Integers **11** (2011), 647–669
DOI 10.1515/INTEG.2011.049

On the Distance Between Smooth Numbers

Jean-Marie De Koninck and Nicolas Doyon

Abstract. Let $P(n)$ stand for the largest prime factor of $n \geq 2$ and set $P(1) = 1$. For each integer $n \geq 2$, let $\delta(n)$ be the distance to the nearest $P(n)$-smooth number, that is, to the nearest integer whose largest prime factor is no larger than that of n. We provide a heuristic argument showing that $\sum_{n \leq x} 1/\delta(n) = (4 \log 2 - 2 + o(1))x$ as $x \to \infty$. Moreover, given an arbitrary real-valued arithmetic function f, we study the behavior of the more general function $\delta_f(n)$ defined by $\delta_f(n) = \min_{1 \leq m \neq n,\, f(m) \leq f(n)} |n - m|$ for $n \geq 2$, and $\delta_f(1) = 1$. In particular, given any positive integers $a < b$, we show that $\sum_{a \leq n < b} 1/\delta_f(n) \geq 2(b-a)/3$ and that if $f(n) \geq f(a)$ for all $n \in [a, b[$, then one has $\sum_{a < n < b} \delta_f(n) \leq (b-a) \log(b-a)/(2 \log 2)$.

Keywords. Arithmetic Function, Smooth Numbers.

2010 Mathematics Subject Classification. 11A25, 11N37.

1 Introduction

A *smooth number* (or a *friable number*) is a positive integer n whose largest prime factor is "small" compared to n. Hence, given an integer $B \geq 2$, we say that an integer n is B-*smooth* if all its prime factors are $\leq B$.

Let $P(n)$ stand for the largest prime factor of n (with $P(1) = 1$).

For each integer $n \geq 2$, let $\delta(n)$ be the distance to the nearest $P(n)$-smooth number, that is, to the nearest integer whose largest prime factor is no larger than that of n. In other words,

$$\delta(n) := \min_{\substack{1 \leq m \neq n \\ P(m) \leq P(n)}} |n - m|.$$

Equivalently, if we let

$$\Psi(x, y) := \#\{n \leq x : P(n) \leq y\},$$

then $\delta(n)$ is the smallest positive integer δ such that either one of the following two equalities occur:

$$\Psi(n + \delta, P(n)) - \Psi(n, P(n)) = 1, \quad \Psi(n, P(n)) - \Psi(n - \delta, P(n)) = 1.$$

Research work supported in part by a grant from NSERC.

For convenience, we set $\delta(1) = 1$. In particular,

$$\delta(2^a) = 2^{a-1} \quad \text{for each integer } a \geq 1. \tag{1}$$

The first 40 values of $\delta(n)$ are

$$\begin{array}{cccccccccccccccccccc} 1, & 1, & 1, & 2, & 1, & 2, & 1, & 4, & 1, & 1, & 1, & 3, & 1, & 1, & 1, & 8, & 1, & 2, & 1, & 2, \\ 1, & 1, & 1, & 3, & 1, & 1, & 3, & 1, & 1, & 2, & 1, & 16, & 1, & 1, & 1, & 4, & 1, & 1, & 1, & 4. \end{array}$$

We call $\delta(n)$ the *index of isolation* of n and we say that an integer n is *isolated* if $\delta(n) \geq 2$ and *non-isolated* if $\delta(n) = 1$. Finally, an integer n is said to be *very isolated* if $\delta(n)$ is "large".

It follows from (1) that the most isolated number $\leq x$ is the largest power of 2 not exceeding x, which implies in particular that $\delta(n) \leq n/2$ for all $n \geq 2$.

Remark. One might think, as a rule of thumb, that the smaller $P(n)$ is, the larger $\delta(n)$ will be, that is, that "smooth numbers have a large index of isolation". But this is not true for small values of n: for instance, $n = 11\,859\,211$ has a small $P(n)$ and nevertheless $\delta(n) = 1$, since

$$\begin{aligned} n &= 11\,859\,211 = 7 \cdot 13 \cdot 19^4, \\ n - 1 &= 11\,859\,210 = 2 \cdot 3^4 \cdot 5 \cdot 11^4. \end{aligned}$$

However, for large values of n, one can say that smooth numbers do indeed have a large index of isolation. Indeed, one can prove (see Lemma 3) that, given $B \geq 3$ fixed, there exist a constant $c = c(B) > 0$ and a number $n_0 = n_0(B)$ such that

$$\delta(n) > \frac{n}{(\log n)^c} \qquad \text{for all } B\text{-smooth integers } n \geq n_0.$$

2 Preliminary Observations and Results

It is clear that $\delta(p) = 1$ for each prime p and also that if p is odd, then $\delta(p^2) = 1$. Each of the following also holds:

$$\begin{array}{lll} \delta(2p) = 1 \ \text{ for } p \geq 5, & \delta(3p) = 1 \ \text{ for } p \geq 3, & \delta(4p) \leq 2 \ \text{ for } p \geq 5, \\ \delta(5p) = 1 \ \text{ for } p \geq 2, & \delta(6p) \leq 2 \ \text{ for } p \geq 3, & \delta(7p) \leq 2 \ \text{ for } p \geq 2, \\ \delta(8p) \leq 3 \ \text{ for } p \geq 7, & \delta(9p) \leq 3 \ \text{ for } p \geq 2, & \delta(10p) \leq 2 \ \text{ for } p \geq 2. \end{array}$$

The above are easily proven. For instance, to prove the second statement, observe that if $p \equiv 1 \pmod 4$, then $3p + 1 \equiv 0 \pmod 4$, in which case $P(3p + 1) < p$, while if $p \equiv 3 \pmod 4$, then $3p-1 \equiv 0 \pmod 4$, in which case $P(3p-1) < p$, so that in both cases $\delta(3p) = 1$.

Observe also that given any prime p, if a is an integer such that $P(a) \le p$, then $\delta(ap) \le a$ because

$$P(ap - a) = P(a(p-1)) \le \max(P(a), P(p-1)) \le p.$$

If follows from this simple observation that

$$\delta(n) = \delta(aP(n)) \le a = \frac{n}{P(n)} \quad (n \ge 2). \tag{2}$$

Moreover, one can easily show that if $P(n)^2|n$, then $\delta(n) \le \dfrac{n}{P(n)^2}$.

Definition. For each integer $n \ge 1$, let

$$\Delta(n) := \sum_{d|n} \delta(d).$$

Trivially, we have $\Delta(n) \ge \tau(n)$.

Lemma 1. *If n is a power of 2, then $\Delta(n) = n$. On the other hand, for all $n > 1$ such that $P(n) \ge 3$, we have*

$$\Delta(n) < n. \tag{3}$$

Proof. The first assertion is obvious since for each integer $\alpha \ge 1$, we have the following: $\Delta(2^\alpha) = \sum_{i=1}^{\alpha} 2^{i-1} = 2^\alpha$.

Now consider the case when n is not a power of 2. First, it is easy to show that if $P(n) = 3$, then (3) holds. Indeed, if $n = 2^\alpha \cdot 3^\beta$ for some integers $\alpha \ge 0$ and $\beta \ge 1$, then in light of (2), we have

$$\begin{aligned}
\Delta(n) &\le \sum_{d|n} \frac{d}{P(d)} = 1 + \sum_{\substack{d|n \\ P(d)=2}} \frac{d}{P(d)} + \sum_{\substack{d|n \\ P(d)=3}} \frac{d}{P(d)} \\
&= 1 + \frac{1}{2}\sum_{i=1}^{\alpha} 2^i + \frac{1}{3}\sum_{\substack{d|n \\ 3|d}} d = 1 + 2^\alpha - 1 + \frac{1}{3}(\sigma(n) - \sigma(2^\alpha)) \\
&= 2^\alpha + \frac{1}{3}\left((2^{\alpha+1}-1)\frac{3^{\beta+1}-1}{2} - (2^{\alpha+1}-1)\right) \\
&= n - \frac{3^\beta}{2} + \frac{1}{2} \le n - 1 < n,
\end{aligned}$$

which proves that (3) holds if $P(n) = 3$.

Hence, from here on, we shall assume that $P(n) \geq 5$. We shall use induction on the number of distinct prime factors of n in order to prove that

$$\sum_{d|n} \frac{d}{P(d)} < n. \tag{4}$$

First observe that the above inequality is true if $\omega(n) = 1$. Indeed, in this case, we have $n = p^b$. It is clear that

$$\begin{aligned}\sum_{d|n} \frac{d}{P(d)} &= 1 + 1 + p + p^2 + \cdots + p^{b-1} = 1 + \frac{p^b - 1}{p - 1} \\ &< 1 + p^b - 1 = p^b = n,\end{aligned}$$

which will clearly establish (3).

Let us now assume that the result holds for all n such that $\omega(n) = r - 1$ and prove that it does hold for n such that $\omega(n) = r$. Take such an integer n with k being the unique positive integer such that $P(n)^k \| n$. We then have

$$\begin{aligned}\sum_{d|n} \frac{d}{P(d)} &= \sum_{d|n/P(n)^k} \frac{d}{P(d)} + \sum_{i=1}^{k} \sum_{d|n/P(n)^k} \frac{dP(n)^i}{P(dP(n)^i)} \\ &= \sum_{d|n/P(n)^k} \frac{d}{P(d)} + \sum_{d|n/P(n)^k} d + P(n) \sum_{d|n/P(n)^k} d \\ &\qquad + \cdots + P(n)^{k-1} \sum_{d|n/P(n)^k} d \\ &= \sum_{d|n/P(n)^k} \frac{d}{P(d)} + (1 + P(n) + \cdots + P(n)^{k-1})\sigma(n/P(n)^k). \end{aligned} \tag{5}$$

Using the identity

$$\sigma\left(\frac{n}{P(n)^k}\right) = \frac{\sigma(n)}{1 + P(n) + P(n)^2 + \cdots + P(n)^k}$$

and the induction argument, it follows from (5) that

$$\begin{aligned}\sum_{d|n} \frac{d}{P(d)} &< \frac{n}{P(n)^k} + \frac{1 + P(n) + \cdots + P(n)^{k-1}}{1 + P(n) + \cdots + P(n)^k}\sigma(n) \\ &< \frac{n}{P(n)} + \frac{\sigma(n)}{P(n)}. \end{aligned} \tag{6}$$

On the other hand, it is clear that for any integer $n > 1$ such that $P(n) \geq 5$ and

$\omega(n) \geq 2$, we have

$$\frac{\sigma(n)}{n} < \prod_{p|n} \frac{p}{p-1} < \frac{3}{4}P(n).$$

Using this in (6), we obtain, since $P(n) \geq 5$, that $\sum_{d|n} \frac{d}{P(d)} < \frac{n}{5} + \frac{3}{4}n < n$, which completes the proof of (4) and thus of (3). □

Lemma 2. *The following are true:*

(i) $\#\{n \leq x : \delta(n) = 1\} \geq \frac{x}{2}$ *for all* $x \geq 2$.

(ii) $\frac{x}{2} \leq \sum_{n \leq x} \frac{1}{\delta(n)} < x$ *for all* $x \geq 4$.

(iii) $\delta(n) = \delta(m) = k \Longrightarrow |n - m| \geq k$.

(iv) $|m - n| \geq \min(\delta(m), \delta(n))$.

(v) $c_k(x) := \#\{n \leq x : \delta(n) = k\} \leq \frac{x}{k}$.

(vi) $\#\{n \leq x : \delta(n) \geq y\} \leq \frac{x}{y}$.

Proof. Since it is clear that if $\delta(n) \geq 2$, then $\delta(n-1) = \delta(n+1) = 1$, (i) follows immediately, along with (ii).

On the other hand, (iii) follows from the definition of $\delta(n)$. From (iii), we easily deduce (iv) and (v).

To prove statement (vi), we proceed as follows. Fix $1 < y \leq x$ and assume that $k := \#\{n \leq x : \delta(n) \geq y\} > x/y$. Let $y \leq n_1 < n_2 < \cdots < n_k$ be the integers $n_i \leq x$ such that $\delta(n_i) \geq y$. By the Pigeonhole Principle, there exist n_r and n_s with $1 \leq r < s \leq k$ such that $n_s - n_r < y$. Without any loss in generality, one can assume that $P(n_r) \leq P(n_s)$, in which case we have $\delta(n_s) \leq n_s - n_r < y$, a contradiction. □

Lemma 3. *Let* $B \geq 3$ *be a fixed integer. Then there exist a constant* $c = c(B) > 0$ *and a number* $n_0 = n_0(B)$ *such that*

$$\delta(n) > \frac{n}{(\log n)^c} \quad \textit{for all } B\textit{-smooth integers } n \geq n_0.$$

Proof. The result follows almost immediately from an estimate of Tijdeman [7], who showed, using the theory of logarithmic forms of Baker (see Theorem 3.1 in the book of Baker [1]), that if $n_1 < n_2 < \cdots$ represents the sequence of B-smooth numbers, then there exist positive constants $c_1(B)$ and $c_2(B)$ such that

$$\frac{n_i}{(\log n_i)^{c_1(B)}} \ll n_{i+1} - n_i \ll \frac{n_i}{(\log n_i)^{c_2(B)}},$$

where $c_2(B) \leq \pi(B) \leq c_1(B)$. Observe that Langevin [6] later provided explicit values for the constants $c_1(B)$ and $c_2(B)$. □

3 Probabilistic Results

In 1978, Erdős and Pomerance [4] showed that the lower density of those integers n for which $P(n) < P(n+1)$ (or $P(n) > P(n+1)$) is positive. Most likely, this density is $\frac{1}{2}$, but this fact remains an open problem. In 2001, Balog [2] showed that the number of integers $n \le x$ with

$$P(n-1) > P(n) > P(n+1) \tag{7}$$

is $\gg \sqrt{x}$ and observed that "undoubtedly" the density of those integers n such that (7) holds is equal to $\frac{1}{6}$.

To establish our next result, we shall make the following reasonable assumption.

Hypothesis A. *Fix an arbitrary integer $k \ge 2$ and let n be a large number. Let $a_1, a_2, \dots, a_k$ be any permutation of the numbers $0, 1, 2, \dots, k-1$. Then,*

$$\text{Prob}[P(n+a_1) < P(n+a_2) < \dots < P(n+a_k)] = \frac{1}{k!}.$$

Theorem 4. *Assuming Hypothesis A and given any integer $k \ge 1$, the expected proportion of integers n for which $\delta(n) = k$ is equal to $\frac{2}{4k^2-1}$. In particular, the proportion of non-isolated numbers is $\frac{2}{3}$.*

Proof. Fix k. Let E_k be the expected proportion of integers n for which $\delta(n) = k$. Given a large integer n, the probability that $\delta(n) > k$ is equal to the probability that

$$\min(P(n \pm 1), \dots, P(n \pm k)) > P(n).$$

Under Hypothesis A, this probability is equal to $\frac{1}{2k+1}$. This implies that

$$E_k = P(\delta(n) > k-1) - P(\delta(n) > k) = \frac{1}{2k-1} - \frac{1}{2k+1} = \frac{2}{4k^2-1},$$

which completes the proof of the theorem. □

Remark. Let $S_k(x) := \#\{n \le x : \delta(n) = k\}$ and choose $x = 10^6$. Then, we obtain the following numerical evidence.

k	1	2	3	4	5
$a = S_k(x)$	664 084	134 239	57 089	32 185	20 145
$b = [x \cdot 2/(4k^2-1)]$	666 666	133 333	57 142	31 746	20 202
a/b	0.996	1.006	0.999	1.013	0.997

Theorem 5. *Assuming Hypothesis* A*,*

$$\lim_{x\to\infty}\frac{1}{x}\sum_{n\le x}\frac{1}{\delta(n)} = 2(2\log 2-1)\approx 0.7725.$$

Proof. According to Theorem 4 (proved assuming Hypothesis A),

$$\frac{1}{x}\#\{n\le x:\delta(n)=k\} = (1+o(1))\frac{2}{4k^2-1}\quad (x\to\infty).$$

Therefore, given a fixed large integer N, we have, as $x\to\infty$,

$$\begin{aligned}\sum_{n\le x}\frac{1}{\delta(n)} &= \sum_{k=1}^{N}\frac{1}{k}\sum_{\substack{n\le x\\ \delta(n)=k}}1 + \sum_{k=N+1}^{[x/2]}\frac{1}{k}\sum_{\substack{n\le x\\ \delta(n)=k}}1\\ &= \sum_{k=1}^{N}\frac{2x}{k(4k^2-1)}(1+o(1)) + O\left(\sum_{k>N}\frac{x}{k^2}\right)\\ &= (1+o(1))xT_1(N)+O(xT_2(N)),\end{aligned}\tag{8}$$

say, where we used Lemma 2 (v). First, one can show that

$$T_1(N)=\sum_{k=1}^{N}\frac{2}{k(4k^2-1)} = 2(2\log 2-1)+O\left(\frac{1}{N}\right)\quad (N\to\infty).\tag{9}$$

To prove (9), we proceed as follows. Assume for now that $N\equiv 3 \pmod 4$. Then, using the estimate $\sum_{k=1}^{N}\frac{1}{k} = \log N+\gamma+O(1/N)$ as $N\to\infty$ (where γ is Euler's constant), we have

$$\begin{aligned}\frac{1}{2}T_1(N) &= -\sum_{k=1}^{N}\frac{1}{k}+\sum_{k=1}^{N}\frac{1}{2k-1}+\sum_{k=1}^{N}\frac{1}{2k+1}\\ &= -\frac{1}{2}+\frac{1}{3}-\cdots+\frac{1}{N}\\ &\qquad +2\left(\frac{1}{N+2}+\frac{1}{N+4}+\cdots+\frac{1}{2N-1}\right)+\frac{1}{2N+1}\\ &= \log 2-1+O(1/N)+2\sum_{j=(N+1)/2}^{N-1}\frac{1}{2j+1}+\frac{1}{2N+1}\\ &= \log 2-1+O(1/N)+\log 2+O(1/N)\\ &= 2\log 2-1+O(1/N),\end{aligned}$$

which proves (9). A similar argument holds if $N\equiv 0, 1$ or $2 \pmod 4$, thus establishing (9).

On the other hand,

$$\sum_{k>N} \frac{1}{k^2} < \int_N^\infty \frac{1}{t^2}\,dt = \frac{1}{N},$$

so that

$$T_2(N) < \frac{1}{N}. \tag{10}$$

Now let $\varepsilon > 0$ be arbitrarily small and let $N = [1/\varepsilon] + 1$. We then have, using (9) and (10) in (8),

$$\sum_{n\le x} \frac{1}{\delta(n)} = 2(2\log 2 - 1)x + O(\varepsilon x) + O(\varepsilon x) \quad (x \to \infty),$$

which completes the proof of the theorem. □

Remark. Using a computer, one obtains that

$$\frac{1}{10^9} \sum_{n\le 10^9} \frac{1}{\delta(n)} = 0.7719\ldots\ .$$

4 The Isolation Index with Respect to a Given Function

The function δ can also be defined relatively to any real-valued arithmetic function f taking its minimal value at $f(1)$ as

$$\delta_f(n) := \min_{\substack{1\le m\ne n\\ f(m)\le f(n)}} |n-m| \quad (n \ge 2)$$

with $\delta_f(1) = 1$.

Examples. • Let $f(n)$ be any monotonic function. Then we have $\delta_f(n) = 1$ for all $n \ge 2$.

• Let $f(n) = \omega(n) := \sum_{p|n} 1$. Then, the first 40 values of $\delta_\omega(n)$ are

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 2, 1, 1, 1, 1, 1, 2, 1,
1, 1, 2, 1, 2, 1, 2, 1, 2, 1, 1, 1, 1, 1, 1, 1, 4, 1, 1, 1.

• Let $f(n) = \Omega(n) := \sum_{p^\alpha \| n} \alpha$. Then, the first 40 values of $\delta_\Omega(n)$ are

1, 1, 1, 1, 2, 1, 2, 1, 1, 1, 2, 1, 2, 1, 1, 1, 2, 1, 2, 1,
1, 1, 4, 1, 1, 1, 1, 1, 2, 1, 2, 1, 1, 1, 1, 1, 4, 1, 1, 1.

- Let $f(n) = \tau(n) := \sum_{d|n} 1$. Then, the first 40 values of $\delta_\tau(n)$ are

$$1,\ 1,\ 1,\ 1,\ 2,\ 1,\ 2,\ 1,\ 2,\ 1,\ 2,\ 1,\ 2,\ 1,\ 1,\ 1,\ 2,\ 1,\ 2,\ 1,$$
$$1,\ 1,\ 4,\ 1,\ 2,\ 1,\ 1,\ 1,\ 2,\ 1,\ 2,\ 1,\ 1,\ 1,\ 1,\ 1,\ 4,\ 1,\ 1,\ 1.$$

Remark 6. It turns out that Hypothesis A holds unconditionally when one replaces the function $P(n)$ by the function $\omega(n)$ or $\Omega(n)$ or $\tau(n)$, and therefore that the equivalent of Theorem 4 for either of these three functions is true without any conditions, that is, that for any fixed positive integer k,

$$\frac{1}{x}\#\{n \le x : \delta_\omega(n) = k\} = \frac{2}{4k^2 - 1} + o(1) \quad (x \to \infty),$$

the same being true for $\Omega(n)$ or $\tau(n)$ in place of $\omega(n)$.

To prove our claim, we first need to prove the following two propositions.

Proposition 7. *Let $a_1, a_2, \ldots, a_k$ be any distinct integers and let $z_1, z_2, \ldots, z_k$ be arbitrary real numbers. Then,*

$$\lim_{x\to\infty} \frac{1}{x}\#\left\{n \le x : \frac{\omega(n + a_j) - \log\log n}{\sqrt{\log\log n}} < z_j,\ 1 \le j \le k\right\} = \prod_{1\le j\le k} \Phi(z_j), \tag{11}$$

where $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt$.

Proof. Given a real-valued vector $(t_1, t_2, \ldots, t_k)$, consider the function

$$H(n) = \sum_{j=1}^{k} t_j \omega(n + a_j).$$

We will now apply Proposition 2 and Theorem 1 of Granville and Soundararajan [5], where instead of considering the function

$$f_p(n) = \begin{cases} 1 - \frac{1}{p} & \text{if } p|n, \\ -\frac{1}{p} & \text{otherwise,} \end{cases}$$

we use the function

$$f_p(n) = \begin{cases} t_r - \frac{1}{p}\sum_{j=1}^{k} t_j & \text{if } p|(n + a_r), \\ -\frac{1}{p}\sum_{j=1}^{k} t_j & \text{otherwise,} \end{cases}$$

where it is clear that, except for a finite number of primes p, each prime p divides $n + a_r$ for at most one a_r.

Using this newly defined function $f_p(n)$ and following exactly the same steps as in the proof of Proposition 2 of Granville and Soundararajan, we obtain that

$$\lim_{x\to\infty}\frac{1}{x}\#\left\{n\le x : \frac{H(n)-\sum_{j=1}^{k} t_j \log\log n}{\sqrt{\log\log n}} < z\right\} = \Phi\left(\frac{z}{\sqrt{\sum_{j=1}^{k} t_j^2}}\right). \quad (12)$$

In other words, $H(n)$ has a Gaussian distribution with mean value

$$\sum_{j=1}^{k} t_j \log\log n$$

and standard deviation

$$\sqrt{\sum_{j=1}^{k} t_j^2 \cdot \log\log n}.$$

Because of the moments of a Gaussian distribution, statement (12) is equivalent to

$$\begin{aligned}\lim_{x\to\infty}\frac{1}{x}\sum_{n\le x}&\left(\frac{H(n)-\sum_{j=1}^{k} t_j \log\log n}{\sqrt{\log\log n}}\right)^m \\ &= G(m)\left(\sum_{j=1}^{k} t_j^2\right)^{m/2} \qquad (m=1,2,\dots),\end{aligned} \quad (13)$$

where, for each positive integer m,

$$G(m) = \begin{cases}\prod_{1\le j\le m/2}(2j-1) & \text{if } m \text{ is odd,}\\ 0 & \text{if } m \text{ is even.}\end{cases}$$

By expanding the left-hand side of (13) using the Multinomial Theorem, we may rewrite it (for each positive integer m) as

$$\sum_{\substack{0\le u_i\le m,\ i=1,\dots,k\\ u_1+\cdots+u_k=m}} \lim_{x\to\infty}\frac{1}{x}\sum_{n\le x}\frac{m!}{u_1!\cdots u_k!}\prod_{1\le j\le k} t_j^{u_j}\left(\frac{\omega(n+a_j)-\log\log n}{\sqrt{\log\log n}}\right)^{u_j}. \quad (14)$$

By considering (14) as a function of $t_1, t_2, \dots, t_k$ and comparing the coefficients with those on the right-hand side of (13), we obtain, for each positive integer m,

$$\begin{aligned}\frac{m!}{u_1!\cdots u_k!}\lim_{x\to\infty}\frac{1}{x}\sum_{n\le x}\prod_{1\le j\le k}&\left(\frac{\omega(n+a_j)-\log\log n}{\sqrt{\log\log x}}\right)^{u_j}\\ &= G(m)\frac{(m/2)!}{(u_1/2)!\cdots(u_k/2)!}\end{aligned}$$

or equivalently

$$\lim_{x\to\infty} \frac{1}{x} \sum_{n\le x} \prod_{1\le j\le k} \left(\frac{\omega(n+a_j) - \log\log n}{\sqrt{\log\log x}} \right)^{u_j} = \prod_{1\le j\le k} G(u_j). \tag{15}$$

Since the right-hand side of (15) corresponds to the centered moments of a multivariate independent Gaussian distribution, the validity of (11) follows, thereby completing the proof. □

Proposition 8. *Let g stand for any of the functions ω, Ω or τ. Let $a_1, a_2, \ldots, a_k$ be any permutation of the integers $0, 1, \ldots, k-1$. Then,*

$$\lim_{x\to\infty} \frac{1}{x} \#\{n \le x : g(n+a_j) < g(n+a_{j+1}),\ j = 1, \ldots, k-1\} = \frac{1}{k!}.$$

Proof. In the case $g = \omega$, the result follows from Proposition 7, namely by simple integration of (11). As for $g = \Omega$, observe that it is easy to show that

$$\lim_{K\to\infty} \lim_{x\to\infty} \frac{1}{x} \#\{n \le x : \Omega(n) - \omega(n) > K\} = 0. \tag{16}$$

Hence, using (16) and integrating (11), Proposition 7 holds for $g = \Omega$. Finally, as Proposition 7 holds for $g = \omega$ and $g = \Omega$, the inequality $2^{\omega(n)} \le \tau(n) \le 2^{\Omega(n)}$ implies that it also holds for $g = \tau$, thus completing the proof. □

Theorem 9. *Let $f(n) = \omega(n)$ or $\Omega(n)$ or $\tau(n)$. Then,*

$$\lim_{x\to\infty} \frac{1}{x} \sum_{n\le x} \frac{1}{\delta_f(n)} = 2(2\log 2 - 1).$$

Proof. In light of Remark 6 and of Proposition 8, the result is immediate. □

Theorem 10. *Let $a < b$ be positive integers. For any real-valued arithmetic function f and any interval $I = [a, b[$ of length $N = b - a$,*

$$\sum_{n\in I} \frac{1}{\delta_f(n)} \ge \frac{2}{3}N - \frac{2}{3}.$$

Proof. We conduct the proof using induction on N. First observe that Theorem 10 holds for small values of N. For instance, if $N = 1$,

$$\sum_{n\in I} \frac{1}{\delta_f(n)} = \frac{1}{\delta_f(a)} > 0 = \frac{2}{3} - \frac{2}{3}.$$

If $N = 2$, then since at least one of $\delta_f(n)$ and $\delta_f(n+1)$ must be 1, it follows that

$$\sum_{n\in I} \frac{1}{\delta_f(n)} = \frac{1}{\delta_f(a)} + \frac{1}{\delta_f(a+1)} > 1 > \frac{2}{3} = \frac{2}{3}\cdot 2 - \frac{2}{3}.$$

We will now assume that the result holds for every integer smaller than N and prove that it must therefore hold for N. We shall do this by distinguishing three possible cases:

(i) either $\delta_f(a) = 1$ or $\delta_f(b-1) = 1$;

(ii) case (i) is not satisfied and there exists a positive integer $k \in]a, b-2[$ such that both $\delta_f(k) = 1$ and $\delta_f(k+1) = 1$;

(iii) neither of the two previous cases holds.

In case (i), we can assume without any loss of generality that $\delta_f(a) = 1$, in which case

$$\sum_{n\in I} \frac{1}{\delta_f(n)} = 1 + \sum_{a+1\le n<b} \frac{1}{\delta_f(n)}.$$

Using our induction hypothesis, we get that

$$1 + \sum_{a+1\le n<b} \frac{1}{\delta_f(n)} \ge 1 + \frac{2}{3}(N-1) - \frac{2}{3} = \frac{2}{3}N - \frac{1}{3} > \frac{2}{3}N - \frac{2}{3},$$

proving the theorem in case (i).

Suppose now that case (ii) is satisfied. Then, there is an integer $k \in]a, b-2[$ such that $\delta_f(k) = \delta_f(k+1) = 1$, in which case

$$\sum_{n\in I} \frac{1}{\delta_f(n)} = \sum_{a\le n<k} \frac{1}{\delta_f(n)} + 2 + \sum_{k+2\le n<b} \frac{1}{\delta_f(n)}.$$

Again, using our induction hypothesis we have that the right-hand side is larger or equal to

$$\frac{2}{3}(k-a) - \frac{2}{3} + 2 + \frac{2}{3}(b-k-2) - \frac{2}{3} = \frac{2}{3}N - \frac{2}{3},$$

proving the theorem in case (ii).

We now consider case (iii). In this situation, N has to be odd, since the sum starts with the term $1/\delta_f(a) < 1$ and ends with the term $1/\delta_f(b-1) < 1$, because every second value of $\delta_f(n)$ must be 1. Assume that a is odd, in which case $b-1$ is odd. The case a and $b-1$ even can be treated in a similar way. Hence, it is at odd integers n that $\delta_f(n) > 1$, in which case we must have both $f(n-1) > f(n)$ and $f(n+1) > f(n)$. Recall the definition

$$\delta_f(n) := \min_{\substack{1 \le m \ne n \\ f(m) \le f(n)}} |m-n|;$$

now since the integer m at which this minimum occurs must be odd (since for the other integers m, the even ones, we have $\delta_f(m) = 1$), it follows that for an odd $n = 2j+1$, we have

$$\delta_f(2j+1) = 2 \min_{\substack{1 \le k \ne j \\ f(2k+1) \le f(2j+1)}} |k-j|,$$

so that we may write

$$\begin{aligned}
\sum_{\substack{n \in [a,b[\\ n \text{ odd}}} \frac{1}{\delta_f(n)} &= \sum_{j \in [\frac{a-1}{2}, \frac{b}{2}[} \frac{1}{\delta_f(2j+1)} \\
&= \frac{1}{2} \sum_{j \in [\frac{a-1}{2}, \frac{b}{2}[} \frac{1}{\min\limits_{\substack{1 \le k \ne j \\ f(2k+1) \le f(2j+1)}} |k-j|} \\
&= \frac{1}{2} \sum_{j \in [\frac{a-1}{2}, \frac{b}{2}[} \frac{1}{\min\limits_{\substack{1 \le k \ne j \\ g(k) \le g(j)}} |k-j|} \\
&= \frac{1}{2} \sum_{j \in [\frac{a-1}{2}, \frac{b}{2}[} \frac{1}{\delta_g(j)},
\end{aligned} \tag{17}$$

where we have set $g(j) := f(2j+1)$. Now since the interval $[\frac{a-1}{2}, \frac{b}{2}[$ is of length $\frac{N+1}{2} < N$, we can apply our induction hypothesis and write that the last expression in (17) is no larger than

$$\frac{1}{2} \cdot \frac{2}{3} \left(\frac{N+1}{2} - 1 \right).$$

It follows from this that

$$\sum_{n \in I} \frac{1}{\delta_f(n)} \ge \frac{N-1}{2} + \frac{1}{2} \cdot \frac{2}{3} \left(\frac{N+1}{2} - 1 \right) = \frac{2}{3}(N-1),$$

as needed to be proved. This completes the proof of the theorem. □

In the statement of Theorem 10, is there any hope that one could replace the constant $\frac{2}{3}$ by a larger one? The answer is 'no', as the following result shows.

Theorem 11. *Let $\alpha_{0,q}(n), \alpha_{1,q}(n), \alpha_{2,q}(n), \ldots, \alpha_{k,q}(n)$ be the digits of n when written in base q, that is,*

$$n = \sum_{j=0}^{k} \alpha_{j,q}(n) q^j.$$

Then the function $f = g_q$ defined by

$$g_q(n) = \sum_{j=0}^{k} \alpha_{j,q}(n) q^{k-j} \tag{18}$$

has the following property:

$$\lim_{x\to\infty} \frac{1}{x} \sum_{n\le x} \frac{1}{\delta_f(n)} = \frac{q}{q+1}. \tag{19}$$

Remark. Observe that $g_q(n)$ is the number obtained by writing the basis q digits of n in reverse order. In a sense, the result claims that the function g_2 is the one that provides the minimal value for the sum of the reciprocals of the index of isolation.

Remark 12. Let n be written in basis $q \ge 2$, that is,

$$n := \sum_{j=0}^{k} \alpha_{j,q}(n) q^j.$$

Let m be the smallest integer such that $\alpha_{m,q}(n)$ is greater than zero. Then, under the assumption that n is not a perfect power of q, it is easy to verify that

$$\delta_{g_q}(n) = q^m.$$

On the other hand, if n is a perfect power of q, say $n = q^k$, we have

$$\delta_{g_q}(n) = q^{k-1}(q-1).$$

Proof of Theorem 11. We shall only consider the case $q = 2$, since the general case can be treated similarly. We observe that $\delta_{g_2}(n) = 1$ if and only if n is odd. More generally, in light of Remark 12, we have

$$\delta_{g_2}(n) = 2^k \text{ if and only if } \frac{n}{2^k} \text{ is an odd integer} \quad (k \ge 0).$$

We can therefore write

$$\sum_{n\le x}\frac{1}{\delta_{g_2}(n)}=\sum_{k=0}^{[\log x/\log 2]}\frac{1}{2^k}\cdot\#\left\{n\le x:\frac{n}{2^k}\equiv 1\pmod 2\right\}.$$

It is easy to see that

$$\#\left\{n\le x:\frac{n}{2^k}\equiv 1\pmod 2\right\}=\frac{x}{2^{k+1}}+O(1),$$

so that

$$\sum_{n\le x}\frac{1}{\delta_f(n)}=\sum_{k=0}^{[\log x/\log 2]}\frac{x}{2^{2k+1}}+O\left(\sum_{k\ge 0}\frac{1}{2^k}\right)=\frac{2}{3}x+O(1),$$

which proves (19) in the case $q=2$, thus completing the proof of the theorem. □

Theorem 13. *For real numbers y, w such that $\frac{2}{3}\le y\le w\le 1$, there exists an arithmetic function f such that*

$$\liminf_{x\to\infty}\frac{1}{x}\sum_{n\le x}\frac{1}{\delta_f(n)}=y\quad\textit{and}\quad\limsup_{x\to\infty}\frac{1}{x}\sum_{n\le x}\frac{1}{\delta_f(n)}=w. \tag{20}$$

Proof. The function $1/\delta_{g_2}$ has a mean value of $\frac{2}{3}$, while it is clear that the mean value of the reciprocal of the index of isolation of any monotone function h is 1. We shall construct a function that behaves piecewise like g_2 and piecewise like h so that the mean value of the reciprocal of its isolation index will be a pondered mean of the values $\frac{2}{3}$ and 1. We first define real numbers $s,t\in[0,1]$ in such a way that $s+(1-s)\frac{2}{3}=y$ and $t+(1-t)\frac{2}{3}=w$. Now consider the intervals

$$I_j:=[2^{2^j},2^{2^{j+1}}[\quad(j=1,2,3,\dots).$$

For j even and for each integer $m\in[1,2^{2^j-j}(2^{2^j}-1)]$, define the families of subintervals $K_{j,m}$ and $L_{j,m}$ as follows:

$$K_{j,m}:=[2^{2^j}+(m-1)2^j,2^{2^j}+(m-1+s)\cdot 2^j[$$

and

$$L_{j,m}:=[2^{2^j}+(m-1+s)2^j,2^{2^j}+m2^j[,$$

so that

$$I_j=\bigcup_{m=1}^{2^{2^j-j}(2^{2^j}-1)}(K_{j,m}\cup L_{j,m}).$$

For j odd, we replace the number s by t in the definition of the subintervals $K_{j,m}$ and $L_{j,m}$. For simplicity, we define implicitly $A_{j,m}, B_{j,m}, C_{j,m}, D_{j,m}$ as

$$K_{j,m} = [A_{j,m}, B_{j,m}[$$

and

$$L_{j,m} = [C_{j,m}, D_{j,m}[\,.$$

We are now ready to define a function f satisfying (20). We define f piecewise in the following manner. If $n \in K_{j,m}$, then set

$$f(n) = n - A_{j,m},$$

while if $n \in L_{j,m}$, set

$$f(n) = g_2(n - C_{j,m}).$$

Assume first that j is even. Then,

$$\begin{aligned}
\sum_{n\in I_j} \frac{1}{\delta_f(n)} &= \sum_m \sum_{n\in K_{j,m}} \frac{1}{\delta_f(n)} + \sum_m \sum_{n\in L_{j,m}} \frac{1}{\delta_f(n)} \\
&= \sum_m \left(B_{j,m} - A_{j,m} + O(1)\right) + \sum_m \frac{2}{3}\left(D_{j,m} - C_{j,m} + O(1)\right) \\
&= \sum_m \left(s2^j + O(1)\right) + \sum_m \left(\frac{2}{3}(1-s)2^j + O(1)\right) \\
&= \sum_m \left(y2^j + O(1)\right) \\
&= y\left(2^{2^{j+1}} - 2^{2^j}\right) + O\left(\frac{2^{2^{j+1}}}{2^j}\right). \qquad (21)
\end{aligned}$$

If j is odd, we obtain in a similar fashion

$$\sum_{n\in I_j} \frac{1}{\delta_f(n)} = w\left(2^{2^{j+1}} - 2^{2^j}\right) + O\left(\frac{2^{2^{j+1}}}{2^j}\right). \qquad (22)$$

Let x be a large real number. Let j^* be the largest integer such that $2^{2^{j^*}} < x$ and let m^* be the largest integer such that $D_{j^*,m^*} \le x$. We then have

$$\begin{aligned}
\sum_{n\le x} \frac{1}{\delta_f(n)} &= \sum_{n<2^{2^{j^*-1}}} \frac{1}{\delta_f(n)} + \sum_{n\in I_{j^*-1}} \frac{1}{\delta_f(n)} + \sum_{2^{2^{j^*}}\le n\le x} \frac{1}{\delta_f(n)} \\
&= \Sigma_1 + \Sigma_2 + \Sigma_3, \qquad (23)
\end{aligned}$$

say. On the one hand, we trivially get

$$\Sigma_1 = O\left(2^{2^{j^*-1}}\right), \tag{24}$$

while assuming $j^* - 1$ even, we obtain, using (21),

$$\Sigma_2 = y \cdot 2^{2^{j^*}} + O\left(\frac{2^{2^{j^*}}}{2^{j^*}}\right). \tag{25}$$

Finally,

$$\Sigma_3 = \sum_{m \le m^*} \sum_{n \in L_{j^*,m}} \frac{1}{\delta_f(n)} + \sum_{m \le m^*} \sum_{n \in K_{j^*,m}} \frac{1}{\delta_f(n)} + O\left(2^{j^*}\right),$$

which yields, in light of (22) (since j^* is odd),

$$\Sigma_3 = w\left(x - 2^{2^{j^*}}\right) + O\left(2^{j^*} + m^*\right). \tag{26}$$

Using (24), (25) and (26) in (23), we get

$$\sum_{n \le x} \frac{1}{\delta_f(n)} = y 2^{2^{j^*}} + w\left(x - 2^{2^{j^*}}\right) + O\left(\frac{x}{\log x}\right),$$

which completes the proof of the theorem. □

5 The Mean Value of the Index of Isolation

While the mean value of the reciprocal of the index of isolation gives information on the local behavior of a function f, the mean value of the index of isolation itself gives information on the very isolated numbers.

Theorem 14. *Let f be a real-valued arithmetic function. Let a and b be two positive integers such that $b - a = N$. Suppose furthermore that for all $m \in [a, b[$, $f(m) \ge f(a)$. Then*

$$\sum_{n \in]a,b[} \delta_f(n) \le \frac{N \log N}{2 \log 2}. \tag{27}$$

Proof. We prove Theorem 14 by induction on N. The result holds for $N = 1$, because the left-hand side of (27) is 0 (since the interval of summation contains no integers), yielding the inequality $0 \le 0$. It also holds for $N = 2$, because in this case we have $\delta_f(a + 1) = 1$, yielding the inequality $1 \le 1$. So, let us assume that

(27) is true for all intervals $]a, b[$ with $a - b = N$ for some integer $N > 1$. We shall prove that under the hypothesis that $b - a = N + 1$ and that for all $m \in]a, b[$, $f(m) \geq f(a)$ and $f(m) \geq f(b)$, we have

$$\sum_{a+1 \leq n \leq b-1} \delta_f(n) \leq \frac{N}{2} \frac{\log N}{\log 2}. \tag{28}$$

We prove (28) by using induction on N. Clearly, the result is true for $N = 1$ and $N = 2$. Assume that it is true for $N - 1$ and let us prove that it is true for N. Let n^* be an integer such that $n^* \in]a, b[$ and $f(m) \geq f(n^*)$ for all $m \in]a, b[$. We thus have

$$\sum_{a+1 \leq n \leq b-1} \delta_f(n) \leq \min(n^* - a - 1, b - n^* - 1) + \sum_{a+1 \leq n \leq n^*-1} \delta_f(n) + \sum_{n^*+1 \leq n \leq b-1} \delta_f(n).$$

Using our induction hypothesis, we have

$$\sum_{a+1 \leq n \leq n^*-1} \delta_f(n) \leq \frac{n^* - a}{2} \frac{\log(n^* - a)}{\log 2}$$

and

$$\sum_{n^*+1 \leq n \leq b-1} \delta_f(n) \leq \frac{b - n^*}{2} \frac{\log(b - n^*)}{\log 2}.$$

Without any loss of generality, we can assume that $n^* \in]a, a + (b - a)/2]$, so that

$$\sum_{a+1 \leq n \leq b-1} \delta_f(n) \leq n^* - a - 1 + \frac{n^* - a}{2} \frac{\log(n^* - a)}{\log 2} + \frac{b - n^*}{2} \frac{\log(b - n^*)}{\log 2}. \tag{29}$$

Assuming for now that n^* is a real variable, and taking the derivative of the right-hand side of (29) with respect to n^*, we obtain

$$1 + \frac{\log(n^* - a)}{2 \log 2} - \frac{\log(b - n^*)}{2 \log 2}.$$

Since the second derivative is positive, the right-hand side of (29) reaches its maximum value at the end points, that is, either when $n^* = a + 1$ or $n^* = (b + a)/2$. In the first case, we get

$$\sum_{a+1 \leq n \leq b-1} \delta_f(n) \leq \frac{(N - 1) \log(N - 1)}{2 \log 2} \leq \frac{N \log N}{2 \log 2}.$$

In the second case, that is, when $n^* = (b+a)/2$, we obtain

$$\sum_{a+1 \le n \le b-1} \delta_f(n) \le \frac{N}{2} - 1 + \frac{N \log(N/2)}{2\log 2} = \frac{N \log N}{2 \log 2} - 1 \le \frac{N \log N}{2 \log 2},$$

thus proving (28) in all cases.

We are now ready to complete the proof of Theorem 14, that is, to remove the condition $f(m) \ge f(b)$. For this, we use induction.

Let $n_0 \in]a, b[$ be an integer such that for all $m \in]a, b[$, $f(m) \ge f(n_0)$. We can write

$$\sum_{a+1 \le n \le b-1} \delta_f(n) \le \sum_{a+1 \le n \le n_0 - 1} \delta_f(n) + n_0 - a - 1 + \sum_{n_0+1 \le n \le b-1} \delta_f(n).$$

Using our induction hypothesis, we obtain

$$\sum_{a+1 \le n \le n_0-1} \delta_f(n) \le \frac{(n_0 - a)\log(n_0 - a)}{2 \log 2}$$

and

$$\sum_{n_0+1 \le n \le b-1} \delta_f(n) \le \frac{(b - n_0)\log(b - n_0)}{2 \log 2}.$$

From the last three estimates, it follows that

$$\sum_{a+1 \le n \le b-1} \delta_f(n) \le n_0 - a - 1 + \frac{(n_0 - a)\log(n_0 - a)}{2 \log 2} + \frac{(b - n_0)\log(b - n_0)}{2 \log 2}. \tag{30}$$

Proceeding as we did to estimate the right-hand side of (29), we obtain that the right-hand side of formula (30) is less than $\frac{N \log N}{2 \log 2}$, which completes the proof of the theorem. □

Theorem 15. *As* $x \to \infty$,

$$\sum_{n \le x} \delta_\omega(n) \ll x \log\log x.$$

Proof. For each $x \ge 2$,

$$\begin{aligned} \sum_{n \le x} \delta_\omega(n) &= \sum_{d} \sum_{\substack{n \le x \\ \omega(n) = d}} \delta_\omega(n) \\ &= \sum_{d \le 10 \log\log x} \sum_{\substack{n \le x \\ \omega(n) = d}} \delta_\omega(n) + \sum_{d > 10 \log\log x} \sum_{\substack{n \le x \\ \omega(n) = d}} \delta_\omega(n). \end{aligned} \tag{31}$$

Clearly, for any fixed $d \geq 1$,

$$\sum_{\substack{n \leq x \\ \omega(n)=d}} \delta_\omega(n) \leq 2x.$$

Therefore,

$$\sum_{d \leq 10 \log\log x} \sum_{\substack{n \leq x \\ \omega(n)=d}} \delta_\omega(n) \leq 20x \log\log x. \tag{32}$$

Let x be large and fixed, and consider the set $S := \{n \leq x : \omega(n) > 10 \log\log x\}$. Write S as the union of disjoint intervals $S = I_1 \cup I_2 \cup \cdots \cup I_k$, and let ℓ_j stand for the length of the interval I_j. We have from Theorem 14 that

$$\sum_{n \in I_j} \delta_\omega(n) \leq \frac{\ell_j \log \ell_j}{2 \log 2}. \tag{33}$$

On the other hand, one can show that

$$\sum_{j=1}^{k} \ell_j = \#S = \sum_{\substack{n \leq x \\ \omega(n) > 10 \log\log x}} 1 = o\left(\frac{x}{\log x}\right) \quad (x \to \infty) \tag{34}$$

(see, for instance, relation (18) in De Koninck, Doyon and Luca [3]).

It follows from (33) and (34) that

$$\sum_{n \in S} \delta_\omega(n) \leq \sum_{j=1}^{k} \frac{\ell_j \log \ell_j}{2 \log 2} \leq \frac{\log x}{2 \log 2} \cdot o\left(\frac{x}{\log x}\right) = o(x) \quad (x \to \infty). \tag{35}$$

Substituting (32) and (35) in (31) completes the proof of the theorem. □

Remark. Using a computer, one can observe that, for $x = 10^9$,

$$\frac{1}{x \log\log x} \sum_{n \leq x} \delta_\omega(n) \approx 0.60.$$

Theorem 16. *Let the function g_q be defined as in* (18). *Then*

$$\sum_{1 \leq n \leq N} \delta_{g_q}(n) = \frac{(q-1) N \log N}{q \log q} + O(N),$$

so that the function g_2 is the function for which the sums of the index of isolation is maximal.

Proof. Let n be written in base $q \geq 2$, that is,

$$n := \sum_{j=0}^{k} \alpha_{j,q}(n) q^j.$$

In light of Remark 12, we get, letting k_0 be the largest integer such that $q^{k_0} \leq N$,

$$\begin{aligned}
\sum_{n \leq N} \delta_{g_q}(n) &= \sum_{\substack{n \leq N \\ n \neq q^k}} \delta_{g_q}(n) + \sum_{\substack{n \leq N \\ n = q^k}} \delta_{g_q}(n) \\
&= \sum_{m \leq \log N / \log q} q^m \cdot \#\{n \leq N : \delta_{g_q}(n) = q^m\} + \sum_{q^k \leq N} q^{k-1}(q-1) \\
&= \sum_{m \leq \log N / \log q} q^m \cdot \left(\frac{q-1}{q^{m+1}} N + O(1) \right) + (q-1) \sum_{q^k \leq N} q^{k-1} \\
&= \frac{q-1}{q} N \frac{\log N}{\log q} + O(N) + (q-1) \frac{q^{k_0} - 1}{q-1} \\
&= \frac{q-1}{q} N \frac{\log N}{\log q} + O(N),
\end{aligned}$$

thus completing the proof of the theorem. □

6 Computational Data and Open Problems

If $n = n_k$ stands for the smallest positive integer n such that

$$\delta(n) = \delta(n+1) = \cdots = \delta(n+k-1) = 1, \tag{36}$$

then we have the following table:

k	1	2	3	4	5
n_k	1	1	1	91	91

k	6	7	8	9	10
n_k	169	2737	26 536	67 311	535 591

k	11	12	13	14	15
n_k	3 021 151	26 817 437	74 877 777	657 240 658	785 211 337

Some open problems concerning the sequence n_k, $k = 1, 2, 3, \ldots$, are the following:

(i) Prove that n_k exists for each integer $k \geq 16$.

(ii) Estimate the size of n_k as a function of k. Also, is it true that $n_k \leq k!$ for each integer $k \geq 5$?

(iii) Prove that for any fixed $k \geq 3$, there are infinitely many integers n such that (36) is satisfied. The fact that the matter is settled for $k = 2$ follows immediately from the Balog result stated at the beginning of Section 3.

Interesting questions also arise from the study of the function $\Delta(n)$ first mentioned in Section 2. For instance, let m_k stand for the smallest number m for which

$$\Delta(m) = \Delta(m+1) = \cdots = \Delta(m+k-1). \tag{37}$$

Then

- $m_2 = 14$, with $\Delta(14) = \Delta(15) = 4$;
- $m_3 = 33$, with $\Delta(33) = \Delta(34) = \Delta(35) = 4$;
- $m_4 = 2\,189\,815$, with $\Delta(m_4 + i) = 12$ for $i = 0, 1, 2, 3$;
- $m_5 = 7\,201\,674$, with $\Delta(m_5 + i) = 14$ for $i = 0, 1, 2, 3, 4$;
- if m_6 exists, then $m_6 > 1\,500\,000\,000$.

Specific questions are the following:

(i) Prove that m_k exists for each integer $k \geq 6$.

(ii) Estimate the size of m_k as a function of k.

(iii) Prove that for any fixed $k \geq 3$, there are infinitely many integers m such that (37) is satisfied.

Acknowledgments. The authors would like to thank the referee for some very helpful suggestions, which led to the improvement of this paper.

Bibliography

[1] A. Baker, *Transcendental Number Theory*, Cambridge University Press, London, 1975.

[2] A. Balog, On triplets with descending largest prime factors, *Studia Sci. Math. Hungar.* **38** (2001), 45–50.

[3] J. M. De Koninck, N. Doyon and F. Luca, Sur la quantité de nombres économiques, *Acta Arithmetica* **127** (2007), no. 2, 125–143.

[4] P. Erdős and C. Pomerance, *On the largest prime factors of n and $n+1$*, *Aequationes Math.* **17** (1978), 311–321.

[5] A. Granville and K. Soundararajan, Sieving and the Erdős–Kac theorem, in: *Equidistribution in Number Theory. An Introduction* (Montreal, Canada, July 11–22, 2005), pp. 15–27, NATO Science Series II: Mathematics, Physics and Chemistry 237, Springer-Verlag, Amsterdam , 2007.

[6] M. Langevin, Quelques applications de nouveaux résultats de Van der Poorten, *Sém. Delange-Pisot-Poitou* **17** (1975/76), No. G12, 11 pp.

[7] R. Tijdeman, On the maximal distance between integers composed of small primes, *Compos. Math.* **28** (1974), 159–162.

Received March 21, 2010; revised November 29, 2010; accepted February 7, 2011.

Author information

Jean-Marie De Koninck, Département de mathématiques et de statistique,
Université Laval, Québec, Québec G1V 0A6, Canada.
E-mail: jmdk@mat.ulaval.ca

Nicolas Doyon, Département de mathématiques et de statistique,
Université Laval, Québec, Québec G1V 0A6, Canada.
E-mail: nicolas.doyon@mat.ulaval.ca