28) of x_1 and x_2, this is equivalent to

$$\frac{p}{\log p} < \frac{x_1}{\log x_1} \quad \text{and} \quad \frac{x_2^2 - x_2}{\log x_2} \leq \frac{p^2 - p}{\log p}.$$

By the increase of $t \mapsto t/\log t$ on $[\exp(1), +\infty[$ and $t \mapsto (t^2 - t)/\log t$ on $[1, +\infty[$, this proves that for $p \geq \exp(1)$, $\alpha_p = 1$ holds if and only if $x_2 \leq p < x_1$. It remains to prove that, when $p = 2$, this equivalence is still true. In this case, $2/\log 2 = (4-2)/\log 2$, and (31) is never satisfied. By (29) we have $x_2 > 2$, and $x_2 \leq 2 < x_1$ is false. Thus, for every prime p, we have $\alpha_p = 1$ if and only if $x_2 \leq p < x_1$.

For $i \geq 2$, $\alpha_p = i$ if and only if $\dfrac{p^i - p^{i-1}}{\log p} < \rho \leq \dfrac{p^{i+1} - p^i}{\log p}$, and, by the definition (28) of x_i and x_{i+1} this is equivalent to

$$\frac{p^i - p^{i-1}}{\log p} < \frac{x_i^i - x_i^{i-1}}{\log x_i} \quad \text{and} \quad \frac{x_{i+1}^{i+1} - x_{i+1}^i}{\log x_{i+1}} \leq \frac{p^{i+1} - p^i}{\log p}$$

or $x_{i+1} \leq p < x_i$. This proves the first equality (30). The second one can be proved by the same way. □

i	$T[i].q$	$T[i].j$	$T[i].p$	$T[i].\ell$
1	2	2	3	7
2	3	2	13	49
3	2	3	13	53
4	2	4	43	301
5	5	2	47	368
6	3	3	67	626
7	7	2	97	1160
8	2	5	107	1487
9	11	2	251	6307
10	2	6	251	6339
11	3	4	271	7453

Figure 3. The first elements of table T associated to E_2.

5. First step of the computation of $g(n)$: getting ρ, N, N'.

5.1. Fixing our notation. When $\rho = 5/\log 5$ we have $N_\rho = 12$ and $\ell(N_\rho) = 7$ (see Fig. 2).

Definition 4. *From now on, $n \geq 7$ will be a fixed integer, and our purpose is to compute $g(n)$. We will denote by ρ the unique real number $\rho \in \mathcal{E}$ such that $\rho \geq 5/\log 5$ and*

$$\ell(N_\rho) \leq n < \ell(N_\rho^+). \tag{32}$$

We will also fix the following notation.

1. $N = N_\rho$, $N' = N_\rho^+$ *and* $N = \prod_p p^{\alpha_p}$ *is the standard factorization of* N.
2. *We define* $x_1 = x_1(\rho) \geq 5$ *and* $x_2 = x_2(\rho) > 2$ *by* (28).
3. *Let* p_k *be the largest prime factor of* $N = N_\rho$. *It follows from* (30) *that*

$$p_k < x_1 \leq p_{k+1} \tag{33}$$

and, actually, $x_1 = p_{k+1}$ *unless* $\rho \in \mathcal{E}''$ *(in this case* $p_k < x_1 < p_{k+1}$*).*

4. *Let us define* B_1 *by*

$$B_1 = \min\left(x_2^2 - 2x_2, \frac{x_1}{2} - \sqrt{x_1}\right) > 0. \tag{34}$$

We have

$$2 < x_2 < \sqrt{x_1} < \rho < x_1. \tag{35}$$

Let us prove (35). Inequalities (29) give $2 < x_2$. With Lemma 2, Point 2., it yields $x_2 < \sqrt{x_1}$. Since for all $t > 1$, $\sqrt{t}/\log t > e/2 > 1$ we have $\sqrt{x}_1/\log x_1 > 1$ and thus $\rho = x_1/\log x_1 > \sqrt{x}_1$.

5.2. The superchampion algorithm. Given n, as already said, the first step in our computation of $g(n)$ is to calculate $\rho, N, N', x_1, x_2, p_k, B_1$ as introduced in Definition 4.

We begin by precomputing in increasing order the first elements of $\mathcal{E}''$ and stop when we get the first $r \in \mathcal{E}''$ such that $\ell(N_r^+) > 10^{15}$. We get a set E_2 with 1360 elements,

$$E_2 = \left\{ \frac{2^2-2}{\log 2}, \frac{3^2-3}{\log 3}, \frac{2^3-2^2}{\log 2}, \cdots \right\}.$$

We construct a table T, indexed from 1 to $\operatorname{card}(E_2) = 1360$. Let $r = (q^{j+1}-q^j)/\log q$ the i^{th} element of E_2. Then $T[i] = [q, j, p, l]$ where $l = \ell(N_r^+)$ and p is the largest prime p such that $p/\log p < r$. The superchampions following N_r^+ are obtained by multiplying it successively by the primes following p. Figure 3 gives the first values of $T[i]$. (In the Maple program the $T[i]$'s are the elements of the table `listesuperchE2`).

The superchampions that are not of the form N_r^+ for an $r \in E_2$ can easily be obtained from this table. For instance, the successive values of $\ell(N)$ between 368 and 626 are $368 + 53 = 421$, $421 + 59 = 480$, $480 + 61 = 541$ and $541 + 67 = 608$.

Two elements of $\mathcal{E}$ can be close. For instance, the smallest difference between two consecutive elements of $\mathcal{E}$ less than $8 \cdot 10^9$ is

$$\begin{aligned} &\frac{43083996283}{\log 43083996283} - \frac{144589^2 - 144589}{\log 144589} \\ &\quad = 1759505912.7146899772 - 1759505912.7146800938 = 0.0000098834 \end{aligned}$$

and thus, working with 20 decimal digits is enough to distinguish the elements of $\mathcal{E}$. For any n up to 10^{15}, Algorithm 2 below determines the superchampion $N = N_\rho$ as in Definition 4.

6. Benefits

6.1. Definition and properties.

Definition 5. *Let $\rho \in \mathcal{E}$ and $N = N_\rho$ (as defined in Definition 4). If M is a positive integer, from* (21)*, we have $\ell(M) - \rho \log M \geq \ell(N) - \rho \log N$. We call* benefit *of M the non-negative quantity*

$$\operatorname{ben}(M) = \ell(M) - \ell(N) - \rho \log \frac{M}{N} \cdot \tag{36}$$

Let $M = \prod_p p^{\beta_p}$ be the standard factorization of M. We define

$$\operatorname{ben}_p(M) = \ell(p^{\beta_p}) - \ell(p^{\alpha_p}) - \rho(\beta_p - \alpha_p) \log p \geq 0, \tag{37}$$

Algorithm 2 : computes $N = N_\rho$ for a given $n \le 10^{15}$.

Construct table T.
$i :=$ the largest index such that $T[i].\ell \le n$.
$\ell' := T[i+1].\ell$, $q' = T[i+1].q$, $j' = T[i+1].j$.
$\{r' = (q'^{j'} - q'^{(j'-1)})/\log q'$ is the smallest element in E_2 such that $\ell(N_{r'}) > n\}$
$t := \ell' - q'^{(j'-1)}(q'-1)$;
{This is the value $\ell(N)$ of the superchampion N preceding N_r^+}
if $t \le n$ **then**
  $\rho := r'$
else
  $n_0 := T[i].\ell + nextprime(T[i].p)$;
  while $n_0 \le n$ **do**
    $p := nextprime(p);\ n_0 := n_0 + p$
  end while
  $\rho := p/\log p$
end if

which implies

$$\text{ben}\,(M) = \sum_p \text{ben}_p\,(M). \tag{38}$$

Geometrically, if we represent $\log M$ in abscissa and $\ell(M)$ in ordinate, the straight line of slope ρ going through the point $(\log M, \ell(M))$ cuts the y axis at the ordinate $y_M = \ell(M) - \rho\log(M)$ and so, the benefit is the difference $y_M - y_N$ (see Fig. 4). Note that $\rho = \dfrac{\ell(N') - \ell(N)}{\log N' - \log N}$ with $N = N_\rho$ and $N' = N_\rho^+$.

Lemma 6. *Let $p \in \mathcal{P}$, $\alpha = \alpha_p = v_p(N)$ and γ a non-negative integer. Then,*

1. $\text{ben}\,(p^\gamma N) = \ell(p^{\gamma+\alpha}) - \ell(p^\alpha) - \rho\gamma\log p$ *is non-decreasing for $\gamma \ge 0$ and tends to infinity with γ.*
2. $\text{ben}\,(N/p^\gamma) = \rho\gamma\log p + \ell(p^{\alpha-\gamma}) - \ell(p^\alpha)$ *is non-decreasing for $0 \le \gamma \le \alpha$.*

Proof. 1. If $\gamma + \alpha \ge 1$, we have

$$\text{ben}\,(p^{\gamma+1}N) - \text{ben}\,(p^\gamma N) = \log p\left(p^\gamma \frac{p^{\alpha+1} - p^\alpha}{\log p} - \rho\right)$$

which is non-negative from (25) and tends to infinity with γ.

If $\alpha = \gamma = 0$, we have $\text{ben}\,(pN) - \text{ben}\,(N) = \log p(p/\log p - \rho)$ which is also non-negative from (25).

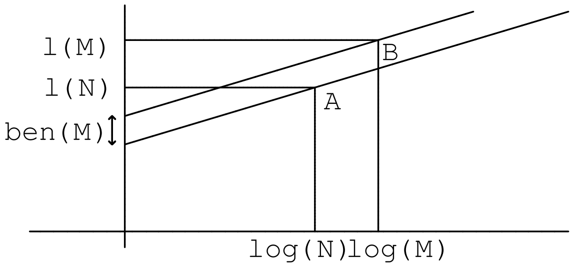

FIGURE 4. $A = (\log N, \ell(N))$ and $B = (\log M, \ell(M))$.

2. If $\alpha \geq 2$ and $0 \leq \gamma \leq \alpha - 2$, we have

$$\text{ben}\left(\frac{N}{p^{\gamma+1}}\right) - \text{ben}\left(\frac{N}{p^{\gamma}}\right) = \log p\left(\rho - \frac{1}{p^{\gamma}}\frac{p^{\alpha} - p^{\alpha-1}}{\log p}\right)$$

which is non-negative from (25).

If $\alpha \geq 1$ and $\gamma = \alpha - 1$,

$$\text{ben}\left(\frac{N}{p^{\gamma+1}}\right) - \text{ben}\left(\frac{N}{p^{\gamma}}\right) = \log p\left(\rho - \frac{p}{\log p}\right)$$

yields the same conclusion.

□

Lemma 7. *Let U/V be an irreducible fraction such that V divides N (as fixed in Definition 4) and $U = U_1U_2$, $V = V_1V_2$ with $(U_1, U_2) = (V_1, V_2) = 1$. Then we have*

1.

$$\ell\left(\frac{UN}{V}\right) - \ell(N) = \ell\left(\frac{U_1N}{V_1}\right) - \ell(N) + \ell\left(\frac{U_2N}{V_2}\right) - \ell(N). \tag{39}$$

2.

$$\text{ben}\left(\frac{UN}{V}\right) = \text{ben}\left(\frac{U_1N}{V_1}\right) + \text{ben}\left(\frac{U_2N}{V_2}\right). \tag{40}$$

Proof. Observing that a prime p divides at most one of the four numbers U_1, U_2, V_1, V_2 we get (39). By the additivity of the logarithm, (40) follows. □

The following proposition will be useful in the sequel.

Proposition 2. *Let M be a positive integer such that $\ell(M) \le n$ (thus, from* (4)*, $M \le g(n)$ holds). Then,*

$$\operatorname{ben} g(n) \le \operatorname{ben} M + \ell(g(n)) - \ell(M)$$

and

$$\operatorname{ben} g(n) \le \operatorname{ben} g(n) + n - \ell(g(n)) \le \operatorname{ben} M + n - \ell(M). \tag{41}$$

Proof. From (36), we have

$$\operatorname{ben} g(n) - \operatorname{ben} M = \ell(g(n)) - \ell(M) - \rho \log \frac{g(n)}{M} \le \ell(g(n)) - \ell(M)$$

which implies the first inequality while the second one follows from (3). □

We shall use Proposition 2 to determine an upper bound B such that

$$\operatorname{ben} g(n) \le \operatorname{ben} g(n) + n - \ell(g(n)) \le B. \tag{42}$$

It has been proved in [13] that $B \le x_1$ and

$$B = \mathcal{O}\left(\frac{x_1}{\log x_1}\right) = \mathcal{O}(\rho), \tag{43}$$

and, by the method of [23], it is possible to show that $B = o(\rho)$. The largest quotient $(\operatorname{ben} g(n) + n - \ell(g(n)))/\rho$ that we have found up to $n = 10^{12}$ is 1.60153 for $n = 45055780$.

6.2. The benefit of large primes.

Proposition 3. *Let N, B_1, x_1 and x_2 as in Definition 4. If M is an integer satisfying* $\operatorname{ben}(M) = \ell(M) - \ell(N) - \rho\log(M/N) < B_1$, *we have*

1.	*if*	$\sqrt{x_1} \le p$	*then*	$v_p(M) \le 1$
2.	*if*	$x_2 \le p < \sqrt{x_1}$	*then*	$v_p(M) \le 2.$

Proof.

1. Let us assume that the prime p satisfies $p \ge \sqrt{x_1}$ and divides M with exponent $k \ge 2$. With (35), we have $p > x_2$ and, from (30), the exponent α_p of p in $N = N_\rho$ is 0 or 1. If $\alpha_p = 1$, from (37) and (25) we have

$$\operatorname{ben}_p(M) = p^k - p - \rho(k-1)\log p = \log p \sum_{i=2}^{k} \Big(\frac{p^i - p^{i-1}}{\log p} - \rho\Big)$$

$$\ge \log p\Big(\frac{p^2 - p}{\log p} - \rho\Big) = p^2 - p - \rho \log p \tag{44}$$

while, if $\alpha_p = 0$,

$$\operatorname{ben}_p M = p^k - \rho\, k \log p = \log p \left(\frac{p}{\log p} - \rho + \sum_{i=2}^{k} \Big(\frac{p^i - p^{i-1}}{\log p} - \rho \Big) \right)$$
$$\geq \log p \Big(\frac{p^2 - p}{\log p} - \rho \Big) = p^2 - p - \rho \log p.$$

So, in both cases, (38) and (37) yield $\operatorname{ben} M \geq \operatorname{ben}_p M \geq f(p)$ with $f(t) = t^2 - t - \rho \log t$. We have $f'(t) = 2t - 1 - \rho/t$, $f''(t) > 0$ and, as $x_2 > 2$ holds, (28) implies

$$f'(x_2) = 2x_2 - 1 - \frac{x_2 - 1}{\log x_2} \geq x_2 \left(2 - \frac{1}{\log x_2} \right) - 1 \geq 2 \left(2 - \frac{1}{\log 2} \right) - 1 > 0$$

and $f(t)$ is increasing for $t \geq x_2$. Thus, since $p \geq \sqrt{x_1}$,

$$\operatorname{ben} M \geq f(p) \geq f(\sqrt{x_1}) = x_1 - \sqrt{x_1} - \frac{x_1}{\log x_1} \log \sqrt{x_1} = \frac{x_1}{2} - \sqrt{x_1} \geq B_1$$

in contradiction with our hypothesis, and 1. is proved.

2. Let p satisfy $2 < x_2 \leq p < \sqrt{x_1}$ so that, from (30), $\alpha_p = v_p(N) = 1$; let us assume that $k = v_p(M) \geq 3$; one would have as in (44)

$$\operatorname{ben} M \geq \log p \sum_{i=2}^{k} \left(\frac{p^i - p^{i-1}}{\log p} - \rho \right) \geq p^3 - p^2 - \rho \log p.$$

The function $f(t) = t^3 - t^2 - \rho \log t$ is easily shown to be increasing for $t \geq x_2$. From (28), $f(x_2) = x_2^3 - x_2^2 - (x_2^2 - x_2)$ and thus

$$\operatorname{ben} M \geq x_2^3 - x_2^2 - (x_2^2 - x_2) = x_2(x_2^2 - 2x_2 + 1) > x_2^2 - 2x_2.$$

From (34), it follows that $\operatorname{ben} M > B_1$ holds, in contradiction with our hypothesis, and 2. is proved.

□

7. Prefixes

7.1. Plain prefixes and suffixes.

Definition 6. *Let j be a positive integer.*

1. *For every positive integer M let us define the fraction*

$$\pi^{(j)}(M) = \prod_{p \leq p_j} p^{v_p(M) - v_p(N)} = \prod_{p \leq p_j} p^{v_p(M) - \alpha_p} \tag{45}$$

and call $\pi^{(j)}(M)$ the j-prefix of M.

2. *We note $\mathcal{T}_j$, and call it the set of* j-prefixes, *the set of fractions*

$$\mathcal{T}_j = \left\{ \delta = \prod_{p \leq p_j} p^{z_p}; \quad z_p \geq -\alpha_p \right\}. \tag{46}$$

3. *For $B' \geq 0$, we define*

$$\mathcal{T}_j(B') = \{\delta \in \mathcal{T}_j\ ;\ \mathrm{ben}\,(N\delta) \leq B'\}. \tag{47}$$

Definition 7. *Le M be a positive integer. Let us define*

$$\pi(M) = \prod_{p<\sqrt{x_1}} p^{v_p(M)-\alpha_p} = \pi^{(j_1)}(M) \tag{48}$$

where p_{j_1} is the largest prime less than $\sqrt{x_1}$, and $\xi(M) = M/(N\pi(M))$. Thus we have

$$M = N\,\pi(M)\;\xi(M). \tag{49}$$

$\pi(M)$ will be called the plain prefix *of M, and $\xi(M)$ the* suffix *of M.*

Let us show that, for each j such that $p_j < \sqrt{x_1}$, we have

$$\mathrm{ben}\,(N\pi^{(1)}(M)) \leq \ldots \leq \mathrm{ben}\,(N\pi^{(j)}(M)) \leq \ldots \leq \mathrm{ben}\,(N\pi(M)) \leq \mathrm{ben}\,M. \tag{50}$$

Indeed, (38) yields $\mathrm{ben}\,(N\pi^{(j)}) = \sum_{i\leq j} \mathrm{ben}_{\,p_i} M$ and $\mathrm{ben}\,M = \sum_p \mathrm{ben}_p\,M$, which implies (50), since, by (37), $\mathrm{ben}_p\,M$ is non-negative.

Definition 8. *From now on, we shall note*

$$\pi^{(j)} = \pi^{(j)}(g(n)), \qquad \pi = \pi(g(n)), \qquad \xi = \xi(g(n)) \tag{51}$$

so that $g(n) = N\pi\xi$ and our work is to compute π and ξ.

Note that π and ξ are coprime and (40) implies

$$\mathrm{ben}\,g(n) = \mathrm{ben}\,(N\pi\xi) = \mathrm{ben}\,(N\pi) + \mathrm{ben}\,(N\xi). \tag{52}$$

Lemma 8. *Let j be a positive integer and $\delta_1 < \delta_2$ be two elements of $\mathcal{T}_j$ satisfying*

$$\ell(\delta_2 N) \leq \ell(\delta_1 N). \tag{53}$$

Then, δ_1 is not the j-prefix of $g(n)$; in other words, $\pi^{(j)} \neq \delta_1$.

Proof. If $\delta_1 = \pi^{(j)}$, equation $g(n) = N\pi\xi$ may be written $g(n) = N\left(\delta_1 \frac{\pi}{\pi^{(j)}}\right)\xi$. Set $M = N\left(\delta_2 \frac{\pi}{\pi^{(j)}}\right)\xi = (\delta_2/\delta_1)g(n)$. From (39), (53) and (3), we get

$$\begin{aligned}\ell(M) &= \ell(\delta_2 N) + \ell\left(N\frac{\pi}{\pi^{(j)}}\right) + \ell(N\xi) - 2\ell(N)\\ &\leq \ell(\delta_1 N) + \ell\left(N\frac{\pi}{\pi^{(j)}}\right) + \ell(N\xi) - 2\ell(N) = \ell(g(n)) \leq n\end{aligned}$$

which, from (4), implies $M \leq g(n)$ and therefore $\delta_2 \leq \delta_1$, in contradiction with our hypothesis. Note that our hypothesis implies $\mathrm{ben}\,(\delta_2 N) < \mathrm{ben}\,(\delta_1 N)$. $\square$

7.2. Computing plain prefixes. Let us suppose that we know an upper bound B such that (42) holds. Then from (50) and (42), for every j such that $p_j < \sqrt{x_1}$, $\operatorname{ben}(N\pi^{(j)}) \leq B$ holds. Let p_{j_1} be the largest prime less than $\sqrt{x_1}$. Then $\pi = \pi^{(j_1)}(g(n))$ is an element of $\mathcal{T}_{j_1}(B)$.

But, we are faced to 2 problems: first, for the moment, we do not know B. Secondly, for a given value B', the sets $\mathcal{T}_j(B')$ are too large to be computed efficiently.

What we can do is the following. Let $B' < B_1$. We shall construct two non-decreasing sequences of sets $\mathcal{U}_j = \mathcal{U}_j(B')$ and $\mathcal{D}_j = \mathcal{D}_j(B')$ with $\mathcal{D}_j \subset \mathcal{U}_j \subset \mathcal{T}_j(B')$ satisfying the following property: *$\mathcal{D}_j$ contains the j-prefix $\pi^{(j)}$ of $g(n)$, provided that* $\operatorname{ben} g(n) \leq B'$ *holds.*

These sequences are defined by the following induction rule. The only element of $\mathcal{T}_0$ is 1. We set $\mathcal{U}_0 = \mathcal{D}_0 = \{1\}$. And, for $j \geq 1$,

- We define $\mathcal{U}_j = \left\{\delta p_j^{\gamma} \mid \delta \in \mathcal{D}_{j-1},\ \gamma \geq -\alpha_{p_j} \text{ and } \operatorname{ben}(N\delta p_j^{\gamma}) \leq B'\right\}$.
- By Lemma 8, if $\delta_1 \in \mathcal{U}_j$ and if there is a δ_2 in $\mathcal{U}_j$ such that $\delta_1 < \delta_2$ and $\ell(N\delta_1) \geq \ell(N\delta_2)$, then δ_1 is not the j-prefix of $g(n)$. The set $\mathcal{D}_j$ is $\mathcal{U}_j$ from which these δ_1's are removed. In other words, $\mathcal{D}_j$ will be the pruned set of $\mathcal{U}_j$ (see Section 2.2).

For each δ in $\mathcal{D}_{j-1}$, δp_j^{γ} belongs to $\mathcal{U}_j$ if $\gamma \geq -\alpha_{p_j}$ and $\operatorname{ben}(N\delta p_j^{\gamma}) \leq B'$ which, according to (40), can be rewritten as

$$\operatorname{ben}(Np_j^{\gamma}) \leq B' - \operatorname{ben}(N\delta). \tag{54}$$

It results from Lemma 6 that $\operatorname{ben}(Np_j^{\gamma})$ is non-increasing for $-\alpha_{p_j} \leq \gamma \leq 0$, non-decreasing for $\gamma \geq 0$, vanishes for $\gamma = 0$ and tends to infinity with γ. Therefore the solutions in γ of (54) form a finite interval containing 0.

Thanks to (50), by induction on j, it can be seen that if $\operatorname{ben} g(n) \leq B'$, the j-prefix $\pi^{(j)}$ of $g(n)$ belongs to $\mathcal{U}_j$ and also to $\mathcal{D}_j$, by Lemma 8.

We set $\mathcal{D}(B') = \mathcal{D}_{j_1}(B')$ and since $\pi = \pi_{j_1}$, $\mathcal{D}(B')$ contains the plain prefix π of $g(n)$, provided that $\operatorname{ben} g(n) \leq B'$ holds.

This construction solves our second problem: at each step of the induction, the pruning algorithm makes $\mathcal{D}_j(B')$ smaller than $\mathcal{U}_j(B')$, and as we progress, $\mathcal{D}_j(B')$ becomes much smaller than $\mathcal{T}_j(B')$.

7.3. Computing B, an upper bound for the benefit. It remains to find an upper bound B such that (42) holds. The key is Proposition 2. Every M such that $\ell(M) \leq n$ gives an upper bound for $\operatorname{ben} g(n) + n - \ell(g(n))$:

$$\operatorname{ben} g(n) \leq \operatorname{ben} g(n) + n - \ell(g(n)) \leq \operatorname{ben} M + n - \ell(M).$$

We choose some B', a provisional value of B satisfying[1] $B' < B_1$. Then we compute the set $\mathcal{D} = \mathcal{D}(B')$, and by using the prefixes belonging to this set we shall construct an integer M to which we apply Proposition 2.

Let us recall that p_k denotes the greatest prime dividing N. To an element $\delta \in \mathcal{D}(B')$ and to an integer ω, we associate

$$\delta_\omega = \begin{cases} \delta p_{k+1}p_{k+2}\dots p_{k+\omega} & \text{if } \omega > 0 \\ \delta & \text{if } \omega = 0 \\ \delta/(p_k p_{k-1}\dots p_{k+\omega+1}) & \text{if } \omega < 0 \text{ and } p_{k+\omega+1} \geq \sqrt{x_1}. \end{cases}$$

From the definition of prefixes, the prime factors of both the numerator and the denominator of $\delta \in \mathcal{D}(B')$ are smaller than $\sqrt{x_1}$, and thus smaller than the primes dividing the numerator or the denominator of δ_ω/δ.

First, to each $\delta \in \mathcal{D}$, let $\omega = \omega(\delta)$ be the greatest integer such that $\ell(N\delta_\omega) \leq n$ (if there is no such $\omega(\delta)$, we just forget this δ). We call $\delta^{(0)}$ an element of $\mathcal{D}$ which minimizes $\operatorname{ben}\big(N\delta_\omega^{(0)}\big) + n - \ell\big(N\delta_\omega^{(0)}\big)$ and set $M = N\delta_\omega^{(0)}$. From the construction of M, we have $\ell(M) \leq n$. By Proposition 2, inequality (42) is satisfied with $B = \operatorname{ben} M + n - \ell(M)$.

If $B \leq B'$, we stop and keep B; otherwise we start again with B instead of B' to eventually obtain a better bound.

For $n = 1000064448$, the value of ρ defined by (32) is equal to $\rho \approx 12661.7$; the table below displays some values of B'/ρ and the corresponding values of $\operatorname{Card}(\mathcal{D}(B'))$ and B/ρ given by the above method.

B'/ρ	0	0.2	0.4	0.6	0.7	0.8	0.9	1	1.1
$\lvert\mathcal{D}(B')\rvert$	1	11	34	76	109	139	165	194	224
B/ρ	7.5	1.15	1.13	1.104	1.098	1.082	1.074	1.055	1.055

In this example, if our first choice for B' is 0.6ρ, we find $B = 1.104\rho$. Starting again the algorithm with $B' = 1.104\rho$, we get the slightly better value $B = 1.055\rho$.

The value of B given by this method is reasonable and less than 10% more than the best possible one: for $n = 1000366$, we find $B \approx 436.04$ while $\operatorname{ben}\big(g(n) + n - \ell(g(n)\big) \approx 406.1$; for $n = 1000064448$, these two numbers are 13361.6 and 13285.7.

7.4. How many plain prefixes are there? Let us denote by $B = B(n)$ the upper bound satisfying (42) as computed in Section 7.3. Let us call $\widetilde{n}$ the integer in the range $\ell(N)..\ell(N') - 1$ such that $B(\widetilde{n})$ is maximal.

Let us denote by $\nu = \nu(n)$ the number of possible plain prefixes as obtained by the algorithm described in Section 7.2. Actually, this number

[1] In view of (43) and after some experiments, our choice is $B' = \rho$ for $2485 \leq n \leq 10^{10}$ while, for greater n's, we take $B' = \rho/2$, and for smaller n's, $B' = B_1 - \varepsilon$ where ε is some very small positive number.

n	$\widetilde{n}$	$B(\widetilde{n})/\rho$	$\nu(\widetilde{n}) = \#$ of plain prefixes	exponent = $\log \nu(\widetilde{n})/\log n$
10^3	$10^3 - 11$	0.9289	14	0.3820
10^4	$10^4 - 10$	0.8453	19	0.3197
10^5	$10^5 - 123$	0.8095	22	0.2685
10^6	$10^6 + 366$	0.9186	51	0.2846
10^7	$10^7 - 1269$	0.7636	59	0.2530
10^8	$10^8 + 639$	1.180	85	0.2412
10^9	$10^9 + 64448$	1.055	212	0.2585
10^{10}	$10^{10} + 88835$	0.6884	252	0.2401
10^{11}	$10^{11} + 1007566$	0.9278	657	0.2561
10^{12}	$10^{12} + 2043578$	1.118	2873	0.2882
10^{13}	$10^{13} + 5276948$	0.8331	3805	0.2754
10^{14}	$10^{14} + 17212588$	0.6669	7048	0.2749
10^{15}	$10^{15} - 44672895$	0.6433	15148	0.2787
10^{16}	$10^{16} - 48912919$	0.5077	25977	0.2759
10^{17}	$10^{17} - 426915678$	0.6001	72341	0.2858
10^{18}	$10^{18} + 385838833$	0.3027	144807	0.2867
10^{19}	$10^{19} - 9639993444$	0.2963	170151	0.2753
10^{20}	$10^{20} + 12041967315$	0.3218	412151	0.2808

FIGURE 5. The number of plain prefixes.

ν depends on $B = B(n)$ and we may think that it is a non-decreasing function on B so that the maximal number of prefixes used to compute $g(m)$ for $\ell(N) \le m < \ell(N')$ should be equal to $\nu(\widetilde{n})$.

For the powers of 10, the table of FIG. 5 displays n, $\widetilde{n}$, the quotient of the maximal benefit $B(\widetilde{n})$ by ρ, the maximal number of plain prefixes $\nu(\widetilde{n})$ and the exponent $\log \nu(\widetilde{n})/\log n$. Note that replacing $\log n$ by $\log \widetilde{n}$ will not change very much this exponent, since with the notation of Definition 4, we have $|\widetilde{n} - n| \le \ell(N') - \ell(N) \le p_{k+1} \lesssim \sqrt{n \log n}$.

The behaviour of $\nu(\widetilde{n})$ looks regular and allows to think that

$$\nu(\widetilde{n}) = O(n^{0.3}). \tag{55}$$

7.5. For ben (M) small, prime factors of $\xi(M)$ are large. If the number B computed as explained in Section 7.3 is greater than B_1 our algorithm fails. Fortunately, we have not yet found any $n \ge 166$ for which that bad event occurs.

Proposition 4. *If B is computed as explained in Section 7.3 (so that* (42) *holds) and satisfies $B < B_1$ (where B_1 is defined in* (34)*) then, in view of* (35)*, there exists a unique real number t_1 such that*

$$2 < x_2 < \sqrt{x_1} < \rho = \frac{x_1}{\log x_1} < t_1 < x_1 \tag{56}$$

and

$$\rho \log t_1 - t_1 = B. \tag{57}$$

Further, if $\operatorname{ben} M \le B$*, we have*

1. *If $x_2 \le p < t_1$ then $v_p(M) \ge 1 = v_p(N)$.*
2. *If $x_2 \le p < \sqrt{x_1}$ then $v_p(M) \in \{1, 2\}$ and $v_p(N) = 1$.*
3. *If $\sqrt{x_1} \le p < t_1$ then $v_p(M) = v_p(N) = 1$.*
4. *If $t_1 \le p < x_1$ then $v_p(M) \in \{0, 1\}$ and $v_p(N) = 1$.*
5. *If $x_1 \le p$ then $v_p(M) \in \{0, 1\}$ and $v_p(N) = 0$.*

Proof. The function $f(t) = \rho \log t - t$ is increasing on $[x_2, \rho]$ and decreasing on $[\rho, x_1]$. From (28) and (34) we have

$$f(\rho) > f(x_2) = \frac{x_2^2 - x_2}{\log x_2} \log x_2 - x_2 = x_2^2 - 2x_2 \ge B_1 > B > 0 = f(x_1)$$

which gives the existence and unicity of t_1, which belongs to $(\rho,\, x_1)$. Now we prove points 1,2,3,4,5.

Let p be a prime number satisfying $x_2 \le p < t_1$. If p does not divide M, from (38) and (37) we have

$$\operatorname{ben} M \ge \operatorname{ben}_p M = \rho \log p - p = f(p) > f(t_1) = B.$$

Since $\operatorname{ben} M \le B$ is supposed to hold, there is a contradiction and 1 is proved.

Since we have assumed that $B < B_1$ holds, Proposition 3 may be applied. Point 2. follows from point 1. and from item 2. of Proposition 3, while point 3. follows from point 1. and from item 1. of Proposition 3. Finally, points 4. and 5. are implied by item 1. of Proposition 3. □

Corollary 1. *Let us assume that B is such that* (42) *and $B < B_1$ hold. Then the suffix $\xi = \xi(g(n))$ from Definition 8 can be written as*

$$\xi = \xi(g(n)) = \frac{p_{i_1} p_{i_2} \cdots p_{i_u}}{p_{j_1} p_{j_2} \cdots p_{j_v}} \qquad u \ge 0,\ v \ge 0 \tag{58}$$

where (we recall that p_k is the largest prime factor of N)

$$2 < x_2 < \sqrt{x_1} < \rho < t_1 \le p_{j_1} < p_{j_2} \cdots < p_{j_v} \le p_k < p_{i_1} < \cdots < p_{i_u}. \tag{59}$$

7.6. Normalized prefix of $g(n)$.

Definition 9. *Let u and v be as defined in* (58) *and $\omega = u - v$. We define the* normalized suffix σ *of $g(n)$ by*

1. *If $\omega \geq 0$*

$$\sigma = \frac{p_{i_1} \cdots p_{i_u}}{p_{j_1} \cdots p_{j_v} p_{k+1} \cdots p_{k+\omega}} = \frac{\xi}{p_{k+1} \cdots p_{k+\omega}}.$$

2. *If $\omega < 0$, we set $\omega' = -\omega$ and*

$$\sigma = \frac{p_{i_1} \cdots p_{i_u} p_k \cdots p_{k-\omega'+1}}{p_{j_1} \cdots p_{j_v}} = \xi p_k \cdots p_{k-\omega'+1}.$$

The normalized prefix Π *of $g(n)$ is defined by*

$$\Pi = \frac{g(n)}{N\sigma} = \begin{cases} \pi p_{k+1} p_{k+2} \cdots p_{k+\omega} & \text{if } \omega \geq 0 \\ \dfrac{\pi}{p_k \cdots p_{k-\omega'+1}} & \text{if } \omega < 0. \end{cases} \tag{60}$$

Proposition 5. *Let σ be the normalized suffix of $g(n)$. Then*

$$\sigma = \frac{Q_1 Q_2 \ldots Q_s}{q_1 q_2 \cdots q_s}$$

where s is a non-negative integer with

1. *If $\omega \geq 0$ then $u \leq s \leq v$ and*

$$\operatorname{ben}(N\Pi) = \operatorname{ben}(N\pi) + \sum_{i=1}^{\omega} \operatorname{ben}(N p_{k+i}) = \operatorname{ben}(N\pi) + \sum_{i=1}^{\omega} (p_{k+i} - \rho \log p_{k+i}), \tag{61}$$

$$\ell(\sigma) = \sum_{i=1}^{s} (Q_i - q_i) = p_{i_1} + \ldots + p_{i_u} - (p_{j_1} + \ldots + p_{j_v}) - (p_{k+1} + \ldots + p_{k+\omega}) \geq 0. \tag{62}$$

2. *If $\omega < 0$ then $v \leq s \leq u$ and, with $\omega' = -\omega = v - u$, we have*

$$\begin{aligned} \operatorname{ben}(N\Pi) &= \operatorname{ben}(N\pi) + \sum_{i=0}^{\omega'-1} \operatorname{ben}\left(\frac{N}{p_{k-i}}\right) \\ &= \operatorname{ben}(N\pi) + \sum_{i=0}^{\omega'-1} (\rho \log p_{k-i} - p_{k-i}) \end{aligned} \tag{63}$$

$$\ell(\sigma) = \sum_{i=1}^{s} (Q_i - q_i) = p_{i_1} + \ldots + p_{i_u} - (p_{j_1} + \ldots + p_{j_v}) + (p_k + \ldots + p_{k-\omega'+1}) \geq 0. \tag{64}$$

In both cases we have also

$$\sqrt{x_1} < \rho < t_1 < q_1 < \cdots < q_s \leq p_{k+\omega} < Q_1 < \cdots < Q_s. \tag{65}$$

Proof. If $u \geq v$ then $\omega = u - v \geq 0$,

$$\sigma = \frac{p_{i_1} \dots p_{i_u}}{p_{j_1} \dots p_{j_v} p_{k+1} \dots p_{k+\omega}} = \frac{\xi}{p_{k+1} \dots p_{k+\omega}}. \tag{66}$$

Since the prime factors $p_{i_1} \dots p_{i_u}$ of the numerator are distinct of the prime factors $p_{j_1} \dots p_{j_v}$ of the denominator, σ can be written after simplification

$$\sigma = \frac{Q_1 Q_2 \dots Q_s}{q_1 q_2 \dots q_s} \tag{67}$$

where $v \leq s \leq u$ and, from (59), we have

$$\sqrt{x_1} < \rho < t_1 < q_1 < q_2 < \dots < q_s \leq p_{k+\omega} < Q_1 < Q_2 < \dots < Q_s$$

which is (65). From (40) we get (61) while (62) follows from (66) and (67).

Similarly, if $u < v$ holds, $\omega' = v - u > 0$. So, $\omega' \leq v$, and from (59), $p_{k-\omega'+1} \geq p_{k-v+1} \geq p_{j_1} > t_1$; (66) and (67) become

$$\sigma = \frac{p_{i_1} \dots p_{i_u} p_k \dots p_{k-\omega'+1}}{p_{j_1} \dots p_{j_v}} = \frac{Q_1 \dots Q_s}{q_1 \dots q_s} \tag{68}$$

where $u \leq s \leq v$ and we have

$$\sqrt{x_1} < \rho < t_1 < q_1 < \dots < q_s \leq p_{k-\omega'} = p_{k+\omega} < Q_1 < \dots < Q_s \tag{69}$$

which is again (65).

By definition, any prime factor of π is smaller than $\sqrt{x_1}$. Therefore, by (69), $p_{k-\omega'+1}$ is greater than any prime factor of π, (40) can be applied and (61) becomes (63) while (62) becomes (64). □

The value of the parameter ω can be computed from the following proposition. It is convenient to set $S_\omega = \sum_{i=1}^{\omega} p_{k+i}$ (for $\omega \geq 0$) and $S_\omega = -\sum_{i=0}^{-\omega-1} p_{k-i}$ (for $\omega < 0$). In both cases, from (39), we have

$$S_\omega = \ell(N\Pi) - \ell(N\pi). \tag{70}$$

Proposition 6. *The relative integer ω which determines the normalized prefix Π of $g(n)$ (cf. (60)) satisfies the following inequalities:*

$$n - \ell(N\pi) - \frac{B}{1 - \rho/t_1} \leq n - \ell(N\pi) - \frac{B - \operatorname{ben}(N\Pi)}{1 - \rho/t_1} \leq S_\omega \leq n - \ell(N\pi) \tag{71}$$

where π is the prefix of $g(n)$ and B and t_1 satisfy (42) and (57).

Proof. Let us prove Proposition 6 for $\omega \geq 0$; the case $\omega < 0$ is similar. From (67), (65) and (62), Lemma 1 1. yields

$$1 \leq \sigma \leq \exp\left(\frac{\ell(\sigma)}{t_1}\right). \tag{72}$$

From (58) and (62), we have

$$\ell(N\xi) - \ell(N) = p_{i_1} + \dots + p_{i_u} - (p_{j_1} + \dots + p_{j_v}) = \ell(\sigma) + S_\omega. \tag{73}$$

So, we get successively

$$\begin{aligned}\operatorname{ben}(N\xi) &= \ell(N\xi) - \ell(N) - \rho\log\xi \quad \text{by (36)}\\ &= \ell(\sigma) + \sum_{i=1}^{\omega}(p_{k+i} - \rho\log p_{k+i}) - \rho\log\sigma \quad \text{by (66)}\\ &\geq \ell(\sigma) + \sum_{i=1}^{\omega}(p_{k+i} - \rho\log p_{k+i}) - \frac{\rho\ell(\sigma)}{t_1} \quad \text{by (72)}\\ &= \ell(\sigma)\left(1 - \frac{\rho}{t_1}\right) + \operatorname{ben}(N\Pi) - \operatorname{ben}(N\pi) \quad \text{by (61).}\end{aligned}$$

From (62), we have $\ell(\sigma) \geq 0$. Since, from (65), $\rho < t_1$ holds, the above result together with (52), (42) and (3) implies that

(74)
$$\begin{aligned}0 \leq \ell(\sigma) \leq \frac{\operatorname{ben}(N\xi) - \operatorname{ben}(N\Pi) + \operatorname{ben}(N\pi)}{1-\rho/t_1} &= \frac{\operatorname{ben} g(n) - \operatorname{ben}(N\Pi)}{1-\rho/t_1}\\ \leq \frac{B - \operatorname{ben}(N\Pi) - n + \ell(g(n))}{1-\rho/t_1} &\leq \frac{B - \operatorname{ben}(N\Pi)}{1-\rho/t_1} - (n - \ell(g(n))).\end{aligned}$$

Now, from (39), and (73), we get

$$\ell(g(n)) = \ell(N\pi\xi) = \ell(N\pi) + \ell(N\xi) - \ell(N) = \ell(N\pi) + \ell(\sigma) + S_\omega. \tag{75}$$

Further, since

$$n - \ell(N\pi) = \ell(g(n)) - \ell(N\pi) + n - \ell(g(n)) = \ell(\sigma) + S_\omega + n - \ell(g(n)), \tag{76}$$

we get from (74) and (3)

$$n - \ell(N\pi) - \frac{B - \operatorname{ben}(N\Pi)}{1-\rho/t_1} \leq S_\omega \leq n - \ell(N\pi) \tag{77}$$

and (71) follows, since $\operatorname{ben}(N\Pi) \geq 0$. Note that (77) implies

$$\operatorname{ben}(N\Pi) \leq B. \tag{78}$$

□

7.7. Computing possible normalized prefixes. In Section 7.2, we have computed B such that (42) holds and a set $\mathcal{D} = \mathcal{D}(B)$ containing the plain prefix π of $g(n)$. By construction, we know that any prime factor of $\pi \in \mathcal{D}$ is smaller than $\sqrt{x_1}$ and thus, from (56), smaller than t_1.

Definition 10. *We call* possible normalized prefix *a positive rational number $\widehat{\Pi} = \widehat{\Pi}(\widehat{\pi}, \omega)$ of the form $\widehat{\Pi} = \widehat{\pi} p_{k+1} \dots p_{k+\omega}$ (with $\omega \geq 0$) or $\widehat{\Pi} = \widehat{\pi}/(p_k \dots p_{k+\omega+1})$ (with $\omega < 0$), where $\widehat{\pi} \in \mathcal{D}(B)$ is a plain prefix, and satisfying*

$$p_{k+\omega+1} \geq t_1 \tag{79}$$

and

$$(80)\quad n-\ell(N\widehat{\pi})-\frac{B}{1-\rho/t_1}\le n-\ell(N\widehat{\pi})-\frac{B-\operatorname{ben}(N\widehat{\Pi})}{1-\rho/t_1}\le S_\omega\le n-\ell(N\widehat{\pi})$$

with $S_\omega=\sum_{i=1}^{\omega}p_{k+i}$ *(if* $\omega\ge 0$*) and* $S_\omega=-\sum_{i=0}^{-\omega-1}p_{k-i}$ *(if* $\omega<0$*).*

Let us denote by $\mathcal{N}$ the set of possible normalized prefixes; $\mathcal{N}$ has been defined in such a way that the normalized prefix Π of $g(n)$ belongs to $\mathcal{N}$. Indeed, from (60), Π has the suitable form, the plain prefix π of $g(n)$ belongs to $\mathcal{D}(B)$, (80) is satisfied by Proposition 6 and (79) by (65).

Let us observe that, if ω increases by 1, by (65), S_ω increases by at least t_1. In practice, $1-\rho/t_1$ is close to 1 and B is much smaller than t_1 so that for most of the $\widehat{\pi}$'s there is no solution to (80) and there are few possible normalized prefixes. For n in the range $[998001, 1000000]$, the number of possible normalized prefixes is 1 (resp. 2 or 3) for 1439 values (resp. 547 or 94). For instance, for $n=998555$, the three possible normalized prefixes are $1, 43/41, 11/10$.

Finally, for a reason given in the next section, for every $\widehat{\Pi}\in\mathcal{N}$, we check that the following inequality holds:

$$(81)\qquad p_{k+\omega+1}-(n-\ell(N\widehat{\Pi}))\ge\sqrt{x_1}.$$

This inequality seems reasonable, since, from (79), we have $p_{k+\omega+1}\ge t_1$ with t_1 close to x_1, and, from (80), $n-\ell(N\widehat{\Pi})=n-\ell(N\widehat{\pi})-S_\omega\le B/(1-\rho/t_1)$ which is much smaller than x_1. We have not found any counterexample to (81).

7.8. The heart of the algorithm. We have now a list $\mathcal{N}$ of possible normalized prefixes containing the normalized prefix Π of $g(n)$. For $\widehat{\Pi}=\widehat{\Pi}(\widehat{\pi},\omega)\in\mathcal{N}$ let us introduce

$$(82)\qquad g(\widehat{\Pi},n)=N\widehat{\Pi}G(p_{k+\omega},n-\ell(N\widehat{\Pi}))=N\widehat{\Pi}\frac{Q_1Q_2\dots Q_s}{q_1q_2\dots q_s}$$

where $G(p_{k+\omega},n-\ell(N\widehat{\Pi}))=\dfrac{Q_1Q_2\dots Q_s}{q_1q_2\dots q_s}$ is defined by (12). We shall use the following proposition to compute $g(n)$.

Proposition 7. *The following formula gives the value of* $g(n)$*:*

$$(83)\qquad g(n)=\max_{\widehat{\Pi}\in\mathcal{N}}g(\widehat{\Pi},n)=\max_{\widehat{\Pi}\in\mathcal{N}}N\widehat{\Pi}G(p_{k+\omega},n-\ell(N\widehat{\Pi})).$$

Proof. Note that (13) and (14) imply either $s=0$ or the smallest prime factor q_s of $G(p_{k+\omega},n-\ell(N\widehat{\Pi}))$ satisfies $p_{k+\omega+1}-q_s\le n-\ell(N\widehat{\Pi})$ which, from (81), implies $q_s\ge\sqrt{x_1}$ and thus, the prime factors of π and those of

$G(p_{k+\omega}, n-\ell(N\widehat{\Pi}))$ are distinct. Therefore, for any $\widehat{\Pi} = \widehat{\Pi}(\widehat{\pi}, \omega) \in \mathcal{N}$ with $\omega \geq 0$, we get from (82), (39) and (15)

$$\begin{aligned}
\ell(g(\widehat{\Pi}, n)) &= \ell(N\widehat{\pi}) + \ell\left(N\frac{p_{k+1}\dots p_{k+\omega}Q_1\dots Q_s}{q_1\dots q_s}\right) - \ell(N) \\
&= \ell(N\widehat{\pi}) + \sum_{i=1}^{\omega} p_{k+i} + \sum_{i=1}^{s}(Q_i - q_i) \\
&= \ell(N\widehat{\Pi}) + \ell(G(p_{k+\omega}, n-\ell(N\widehat{\Pi}))) \\
&\leq \ell(N\widehat{\Pi}) + n - \ell(N\widehat{\Pi}) = n.
\end{aligned}$$

Inequality $\ell(g(\widehat{\Pi}, n)) \leq n$ can be proved similarly in the case $\omega < 0$.

Since $\ell(g(\widehat{\Pi}, n)) \leq n$ holds, (4) implies for all $\widehat{\Pi} \in \mathcal{N}$

$$g(\widehat{\Pi}, n) \leq g(n). \tag{84}$$

From (60), we get $g(n) = N\Pi\sigma$ where Π is the normalized prefix of $g(n)$. Now, if $\omega \geq 0$, from (62), (75), (60) and (3), we have

$$\begin{aligned}
\ell(\sigma) = \sum_{i=1}^{s}(Q_i - q_i) &= \ell(g(n)) - \ell(N\pi) - \sum_{i=1}^{\omega} p_{k+i} \\
&= \ell(g(n)) - \ell(N\Pi) \leq n - \ell(N\Pi)
\end{aligned} \tag{85}$$

($\ell(\sigma) \leq n - \ell(N\Pi)$ still holds for $\omega < 0$). Therefore, in view of (65) and of Definition (12) of function G, we have

$$g(n) = N\Pi\sigma \leq N\Pi G(p_{k+\omega}, n-\ell(N\Pi)) = g(\Pi, n). \tag{86}$$

Since $\Pi \in \mathcal{N}$, (86) and (84) prove (83). $\square$

7.9. The fight of normalized prefixes. Let $\widehat{\Pi}_1$ and $\widehat{\Pi}_2$ be two normalized prefixes. By using Inequalities (90) below, it is sometimes possible to eliminate $\widehat{\Pi}_1$ or $\widehat{\Pi}_2$.

Indeed, from (90), we deduce a lower and an upper bound for $g(\widehat{\Pi}, n)$ (defined in (82)):

$$g'(\widehat{\Pi}, n) \leq g(\widehat{\Pi}, n) \leq g''(\widehat{\Pi}, n).$$

If, for instance, $g''(\widehat{\Pi}_1, n) < g'(\widehat{\Pi}_2, n)$ holds, then clearly $\widehat{\Pi}_1$ cannot compete in (83) to be the maximum.

By this simple trick, it is possible to shorten the list $\mathcal{N}$ of normalized prefixes. For instance, for $n = 10^{15}$, the number of normalized prefixes is reduced from 9 to 1, while, for $n = 10^{15} + 123850000$, it is reduced from 37 to 2.

8. A first way to compute $G(p_k, m)$

8.1. Function G. In this section, we study the function G introduced in (12). First, for $k \geq 3$ and $0 \leq m \leq p_{k+1} - 3$, we consider the set

$$\mathcal{G}(p_k, m) = \left\{ F = \frac{Q_1 Q_2 \dots Q_s}{q_1 q_2 \cdots q_s}; \quad \ell(F) = \sum_{i=1}^{s} (Q_i - q_i) \leq m, \quad s \geq 0 \right\} \tag{87}$$

where the primes $Q_1, Q_2, \dots, Q_s, q_1, q_2, \dots, q_s$ satisfy (13).

The parameter $s = s(F)$ in (87) is called the number of factors of the fraction F. If $s = 0$, we set $F = 1$ and $\ell(F) = 0$ so that $\mathcal{G}(p_k, m)$ contains 1 and is never empty. The definition (12) can be rewritten as

$$G(p_k, m) = \max_{F \in \mathcal{G}(p_k, m)} F. \tag{88}$$

Obviously, $G(p_k, m)$ is non-decreasing on m and $G(p_k, 2m+1) = G(p_k, 2m)$. Note that the maximum in (88) is unique (from the unicity of the standard factorization into primes). It follows from (13) that, if $0 \leq m < p_{k+1} - p_k$, the set $\mathcal{G}(p_k, m)$ contains only 1, and therefore,

$$0 \leq m < p_{k+1} - p_k \quad \Longrightarrow \quad G(p_k, m) = 1. \tag{89}$$

Proposition 8.

1. *Let q be the smallest prime satisfying $q \geq p_{k+1} - m$. The following inequality holds*

$$\frac{p_{k+1}}{q} \leq G(p_k, m) \leq \frac{p_{k+1}}{p_{k+1} - m}. \tag{90}$$

   Note that if $q = p_{k+1} - m$ is prime, then (90) *yields the exact value of $G(p_k, m)$.*

2. *Now, let $F = \dfrac{Q_1 Q_2 \dots Q_s}{q_1 q_2 \cdots q_s}$ be any element of $\mathcal{G}(p_k, m)$; we have*

$$G(p_k, m) \geq F \geq 1 + \frac{\ell(F)}{p_k}. \tag{91}$$

Proof. The lower bound in (90) is obvious. Let us prove the upper bound. If $0 \leq m < p_{k+1} - p_k$, the upper bound of (90) follows by (89). If $m \geq p_{k+1} - p_k$, $\dfrac{p_{k+1}}{p_k} \in \mathcal{G}(p_k, m)$ and thus $G(p_k, m) \geq \dfrac{p_{k+1}}{p_k} > 1$. Moreover, with the notation (87), if $G(p_k, m) = F = \dfrac{Q_1 Q_2 \dots Q_s}{q_1 q_2 \cdots q_s}$, we have $s \geq 1$ and Lemma 1 2. implies

$$G(p_k, m) \leq \frac{Q_s}{Q_s - \ell(F)} \leq \frac{Q_s}{Q_s - m} \leq \frac{p_{k+1}}{p_{k+1} - m} \tag{92}$$

where the last inequality follows from (13) and the decrease of $t \mapsto t/(t-m)$.

Let us now prove (91). This inequality holds if $\ell(F) = 0$ (i.e., $F = 1$ and $s = 0$). If $s > 0$, from (13), we get

$$\frac{Q_i}{q_i} = 1 + \frac{Q_i - q_i}{q_i} \geq 1 + \frac{Q_i - q_i}{p_k}, \qquad i = 1, 2, \ldots, s$$

and

$$F = \prod_{i=1}^{s} \frac{Q_i}{q_i} \geq \prod_{i=1}^{s} \left(1 + \frac{Q_i - q_i}{p_k}\right) \geq 1 + \frac{\sum_{i=1}^{s}(Q_i - q_i)}{p_k} = 1 + \frac{\ell(F)}{p_k}.$$

□

8.2. Function *H*. Let $M \leq p_{k+1} - 3$; we want to calculate $G(p_k, m)$ for $0 \leq m \leq M$. Let us introduce a family of consecutive primes $P_0 < P_1 < \ldots < P_K = p_k < P_{K+1} < \ldots < P_R < P_{R+1}$ (so that $P_i = p_{k+i-K}$ for $0 \leq i \leq R+1$) with the properties

$$P_{R+1} - P_K > M, \quad R \geq K+1, \quad P_{K+1} - P_0 > M, \quad P_1 \geq 3. \tag{93}$$

It follows from (87) and (13) that the prime factors $Q_1, \ldots, Q_s, q_1, \ldots, q_s$ of any element of $\mathcal{G}(p_k, m) = \mathcal{G}(P_K, m)$ should satisfy

$$P_1 \leq q_s < \ldots < q_1 \leq P_K = p_k < P_{K+1} \leq Q_1 < \ldots < Q_s \leq P_R. \tag{94}$$

Of course, in (93) we may choose P_R (resp. P_1) as small (resp. large) as possible, but it is not an obligation.

Let us denote by $Q'_1, Q'_2, \ldots, Q'_{R-K-s}$ the primes among $P_{K+1}, \ldots, P_R$ which are different from $Q_1, \ldots, Q_s$; we have

$$Q'_1 + Q'_2 + \ldots + Q'_{R-K-s} = P_{K+1} + \ldots + P_R - (Q_1 + \ldots + Q_s) \tag{95}$$

and (88) becomes

$$G(P_K, m) = \max \frac{P_{K+1}P_{K+2}\ldots P_R}{Q'_1 \ldots Q'_{R-K-s} q_1 \ldots q_s} = \frac{P_{K+1}P_{K+2}\ldots P_R}{\min(q'_1 \ldots q'_{R-K})} \tag{96}$$

where the minimum is taken over all the subsets $\{q'_1, q'_2, \ldots, q'_{R-K}\}$ of $R-K$ elements of $\{P_1, \ldots, P_R\}$ satisfying from (14) and (95)

$$\begin{aligned} q'_1 + q'_2 + \ldots + q'_{R-K} &= Q'_1 + Q'_2 + \ldots + Q'_{R-K-s} + q_1 + q_2 + \ldots + q_s \\ &= P_{K+1} + P_{K+2} + \ldots + P_R - \sum_{i=1}^{s}(Q_i - q_i) \\ &\geq P_{K+1} + P_{K+2} + \ldots + P_R - m. \qquad (97) \end{aligned}$$

(Note that, from (93), $R - K \geq 1$ holds).

Definition 11. *For $1 \leq r \leq R$, $1 \leq j \leq \min(r, R-K) \leq R$ and $m \geq 0$, we define*

$$H(j, P_r; m) = \min(q'_1 q'_2 \ldots q'_j) \tag{98}$$

where the minimum is taken over the j-uples of primes $(q'_1, q'_2, \ldots, q'_j)$ satisfying

$$P_1 \le q'_1 < q'_2 < \ldots < q'_j \le P_r \tag{99}$$

and

$$q'_1 + q'_2 + \ldots + q'_j \ge P_{K+1} + P_{K+2} + \ldots + P_{K+j} - m. \tag{100}$$

If there is no $(q'_1, q'_2, \ldots, q'_j)$ such that (99) *and* (100) *hold, we set*

$$H(j, P_r; m) = +\infty. \tag{101}$$

By the unicity of the standard factorization into primes, the minimum in (98) is unique and (96) and (98) yield

$$G(p_k, m) = G(P_K, m) = \frac{P_{K+1}P_{K+2}\ldots P_R}{H(R-K, P_R; m)}. \tag{102}$$

For $j = R - K$ and $r = R$, the j-uple $q'_1, q'_2, \ldots, q'_j$ defined by $q'_i = P_{K+i}$ satisfies (99) and (100) for all $m \ge 0$; so, $H(R-K, P_R; m)$ is at most $P_{K+1}P_{K+2}\ldots P_R$ and is finite.

8.3. A combinatorial algorithm to compute H and G.

Definition 12. *For every integers (r, j), $1 \le r \le R$ and $1 \le j \le R - K$, we define*

$$m_j(P_r) = \begin{cases} P_{K+1} + P_{K+2} + \ldots + P_{K+j} - (P_r + P_{r-1} + \ldots + P_{r-j+1}) & \text{if } j \le r \\ +\infty & \text{if } j > r. \end{cases} \tag{103}$$

Remark: If $j \ge r+1$, (99) cannot be satisfied and, from (101), $H(j, P_r; m) = +\infty$ for all $m \ge 0$. If $j \le r$, from (100), it follows that, if $m \ge m_j(P_r)$, $H(j, P_r; m) \le P_r P_{r-1} \ldots P_{r-j+1}$ while, by (101), if $m < m_j(P_r)$, $H(j, P_r; m) = +\infty$. So that, in all cases, if $m < m_j(P_r)$, $H(j, P_r; m) = +\infty$.

Note that, for j fixed, $m_j(P_r)$ is non-increasing on r since, for $j \le r$,

$$m_j(P_{r-1}) - m_j(P_r) = \begin{cases} +\infty & \text{if } j = r \\ P_r - P_{r-j} > 0 & \text{if } 1 \le j \le r-1, \end{cases} \tag{104}$$

and, for $j \ge r+1$, $m_j(P_{r-1})$ and $m_j(P_r)$ are both $+\infty$. On the other hand, if $j \le \min(r, R-K)$ for every m such that

$$m \ge M_j(P_r) = P_{K+1} + P_{K+2} + \ldots + P_{K+j} - (P_1 + P_2 + \ldots + P_j),$$

$H(j, P_r; m)$ is equal to $P_1 P_2 \ldots P_j$.

Proposition 9. *For $j = 1$, from (98), (99) and (100), we have*
(105)

$$H(1, P_r; m) = \begin{cases} P_1 & \text{if } m \geq M_1(P_r) = P_{K+1} - P_1 \\ \ldots & \\ P_i & \text{if } 1 < i < r \text{ and } P_{K+1} - P_i \leq m < P_{K+1} - P_{i-1} \\ \ldots & \\ P_r & \text{if } m_1(P_r) = P_{K+1} - P_r \leq m < P_{K+1} - P_{r-1} \\ \infty & \text{if } m < m_1(P_r) = P_{K+1} - P_r. \end{cases}$$

Further, we have the induction formula:
(106)

$$H(j, P_r; m) = \min\left(H(j, P_{r-1}; m), P_r H(j-1, P_{r-1}; m - P_{K+j} + P_r)\right).$$

Proof. The calculation of $H(1, P_r; m)$ is easy. Let us show the induction formula (106). Either P_r does not divide $H(j, P_r; m)$ and $H(j, P_r; m) = H(j,\ P_{r-1}; m)$ or $P_r = q'_j$ is the greatest prime factor of $H(j, P_r; m) = q'_1 q'_2 \ldots q'_j$ and from (100), we get $q'_1 + \ldots + q'_{j-1} \geq P_{K+1} + \ldots + P_{K+j-1} - (m - P_{K+j} + P_r)$. □

Note that if $m \geq m_j(P_r)$, $m - P_{K+j} + P_r \geq m_{j-1}(P_{r-1})$ since $m_j(P_r) = m_{j-1}(P_{r-1}) + P_{K+j} - P_r$ so that $H(j, P_r; m)$ and $H(j-1, P_{r-1}; m - P_{K+j} + P_r)$ are simultaneously finite or infinite. (104) implies that $m_j(P_r)$ and $m_j(P_{r-1})$ are both infinite or $m_j(P_{r-1}) > m_j(P_r)$. For $m_j(P_r) \leq m < m_j(P_{r-1})$, (106) reduces to

$$H(j, P_r; m) = P_r H(j-1, P_{r-1}; m - P_{K+j} + P_r) \tag{107}$$

while, for $m \geq m_j(P_{r-1})$, the three values of the function H in (106) are finite.

From (105), we may remark that, if we set

$$H(0, P_r; m) = 1 \qquad \text{for all } r \geq 1 \text{ and } m \geq 0, \tag{108}$$

the induction formula (106) still holds for $j = 1$.

In view of (102), for $1 \leq r \leq R$, $1 \leq j \leq \min(r, R-K)$ and $m_j(P_r) \leq m \leq M$, we calculate $H(j, P_r; m)$ by induction, using for that (108), (106) and (107). If $K + 2 \leq r \leq R$, it is useless to calculate $H(j, P_r; m)$ for $j < r - K$.

Finally, after getting the value of $H(R-K, P_R; m)$ for $m_{R-K}(P_R) = 0 \leq m \leq M$, we compute $G(p_k, m)$ by (102).

8.4. Bounding the largest prime. It turns out that the largest prime used in the computation of $G(p_k, m)$ for $0 \leq m \leq M$ is much smaller than P_R defined in (93). For instance, for $p_k = P_K = 150989$ and $M = 5000$, R defined by (93) is at least equal to $K + 425$ while only the primes up to $p_{k+5} = P_{K+5} = 151027$ are used.

So, the idea is to replace R by a smaller number $\widehat{R}$, $K+1 \le \widehat{R} < R$, and to calculate by induction $H(\widehat{R}-K, P_{\widehat{R}}; m)$ instead of $H(R-K, P_R; m)$. We get the fraction $\widehat{F} = \dfrac{P_{K+1}P_{K+2}\dots P_{\widehat{R}}}{H(\widehat{R}-K, P_{\widehat{R}}; m)}$ which satisfies $\widehat{F} \le G(p_k, m)$. Now we have the following lemma.

Lemma 9. *Let F be a real number satisfying $1 < F \le G(p_k, m) = \dfrac{Q_1Q_2\dots Q_s}{q_1q_2\dots q_s}$. Then, the largest prime factor Q_s of the numerator of $G(p_k, m)$ is bounded above by*

$$Q_s \le \min\left(p_k + m, \frac{mF}{F-1}\right). \tag{109}$$

Proof. Using Lemma 1 and (15), we write

$$F \le G(p_k, m) = \frac{Q_1Q_2\dots Q_s}{q_1q_2\dots q_s} \le \frac{Q_s}{Q_s - \ell(G(p_k, m))} \le \frac{Q_s}{Q_s - m}$$

which yields $Q_s \le \dfrac{mF}{F-1}$. On the other hand, Inequality (13) together with (14) implies $Q_s - p_k \le Q_s - q_s \le m$ which completes the proof of (109). $\square$

If $\widehat{F} = \dfrac{P_{K+1}P_{K+2}\dots P_{\widehat{R}}}{H(\widehat{R}-K, P_{\widehat{R}}; m)} > 1$ and if $P_{\widehat{R}} > \min\left(P_K + m, \dfrac{m\widehat{F}}{\widehat{F}-1}\right)$, it follows from Lemma 9 that $G(p_k, m) = \widehat{F}$. If not, we start again by choosing a new value of $P_{\widehat{R}}$ greater than $\min\left(P_K + m, \dfrac{m\widehat{F}}{\widehat{F}-1}\right)$. Actually, Inequality (109) gives a reasonably good upper bound for Q_s. In the program, our first choice is $\widehat{R} = K + 10$.

8.5. Conclusion. The running time of the algorithm described in Sections 8.3 and 8.4 to calculate $G(p, m)$ for $m \le M$ grows about quadratically in M, so, it is rather slow when M is large.

For instance, the computation of $g(10^{15} - 741281)$ leads to the evaluation of $G(p, 688930)$ for $p = 192678883$, and this is not doable by the above combinatorial algorithm.

In the next section, we present a faster algorithm to compute $G(p_k, m)$ when m is large, but which does not work for small m's so that the two algorithms are complementary.

9. Computation of $G(p_k, m)$ for m large

The algorithm described in this section starts from the following two facts:

- if $G(p_k, m) = \dfrac{Q_1 Q_2 \dots Q_s}{q_1 q_2 \dots q_s}$ 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 $Q_1, \dots, Q_s, q_1, \dots, q_{s-1}$ are close to p_k. More precisely, $G(p_k, m)$ is equal to $\dfrac{p_{k+1}}{q_s} G(p_{k+1}, d)$ where $d = m - p_{k+1} + q_s$ is small.

  Note that when m is small $G(p_k, m)$ is not always equal to $\dfrac{p_{k+1}}{q_s} G(p_{k+1}, m - p_{k+1} + q_s)$. For instance, $G(103, 22) = \dfrac{107 \times 113}{97 \times 101}$ while $G(107, 12) = \dfrac{109}{97} < \dfrac{113}{101}\cdot$
- In (91), we have seen that $\ell(G(p, m)) = m$ implies $G(p, m) \geq 1 + \frac{m}{p_k}$, and it turns out that this last inequality seems to hold for m large enough.

9.1. A second way to compute $G(p_k, m)$. We want to compute $G(p_k, m)$ for a large m. The following proposition says that if, for some small δ, $p_k - m + \delta$ is prime and such that $G(p_{k+1}, \delta)$ 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'.

Proposition 10. *We want to compute $G(p_k, m)$ as defined in (12) or (88) with p_k odd and $p_{k+1} - p_k \leq m \leq p_{k+1} - 3$. We assume that we know some even non-negative integer δ satisfying*

$$p_{k+1} + \delta - m \quad \textit{is prime,} \tag{110}$$

$$G(p_{k+1}, \delta) \geq 1 + \frac{\delta}{p_{k+1}} \tag{111}$$

and

$$\delta < \frac{2m}{9} < \frac{2p_{k+1}}{9}\cdot \tag{112}$$

If $\delta = 0$, we know from Proposition 8 that $G(p_k, m) = \dfrac{p_{k+1}}{p_{k+1} - m}\cdot$ If $\delta > 0$, we have

$$G(p_k, m) = \max_{\substack{q \ \text{prime} \\ p_{k+1} - m \leq q \leq \widehat{q}}} \frac{p_{k+1}}{q} G(p_{k+1}, m - p_{k+1} + q), \tag{113}$$

where $\widehat{q}$ is defined by

$$\widehat{q} = \frac{p_{k+1} p_{k+2} (p_{k+1} - m + \delta)}{(p_{k+1} + \delta)(p_{k+1} - 3\delta/2)} \leq p_{k+2} - m + \frac{3\delta}{2}\cdot \tag{114}$$

Before proving Proposition 10 in Section 9.3, we shall first think to the possibility of applying it to compute $G(p_k, m)$.

9.2. Large differences between consecutive primes. For $x \geq 3$, let us define

$$\Delta(x) = \max_{p_j \leq x}(p_j - p_{j-1}). \tag{115}$$

Below, we give some values of $\Delta(x)$:

x	10^2	10^3	10^4	10^5	10^6	10^7	10^8	10^9	10^{10}	10^{11}	10^{12}
$\Delta(x)$	8	20	36	72	114	154	220	282	354	464	540
$(\log x)^2$	21	48	85	133	191	260	339	429	530	642	763

A table of $\Delta(x)$ up to $4 \cdot 10^{12}$ calculated by D. Shanks, L.J. Lander, T.R. Parkin and R. Brent can be found in [26], p. 85. There is a longer table (up to $8 \cdot 10^{16}$) on the web site [16]. H. Cramér conjectured in [4] that $\lim_{x\to\infty} \frac{\Delta(x)}{(\log x)^2} = 1$. For $x \leq 8 \cdot 10^{16}$, $\Delta(x) \leq 0.93(\log x)^2$ holds.

Let us set $\Delta = \Delta(p_{k+1})$; let us denote by $\delta_1 = \delta_1(p_k)$ the smallest even integer such that $\delta_1 \geq \Delta$ and

$$G(p_{k+1}, d) \geq 1 + \frac{d}{p_{k+1}}, \quad d = \delta_1 - \Delta + 2, \delta_1 - \Delta + 4, \ldots, \delta_1. \tag{116}$$

By using the combinatorial algorithm described in 8.3, we have computed that for all primes $p_k \leq 3 \cdot 10^8$, we have $\delta_1(p_k) \leq 900 = \delta_1(252314747)$ and

$$\delta_1(p_k) \leq 2.55(\log p_k)^2. \tag{117}$$

To compute the suffix of $g(n)$ for $n \leq 10^{15}$, we do not have to deal with larger values of p_k. However, for larger p_k's, we conjecture that $\delta_1(p_k)$ exists and is not too large.

Lemma 10. *Let p_k satisfy $5 \leq p_k \leq 3 \cdot 10^8$, m be an even integer such that $p_{k+1} - p_k \leq m \leq p_{k+1} - 3$, and $\delta_1 = \delta_1(p_k)$ defined by (116). If $m \geq \frac{9}{2}\delta_1(p_k)$, then there exists an even non-negative integer*

$$\delta = \delta(p_k, m) \leq \delta_1(p_k) \leq 2.55(\log p_k)^2 \tag{118}$$

such that (110), (111) and (112) hold. Therefore, Proposition 10 can be applied to compute $G(p_k, m)$.

Proof. Let us set $a = p_{k+1} + \delta_1(p_k) - m$. We have

$$a = p_{k+1} + \delta_1(p_k) - m \leq p_{k+1} - \frac{7}{2}\delta_1(p_k) \leq p_{k+1} - \frac{7}{2}\Delta < p_{k+1}.$$

Since $\delta_1 \geq \Delta$ and $m \leq p_{k+1} - 3$, $a \geq \Delta + 3$ holds. From the definition of $\Delta = \Delta(p_{k+1})$, there exists an even number b, $0 \leq b \leq \Delta - 2$ such that $a - b = p_{k+1} - m + (\delta_1 - b)$ is prime. From the definition of $\delta_1(p_k)$, we know that $G(p_{k+1}, \delta_1 - b) \geq 1 + \frac{\delta_1 - b}{p_{k+1}}$. Therefore, $\delta = \delta_1 - b$ satisfies (110), (111), (112) and $0 \leq \delta \leq \delta_1(p_k)$. The last upper bound of (118) follows from (117). $\square$

9.3. Proof of Proposition 10.

A polynomial equation of degree 2.

Lemma 11. *Let us consider real numbers T_1, T_2, δ satisfying*

$$0 < T_1 < T_2 \tag{119}$$

and

$$(\delta = 0 \;\; or \;\; \delta \geq T_2 - T_1) \quad and \quad \delta < \frac{2T_1}{9}\cdot \tag{120}$$

Note that (119) *and* (120) *imply*

$$T_1 + \delta \leq \frac{T_1 T_2}{T_2 - \delta}\cdot \tag{121}$$

Let m be a parameter satisfying

$$0 \leq \frac{9\delta}{2} \leq m < T_1. \tag{122}$$

We set

$$E(X) = X^2 - (T_1 + T_2 - m)X + \frac{T_1 T_2 (T_1 + \delta - m)}{T_1 + \delta}\cdot \tag{123}$$

1. *The equation $E(X) = 0$ has two roots X_1 and X_2 satisfying*

$$0 < X_1 < \frac{T_1 + T_2 - m}{2} < X_2 \leq T_2 - \delta. \tag{124}$$

2. *For T_1, T_2 and δ fixed and m in the range* (122), *X_2 is a non-decreasing function of m.*
3. *We have*

$$T_1 - \frac{3\delta}{2} < \frac{T_1 + 2T_2}{3} - \frac{3\delta}{2} \leq X_2 \leq T_2 - \delta. \tag{125}$$

4. *Let Y_1 and Y_2 be two positive real numbers satisfying*

$$Y_1 < Y_2, \quad Y_1 + Y_2 = T_1 + T_2 - m \;\; and \;\; \frac{T_1 T_2}{Y_1 Y_2} \geq \frac{T_1 + \delta}{T_1 + \delta - m}\cdot \tag{126}$$

We have

$$Y_2 \geq X_2 \geq T_1 - \frac{3\delta}{2} \qquad and \qquad Y_1 \leq X_1 \leq T_2 - m + \frac{3\delta}{2}. \tag{127}$$

Proof. 1. The discriminant D of (123) can be written as

$$\begin{aligned} D &= (T_1 + T_2 - m)^2 - 4\frac{T_1 T_2 (T_1 + \delta - m)}{T_1 + \delta} \\ &= (m + T_2 - T_1)^2 \left[1 - \frac{4\delta}{m} \frac{m^2 T_2}{(m + T_2 - T_1)^2 (T_1 + \delta)}\right], \end{aligned} \tag{128}$$

since, from (119) and (122), $m+T_2-T_1$ does not vanish. If $\delta=0$, the above bracket is 1 while if $\delta \geq T_2-T_1>0$, the fractions $\dfrac{T_2}{T_1+\delta}$ and $\dfrac{m}{m+T_2-T_1}$ are at most 1, so that in both cases (128) yields

$$D \geq (m+T_2-T_1)^2\left[1-\frac{4\delta}{m}\right]. \tag{129}$$

Therefore, from (122) and (119), $D \geq \dfrac{(m+T_2-T_1)^2}{9}>0$ holds.

The sum X_1+X_2 of the two roots is T_1+T_2-m which explains the second and the third inequality of (124). Further, since $T_1<T_2$ and $m \geq 2\delta$, $\dfrac{T_1+T_2-m}{2} \leq T_2-\delta$ holds. By (123), (122) and (121),

$$E(T_2-\delta)=(T_1+\delta-m)\left(\frac{T_1T_2}{T_1+\delta}-(T_2-\delta)\right) \geq 0$$

which proves the last inequality of (124).

Remark: If $\delta=0$, the roots of (123) are $X_1=T_1-m$ and $X_2=T_2$. If $\delta=T_2-T_1$, they are $X_1=T_2-m$ and $X_2=T_1$.

2. By (123), X_2 is implicitely defined in terms of m and, through (121), we have

$$\frac{d\,X_2}{d\,m}=\frac{-\frac{\partial E}{\partial m}}{\frac{\partial E}{\partial X}}=\frac{\frac{T_1T_2}{T_1+\delta}-X_2}{2X_2-(T_1+T_2-m)} \geq \frac{T_2-\delta-X_2}{2X_2-(T_1+T_2-m)}$$

which is non-negative from (124).

3. For $m=\frac{9\delta}{2}$, (129) yields $\sqrt{D} \geq \frac{m+T_2-T_1}{3}=\frac{3\delta}{2}+\frac{T_2-T_1}{3}$ and

$$X_2=\frac{T_1+T_2-m+\sqrt{D}}{2} \geq \frac{T_1+2T_2}{3}-\frac{3\delta}{2} \geq T_1-\frac{3\delta}{2}\cdot$$

Further, for $m \geq \frac{9\delta}{2}$, the lower bound in (125) follows from 2.

4. Conditions (126) imply $E(Y_1)=E(Y_2)=-Y_1Y_2+\dfrac{T_1T_2(T_1+\delta-m)}{T_1+\delta}$ ≥ 0 so that $Y_1 \leq X_1$ and $Y_2 \geq X_2$; (127) follows from (125) and from $X_1=T_1+T_2-m-X_2$.

□

Structure of the fraction $G(p_k,m)$.

Lemma 12. *Let k and m be integers such that $k \geq 3$ and $p_{k+1}-p_k \leq m \leq p_{k+1}-3$. We write*

$$G(p_k,m)=F=\frac{Q_1Q_2\dots Q_s}{q_1q_2\dots q_s} \tag{130}$$

with $s \geq 1$ *and* $Q_1, \dots, Q_s, q_1, \dots, q_s$ *primes satisfying*

$$3 \leq q_s < q_{s-1} < \dots < q_1 \leq p_k < p_{k+1} \leq Q_1 < \dots Q_{s-1} < Q_s, \tag{131}$$

$$p_{k+1} - p_k \leq \ell(F) = \sum_{i=1}^{s} (Q_i - q_i) \leq m \leq p_{k+1} - 3 < p_{k+1} \tag{132}$$

and we assume that there exists an integer δ *such that*

$$0 \leq \delta < \frac{2m}{9}, \quad \text{and} \quad (\delta = 0 \text{ or } \delta \geq p_{k+2} - p_{k+1}) \tag{133}$$

and

$$F \geq \frac{p_{k+1} + \delta}{p_{k+1} - m + \delta}. \tag{134}$$

We apply Lemma 11 with $T_1 = p_{k+1}$ *and* $T_2 = p_{k+2}$, δ *and* m, *and we denote by* X_1 *and* X_2 *the two roots of equation* (123), $E(X) = 0$. *Then we have*

1. $$Q_s \leq p_{k+1} + \delta,$$

2. *for* $s \geq 2$ *and* $1 \leq i \leq s-1$, $\qquad \lambda_i \stackrel{def}{=\!=} Q_i - q_i \leq p_{k+2} - X_2,$

3. *for* $s \geq 2$ *and* $1 \leq j \leq s-1$, $\qquad \Lambda_j \stackrel{def}{=\!=} \sum_{i=1}^{j} \lambda_i \leq p_{k+2} - X_2.$

Moreover, if we write $F = UV$ *with*

$$U = \frac{Q_1 Q_2 \dots Q_{s-1} Q_s}{q_1 q_2 \dots q_{s-1} p_{k+1}} \quad \text{and} \quad V = \frac{p_{k+1}}{q_s}, \tag{135}$$

we have, for $s \geq 1$

4. $$\ell(U) = \Lambda_{s-1} + Q_s - p_{k+1} \leq p_{k+2} - X_2 \leq p_{k+2} - p_{k+1} + \frac{3\delta}{2}$$

and

5. $$p_{k+1} - m \leq q_s \leq \widehat{q} = \frac{p_{k+1} p_{k+2} (p_{k+1} - m + \delta)}{(p_{k+1} + \delta)(p_{k+1} - 3\delta/2)}.$$

Proof. 1. First, we observe that (131) implies

$$Q_i \geq p_{k+i} \geq p_{k+1}, \qquad 1 \leq i \leq s. \tag{136}$$

Lemma 1 and (132) yield respectively $F \leq \frac{Q_s}{Q_s - \ell(F)}$ and $\ell(F) \leq m$, so that, together with (134), we get

$$\frac{p_{k+1} + \delta}{p_{k+1} + \delta - m} \leq F \leq \frac{Q_s}{Q_s - \ell(F)} \leq \frac{Q_s}{Q_s - m}$$

which, with the decrease of $t \mapsto \frac{t}{t-m}$, gives $Q_s \leq p_{k+1} + \delta$.

2. From the definition of λ_i and (131), λ_i is positive and increasing on i, and it suffices to show $\lambda_{s-1} \le p_{k+2} - X_2$. We write $F = F_1 F_2$ with $F_1 = \frac{Q_{s-1}}{q_{s-1}}$ and $F_2 = \prod_{i\neq s-1} \frac{Q_i}{q_i}$. From (132) and (131), we have

$$p_{k+1} > m > m - \lambda_{s-1} \ge \ell(F) - \lambda_{s-1} = \lambda_1 + \ldots + \lambda_{s-2} + \lambda_s \ge \lambda_s > \lambda_{s-1}$$

which implies

$$p_{k+2} - \lambda_{s-1} > p_{k+1} - \lambda_{s-1} > p_{k+1} - (m - \lambda_{s-1}). \tag{137}$$

Further, Lemma 1, (132) and the increase of $t \mapsto \frac{Q_s}{Q_s - t}$, (136) and the decrease of $t \mapsto \frac{t}{t-(m-\lambda_{s-1})}$, imply

$$\begin{aligned} F_2 &\le \frac{Q_s}{Q_s - \ell(F_2)} = \frac{Q_s}{Q_s - (\ell(F) - \lambda_{s-1})} \\ &\le \frac{Q_s}{Q_s - (m - \lambda_{s-1})} \le \frac{p_{k+1}}{p_{k+1} - (m - \lambda_{s-1})}. \end{aligned} \tag{138}$$

If $s \ge 3$ or $Q_1 \ge p_{k+2}$, (131) implies $Q_{s-1} \ge p_{k+2}$ which yields $F_1 = \frac{Q_{s-1}}{Q_{s-1}-\lambda_{s-1}} \le \frac{p_{k+2}}{p_{k+2}-\lambda_{s-1}}$ so that, from (134) and (138), we get

$$\frac{p_{k+1} + \delta}{p_{k+1} + \delta - m} \le F = F_1 F_2 \le \frac{p_{k+2}}{p_{k+2} - \lambda_{s-1}} \, \frac{p_{k+1}}{p_{k+1} - (m - \lambda_{s-1})}. \tag{139}$$

Let us set $Y_2 = p_{k+2} - \lambda_{s-1}$, $Y_1 = p_{k+1} - (m - \lambda_{s-1})$; from (137), $Y_2 > Y_1$ holds and, in view of (139), we may apply Lemma 11, Point 4. to get $Y_2 = p_{k+2} - \lambda_{s-1} \ge X_2$ which implies 2.

If $s = 2$ and $Q_1 = p_{k+1}$, $F = \frac{p_{k+1}}{q_1}\frac{Q_2}{q_2} = \frac{p_{k+1}}{q_1}\frac{Q_2}{Q_2-(Q_2-q_2)}$. From (136) we have $Q_2 \ge p_{k+2}$ and $F \le \frac{p_{k+1}}{q_1}\frac{p_{k+2}}{p_{k+2}-(Q_2-q_2)}$. Here we set $Y_2 = q_1$ and $Y_1 = p_{k+2} - (Q_2 - q_2) = q_2 - (Q_2 - p_{k+2})$; by (131) and (132), we get

$$\begin{aligned} Y_2 = q_1 > q_2 \ge Y_1 = q_2 - (Q_2 - p_{k+2}) &= p_{k+2} - \lambda_2 \\ &\ge p_{k+2} - \sum_{i=1}^{2} \lambda_i = p_{k+2} - \ell(F) \ge p_{k+2} - m > 0; \end{aligned}$$

we may still apply Lemma 11 Point 4. to get $Y_2 = q_1 = p_{k+1} - \lambda_1 \ge X_2$, which implies 2.

3. This time, we write $F = F_1 F_2$ with $F_1 = \prod_{i=1}^{j} \frac{Q_i}{q_i}$ and $F_2 = \prod_{i=j+1}^{s} \frac{Q_i}{q_i}$ so that $\ell(F_1) = \Lambda_j$ and $\ell(F_2) = \ell(F) - \Lambda_j \le m - \Lambda_j$.

For $2 \le j \le s-1$, from (134), Lemma 1, (136), and (132) we get

$$\frac{p_{k+1}+\delta}{p_{k+1}+\delta-m} \le F = F_1F_2 \le \frac{Q_j}{Q_j-\ell(F_1)}\,\frac{Q_s}{Q_s-\ell(F_2)} \le \frac{p_{k+2}}{p_{k+2}-\Lambda_j}\,\frac{p_{k+1}}{p_{k+1}-(m-\Lambda_j)}.$$

Therefore, we apply Lemma 11 Point 4., but we do not know whether $p_{k+2}-\Lambda_j$ is greater than $p_{k+1}-(m-\Lambda_j)$, so that, either

$$p_{k+2}-\Lambda_j \ge X_2 \tag{140}$$

or

$$p_{k+2}-\Lambda_j \le X_1. \tag{141}$$

For $j=1$, as $\Lambda_1=\lambda_1$, (140) holds, from 2. Since Λ_j is increasing on j, if (140) holds for some $j=j_0$, it also holds for $j\le j_0$. If (140) holds for $j=s-1$, 3. is proved; so, let us assume that the greatest value j_0 for which (140) holds satisfies $1\le j_0 < s-1$; we should have

$$p_{k+2}-\Lambda_{j_0} \ge X_2 \qquad \text{and} \qquad p_{k+2}-\Lambda_{j_0+1} \le X_1. \tag{142}$$

From 2., (142) and because X_1, X_2 are solutions of (123), we should get

$$p_{k+2}-X_2 \ge \lambda_{j_0+1} = \Lambda_{j_0+1}-\Lambda_{j_0} \ge X_2-X_1 = 2X_2+m-p_{k+1}-p_{k+2}$$

which, would imply $m \le 2p_{k+2}+p_{k+1}-3X_2$ and, through the second inequality of (125), $m\le \frac{9\delta}{2}$, in contradiction with (133). Therefore, $j_0 \ge s-1$ and 3. is proved.

4. If $s=1$ we have to show $\ell(U)=Q_1-p_{k+1}\le p_{k+2}-X_2$ which is true since, from 1., $Q_1-p_{k+1}\le\delta$ and from (125), with $T_2=p_{k+2}$, $\delta\le p_{k+2}-X_2$.

   So, we assume $s\ge 2$. If $Q_1=p_{k+1}$, U simplifies itself; and, in all cases, from (131), the prime factors of the numerator of U are at least p_{k+2} and those of the denominator are at most p_{k+1}. So, we may apply Lemma 1 which, with (136) and the decrease of $t\mapsto t/(t-\ell(U)$, yields

$$U \le \frac{Q_s}{Q_s-\ell(U)} \le \frac{p_{k+2}}{p_{k+2}-\ell(U)}, \qquad V=\frac{p_{k+1}}{p_{k+1}-\ell(V)}. \tag{143}$$

It follows from (132) that $\ell(U)+\ell(V)=\ell(F)\le m$ and, from (134), we get

$$\frac{p_{k+1}+\delta}{p_{k+1}+\delta-m} \le F = UV \le \frac{p_{k+2}p_{k+1}}{(p_{k+2}-\ell(U))(p_{k+1}-(m-\ell(U)))}.$$

Applying Lemma 11 Point 4. with $(Y_1, Y_2) = (p_{k+2} - \ell(U), p_{k+1} - (m - \ell(U)))$ yields

$$p_{k+2} - \ell(U) \geq X_2 \qquad \text{or} \qquad p_{k+2} - \ell(U) \leq X_1. \tag{144}$$

But, from 1. and 3., we have $\ell(U) = \Lambda_{s-1} + Q_s - p_{k+1} \leq p_{k+2} - X_2 + \delta$ which, together with (X_1, X_2) solutions of (123), the second inequality in (125) and (133), give

$$\begin{aligned} X_1 + \ell(U) - p_{k+2} \leq X_1 - X_2 + \delta &= \delta + p_{k+1} + p_{k+2} - m - 2X_2 \\ &\leq \delta + p_{k+1} + p_{k+2} - m - \frac{2}{3}(p_{k+1} + 2p_{k+2}) + 3\delta \\ &= 4\delta + \frac{p_{k+1} - p_{k+2}}{3} - m < 0. \end{aligned}$$

Therefore, $p_{k+2} - \ell(U) \leq X_1$ does not hold, and, from (144), we have $p_{k+2} - \ell(U) \geq X_2$ which shows the first inequality in 4. The second inequality comes from (125).

5. From (131) and (132), we have $\ell(V) = p_{k+1} - q_s \leq Q_s - q_s \leq \ell(F) \leq m$ which proves the lower bound of 5.

If $s = 1$ and $Q_1 = p_{k+1}$, $U = 1$ and $F = V$ so that, from (134),

$$q_s = \frac{p_{k+1}}{F} \leq \frac{p_{k+1}(p_{k+1} - m + \delta)}{p_{k+1} + \delta} \leq \widehat{q} = \frac{p_{k+1}p_{k+2}(p_{k+1} - m + \delta)}{(p_{k+1} + \delta)(p_{k+1} - 3\delta/2)}.$$

If $s \geq 2$ or $Q_1 \geq p_{k+2}$, (143) holds and gives with (134) and 4.

$$q_s = \frac{p_{k+1}}{V} = \frac{p_{k+1}U}{F} \leq \frac{p_{k+1}p_{k+2}(p_{k+1} - m + \delta)}{(p_{k+1} + \delta)(p_{k+2} - \ell(U))} \leq \widehat{q}.$$

□

Proof of Proposition 10. Let us assume $\delta > 0$. (111) and (89) imply

$$\delta \geq p_{k+2} - p_{k+1}. \tag{145}$$

First, we prove the upper bound (114). We have to show that the quantity below is positive:

$$(p_{k+2} - m + \delta)(p_{k+1} + \delta)\Big(p_{k+1} - \frac{3\delta}{2}\Big) - p_{k+1}p_{k+2}(p_{k+1} - m + \delta).$$

But this quantity is equal to

$$\begin{aligned} &(p_{k+2} - p_{k+1})\Big((p_{k+1} - \delta)(m - \frac{3\delta}{2}) + \delta(m - 3\delta)\Big) \\ &\qquad\qquad + p_{k+1}\frac{\delta}{2}\Big(m - \frac{9\delta}{2}\Big) + \frac{3\delta^2}{4}\Big(m - \frac{3\delta}{2}\Big) \end{aligned}$$

which is clearly positive since, from (112), $p_{k+1} > m > \frac{9\delta}{2}$ holds and (114) is proved.

Let q be a prime satisfying $p_{k+1} - m \le q \le \widehat{q}$. In view of proving (113), let us show that

$$\frac{p_{k+1}}{q} G(p_{k+1}, m - p_{k+1} + q) \le G(p_k, m) \tag{146}$$

holds. Let q' be any prime dividing the denominator of $G(p_{k+1}, m-p_{k+1}+q)$; we should have $p_{k+2} - q' \le m - p_{k+1} + q$ i.e., $q' \ge p_{k+1} + p_{k+2} - m - q$ which yields from (114), (145) and (112)

$$\begin{aligned} q' - q &\ge p_{k+1} + p_{k+2} - m - 2q \ge p_{k+1} + p_{k+2} - m - 2\widehat{q} \\ &\ge p_{k+1} + p_{k+2} - m - 2\left(p_{k+2} - m + \frac{3\delta}{2}\right) = p_{k+1} - p_{k+2} + m - 3\delta \\ &\ge p_{k+1} - (\delta + p_{k+1}) + m - 3\delta = m - 4\delta > 0. \end{aligned}$$

Therefore, $q' \ne q$, and after a possible simplification by p_{k+1}, $\frac{p_{k+1}}{q} G(p_{k+1}, m - p_{k+1} + q) \in \mathcal{G}(p_k, m)$ (defined in (87)), which, from (88), implies (146).

From (145) and (112), we have $0 < 2\delta < m$, and the prime $p = p_{k+1} + \delta - m$ satisfies $p < p_{k+2} - \delta$, and thus is smaller than any prime factor of the denominator of $G(p_{k+1}, \delta)$. Therefore, after possibly simplifying by p_{k+1}, the fraction $\Phi = \frac{p_{k+1}}{p} G(p_{k+1}, \delta)$ belongs to $\mathcal{G}(p_k, m)$ and we have from (88) and (111)

$$G(p_k, m) \ge \Phi \ge \frac{p_{k+1}}{p_{k+1} + \delta - m}\left(1 + \frac{\delta}{p_{k+1}}\right) = \frac{p_{k+1} + \delta}{p_{k+1} + \delta - m}.$$

So, hypotheses (133) and (134) being fullfilled, we may apply Lemma 12 5. which, under the notation (135), asserts that

$$G(p_k, m) = UV = U\frac{p_{k+1}}{q_s} \tag{147}$$

with $q_s \in [\,p_{k+1} - m, \widehat{q}\,]$ and $\ell(U) + \ell(V) = \ell(G(p_k, m))$ which, from (15), implies $\ell(U) \le m - \ell(V) = m - p_{k+1} + q_s$. After a possible simplification by p_{k+1}, U belongs to $\mathcal{G}(p_{k+1}, \ell(U)) \subset \mathcal{G}(p_{k+1}, m - p_{k+1} + q_s)$. So, from (88), $U \le G(p_{k+1}, m - p_{k+1} + q_s)$, and (147) gives

$$G(p_k, m) \le \frac{p_{k+1}}{q_s} G(p_{k+1}, m - p_{k+1} + q_s)$$

which, with (146), completes the proof of (113) and of Proposition 10. $\square$

10. Some results

With the MAPLE program available on the web-site of J.-L. Nicolas, the factorization of $g(n)$ has been computed for some values of n. The results for $n = 10^6, 10^9, 10^{12}, 10^{15}$ are displayed in FIG. 6. For primes $q_1 < q_2$ let us denote by $[q_1 - q_2]$ the product $\prod_{q_1 \le p \le q_2} p$. The bold factors in the values of $g(n)$ are the factors of the plain prefix π of $g(n)$, defined in (8).

$$n = 10^6, \qquad N = 2^9 3^6 5^4 7^3 [11-41]^2 [43-3923],$$
$$\ell(N) = 998093, \qquad g(10^6) = g(10^6-1) = \frac{\mathbf{43}\cdot 3947}{3847} N.$$

$$n = 10^9, \quad N = 2^{14} 3^9 5^6 7^5 11^4 13^4 [17-31]^3 [37-263]^2 [269-150989],$$
$$\ell(N) = 999969437, \qquad g(10^9) = g(10^9-1) = \frac{\mathbf{37}\cdot 150991}{\mathbf{2}\cdot\mathbf{3}\cdot 148399} N.$$

$$n = 10^{12}, \qquad N = 2^{18} 3^{12} 5^8 7^6 11^5 13^5 [17-31]^4 [37-113]^3 \times [127-1613]^2 [1619-5476469],$$
$$\ell(N) = 999997526071, g(10^{12}) = \frac{\mathbf{1621}\cdot\mathbf{1627}\cdot\mathbf{1637}\cdot 5476483}{5475739\cdot 5476469} N.$$

$$n = 10^{15}, \quad N = 2^{23} 3^{15} 5^{10} 7^8 11^7 13^6 17^6 [19-31]^5 [37-79]^4 [83-389]^3 \times [397-9623]^2 [9629-192678817],$$
$$\ell(N) = 999999940824564,$$
$$g(10^{15}) = g(10^{15}-1) = \frac{192678823\cdot 192678853\cdot 192678883\cdot 192678917}{\mathbf{389}\cdot\mathbf{9539}\cdot\mathbf{9587}\cdot\mathbf{9601}\cdot\mathbf{9619}\cdot\mathbf{9623}\cdot 192665881} N.$$

FIGURE 6. The values $g(n)$ for $n = 10^6, 10^9, 10^{12}, 10^{15}$.

On a 3GHz Pentium 4, the time of computation of $g(n)$ is about 0.02 second for an integer n of 6 decimal digits and 10 seconds for 15 digits.

11. Open problems

11.1. An effective bound for the benefit. Let us define $\operatorname{ben} g(n)$ by (36) with N and ρ defined by (32) and (30). Is it possible to get an effective form of (42), i.e.,

$$\operatorname{ben} g(n) + n - \ell(g(n)) \le C\rho$$

for some absolute constant C to determine?

A hint is to apply Proposition 2 with $M = \frac{P_1 P_2 \dots P_r}{q_1 q_2 \dots q_{2r}}$ for some r, where the P_i's are the r smallest primes not dividing N and the q_i's are the $2r$ largest primes such that $v_{q_i}(N) = 2$, and, further, to apply effective results on the Prime Number Theorem like those of [28] or [5].

11.2. Increasing subsequences of $g(n)$. An increasing subsequence of g is a set of k consecutive integers $\{n, n+1, \dots, n+k-1\}$ such that

$$g(n-1) = g(n) < g(n+1) < \dots < g(n+k-1) = g(n+k). \tag{148}$$

Due to a parity phenomenom, these maximal sequences are rare. For $n \le 10^6$, there are only 9 values on n with $k \ge 7$. The record is $n = 35464$ with $k = 20$.

Are there arbitrarily long maximal sequences? It seems to be a very difficult question. In [21], (1.7), it is conjectured that there are infinitely many maximal sequences with $k \ge 2$.

11.3. The second minimum. Let us write $g_1(n) = g(n) > g_2(n) > \dots > g_I(n) = 1$ all the integers such that, if $\sigma \in \mathfrak{S}_n$, the order of σ is equal to $g_i(n)$ for some $i \in \{1, 2, \dots, I\}$. From (5), I is equal to the number of positive integers M satisfying $\ell(M) \le n$.

We might be interested in the computation of $g_2(n)$ or more generally, in the computation of $g_i(n)$ for $1 \le i \le i_0$ where i_0 is some (small) fixed constant.

The basic algorithm (see Section 2) can be easily adapted for this purpose. It seems reasonable to think that our algorithm, as sketched in 1.3, can also be extended to get $g_i(n)$.

11.4. Computing $h(n)$. Let $h(n)$ be the maximal product of primes $p_{i_1}, p_{i_2}, \dots, p_{i_r}$ under the condition $p_{i_1} + p_{i_2} + \dots + p_{i_r} \le n$ (r is not fixed); $h(n)$ can be interpreted as the maximal order of a permutation of the symmetric group $\mathfrak{S}_n$ such that the lengths of its cycles are all primes.

A formula similar to (2) can be written:

$$h(n) = \max_{\substack{M \text{ squarefree} \\ \ell(M) \le n}} M.$$

The superchampion numbers are the product of the first primes.

A related problem is to find an algorithm to compute $h(n)$ for n up to 10^{15}.

11.5. Maximum order in $GL(n, \mathbb{Z})$. Let $G(n)$ be the maximum order of torsion elements in $GL(n, \mathbb{Z})$. It has been shown in [10] that

$$G(n) = \max_{L(M) \le n} M \tag{149}$$

where L is the additive function defined by $L(1) = L(2) = 0$ and $L(p^\alpha) = \varphi(p^\alpha) = p^\alpha - p^{\alpha-1}$ if $p^\alpha \ge 3$.

From (149) and (2), it follows that $g(n) \leq G(n)$ holds for all n's and it has been shown in [22] that $\lim_{n\to\infty} G(n)/g(n) = \infty$.

Is it possible to adapt the algorithm described in this paper to compute $G(n)$ up to 10^{15}?

References

[1] E. BACH, *Sums over Primes*. Conference on Algorithmic Number Theory, Turku, May 8-11, 2007.

[2] E. BACH and J. SORENSON, *Computing prime harmonic sums*. Submitted to Math. Comp.

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

[4] H. CRAMÉR, *On the order of magnitude of the difference between consecutive prime numbers*. Acta Arithmetica **2** (1936), 23–46.

[5] 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.

[6] P. ERDŐS and P. TURÁN, *On some problems of a statistical group theory, IV*. Acta Math. Acad. Sci. Hungar. **19** (1968), 413–435.

[7] H. GERLACH, *Über die Elemente einer Menge verallgemeinerter ganzer Zahlen, die klein sind bezüglich einer auf dieser Menge definierten reellwertigen Abbildung*. Thesis of the University of Kaiserslautern, 1986.

[8] J. GRANTHAM, *The largest prime dividing the maximal order of an element of S_n*. Math. Comp. **64** (1995), 407–410.

[9] E. LANDAU, *Über die Maximalordnung der Permutationen gegebenen Grades*. Archiv. der Math. und Phys., Sér 3, **5** (1903), 92–103. Handbuch der Lehre von der Verteilung der Primzahlen, I, 2nd ed, Chelsea, New-York, 1953, 222–229.

[10] G. LEVITT and J.-L. NICOLAS, *On the Maximum Order of Torsion Elements in $GL(n,\mathbb{Z})$ and $Aut(F_n)$*. Journal of Algebra **208** (1998), 630–642.

[11] J.-P. MASSIAS, *Majoration explicite de l'ordre maximum d'un élément du groupe symétrique*. Ann. Fac. Sci. Toulouse Math. **6** (1984), 269–280.

[12] J.-P. MASSIAS, J.-L. NICOLAS et G. ROBIN. *Évaluation asymptotique de l'ordre maximum d'un élément du groupe symétrique*. Acta Arithmetica **50** (1988), 221–242.

[13] J.-P. MASSIAS, J.-L. NICOLAS and G. ROBIN, *Effective Bounds for the Maximal Order of an Element in the Symmetric Group*. Math. Comp. **53** (1989), 665–678.

[14] W. MILLER, *The Maximal Order of an Element of a Finite Symmetric Group*. Amer. Math. Monthly **94** (1987), 497–506.

[15] F. MORAIN, *Table de $g(n)$ pour $1 \leq n \leq 32000$*. Internal document, INRIA, 1988.

[16] T. R. NICELY, `http://www.trnicely.net/gaps/gaplist.html`

[17] J.-L. NICOLAS, *Sur l'ordre maximum d'un élément dans le groupe S_n des permutations*. Acta Arithmetica **14** (1968), 315–332.

[18] J.-L. NICOLAS, *Ordre maximal d'un élément du groupe des permutations et highly composite numbers*. Bull. Soc. Math. France **97** (1969), 129–191.

[19] J.-L. NICOLAS, *Calcul de l'ordre maximum d'un élément du groupe symétrique S_n*. Rev. Française Informat. Recherche Opérationnelle, Sér. R-2 **3** (1969), 43–50.

[20] J.-L. NICOLAS, *On highly composite numbers. Ramanujan revisited*, edited by G. E. Andrews, R. A. Askey, B. C. Berndt, K. G. Ramanathan, R. A. Rankin, Academic Press, 1988, 216–244.

[21] J.-L. NICOLAS, *On Landau's function $g(n)$. The Mathematics of Paul Erdős*, vol. I, R. L. Graham and J. Nešetřil editors, Springer Verlag, Algorithms and Combinatorics n°**13** (1997), 228–240.

[22] J.-L. NICOLAS, *Comparaison des ordres maximaux dans les groupes S_n et $GL(n,\mathbb{Z})$*. Acta Arithmetica **96** (2000), 175–203.

[23] J.-L. NICOLAS and N. ZAKIC, *Champion numbers for the number of representations as a sum of six squares*. In preparation.

[24] S. Ramanujan, *Highly composite numbers. Proc. London Math. Soc.* Serie 2, **14** (1915), 347–409. Collected papers, Cambridge University Press, 1927, 78–128.
[25] S. Ramanujan, *Highly composite numbers, annotated by J.-L. Nicolas and G. Robin.* The Ramanujan J. **1** (1997), 119–153.
[26] H. Riesel, *Prime Numbers and Computer Methods for Factorization.* Birkhäuser, 1985.
[27] G. Robin, *Méthodes d'optimisation pour un problème de théorie des nombres.* R.A.I.R.O. Informatique Théorique **17** (1983), 239–247.
[28] J. B. Rosser and L. Schoenfeld, *Approximate Formulas for Some Functions of Prime Numbers.* Illinois. J. Math **6** (1962), 64–94.
[29] S. M. Shah, *An inequality for the arithmetical function $g(x)$.* J. Indian Math. Soc. **3** (1939), 316–318.
[30] M. Szalay, *On the maximal order in S_n and S_n^*.* Acta Arithmetica **37** (1980), 321–331.
[31] P. M. B. Vitányi, *On the size of DOL languages.* Lecture Notes in Computer Science, vol. **15**, Springer-Verlag, 1974, 78–92 and 327–338.

Marc Deléglise
Université de Lyon,
Université de Lyon 1, CNRS,
Institut Camille Jordan,
Bât. Doyen Jean Braconnier,
21 Avenue Claude Bernard,
F-69622 Villeurbanne cedex, France.
E-mail: `deleglis@math.univ-lyon1.fr,jlnicola@in2p3.fr`
URL: `http://math.univ-lyon1.fr/~deleglis`

Jean-Louis Nicolas
Université de Lyon,
Université de Lyon 1, CNRS,
Institut Camille Jordan,
Bât. Doyen Jean Braconnier,
21 Avenue Claude Bernard,
F-69622 Villeurbanne cedex, France.
E-mail: `jlnicola@in2p3.fr`
URL: `http://math.univ-lyon1.fr/~nicolas/`

Paul Zimmermann
Centre de Recherche INRIA Nancy Grand Est
Projet CACAO -bâtiment A
615 rue du Jardin Botanique,
F-54602 Villers-lès-Nancy cedex, France.
E-mail: `Paul.Zimmermann@loria.fr`
URL: `http://www.loria.fr/~zimmerma/`