

ON THE DISTRIBUTION OF THE NUMBER OF DIGITS NEEDED TO WRITE THE FACTORIZATION OF AN INTEGER

J.-M. DeKoninck, N. Doyon and P. Letendre
(Québec, Canada)

*Dedicated to Professor Imre Kátai
on the occasion of his 70th birthday*

Abstract. Let $F_q(n)$ be the number of digits needed to write the factorization of n in base q . Several authors have studied the cardinality of the set of economical numbers, that is those integers n for which $F_q(n) \leq \left\lfloor \frac{\log n}{\log q} \right\rfloor + 1$. The fact that the set of economical numbers is of zero density in the set of integers reveals nothing about the *normal* behavior of $F_q(n)$. In this note we study the central distribution of the function $F_q(n)$ and show that it is Gaussian.

1. Introduction and notations

Let $F_q(n)$ be the number of digits needed to write the factorization of n in base q . For example, $F_{10}(125) = F_{10}(5^3) = 2$ and $F_{10}(30) = F_{10}(2 \cdot 3 \cdot 5) = 3$. In 1995, Santos [7] introduced the notation of *economical numbers* in base q , $q \geq 2$, namely those integers n for which $F_q(n) \leq \left\lfloor \frac{\log n}{\log q} \right\rfloor + 1$, meaning that the number of digits needed to write the factorization of n is smaller or equal to the number of digits appearing in its digital expansion in base q . Since then, several authors have studied the counting function of economical numbers, in particular De Koninck and Luca [3], [4], and more recently De Koninck, Doyon and Luca [5]. Here, for a fixed $q \geq 2$, we study

the distribution function $H_q(x, y) := \#\{n < x : F_q(n) < y\}$ and more precisely the case where $y = y(x, c) = \frac{\log x}{\log q} + \frac{1}{2} \log \log x + c\sqrt{\log \log x}$. We show that in this case, the expression $G(c) = \lim_{x \rightarrow \infty} \frac{1}{x} H_q(x, y)$ is well defined and that $G(c) = \Phi(\sqrt{3}c)$, where $\Phi(y) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{t^2}{2}} dt$ is the distribution function of the standard normal law.

For real number $y \geq 0$, we let $[y]$ stand for the largest integer smaller or equal to y and we write $\{y\} := y - [y]$ for its fractional part. As usual, the letter p will always denote a prime number, while $\pi(x)$ will stand for the number of prime numbers $p \leq x$. On the other hand, $\phi(y) := \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$ stands for the density function of the standard normal law. Moreover, we let $\omega(n)$ stand for the number of distinct prime factors of n and we let $\gamma(n) := \prod_{p|n} p$ be the *kernel* of n . Finally, by $\log \log x$ we mean $\max(1, \log \log x)$.

2. The main results

It is clear that

$$F_q(n) := \sum_{p|n} \left(\left\lfloor \frac{\log p}{\log q} \right\rfloor + 1 \right) + \sum_{\substack{a \geq 2 \\ p^a \parallel n}} \left(\left\lfloor \frac{\log a}{\log q} \right\rfloor + 1 \right).$$

The first sum counts the number of digits needed to write the prime factors of n while the second counts the number of digits needed to write the exponents ≥ 2 . Using the identities

$$\left\lfloor \frac{\log p}{\log q} \right\rfloor = \frac{\log p}{\log q} - \left\{ \frac{\log p}{\log q} \right\} \quad \text{and} \quad \sum_{p|n} \frac{\log p}{\log q} = \frac{\log \gamma(n)}{\log q},$$

it is easily seen that

$$F_q(n) := \frac{\log \gamma(n)}{\log q} + \sum_{p|n} \left(1 - \left\{ \frac{\log p}{\log q} \right\} \right) + \sum_{\substack{a \geq 2 \\ p^a \parallel n}} \left(\left\lfloor \frac{\log a}{\log q} \right\rfloor + 1 \right),$$

which can also be written as

$$(1) \quad F_q(n) = \frac{\log n}{\log q} - h_1(n) + h_2(n) + h_3(n),$$

where

$$\begin{aligned}
 h_1(n) &:= \frac{\log(n/\gamma(n))}{\log q}, \\
 h_2(n) &:= \sum_{\substack{a \geq 2 \\ p^a \parallel n}} \left(\left\lfloor \frac{\log a}{\log q} \right\rfloor + 1 \right), \\
 h_3(n) &:= \sum_{p|n} \left(1 - \left\lfloor \frac{\log p}{\log q} \right\rfloor \right).
 \end{aligned}$$

Let $H_q(x, y)$ be the distribution function of F_q , that is

$$H_q(x, y) := \sharp\{n < x : F_q(n) < y\}$$

and consider the function

$$G(c) = \lim_{x \rightarrow \infty} \frac{1}{x} H_q \left(x, \frac{\log x}{\log q} + \frac{1}{2} \log \log x + c\sqrt{\log \log x} \right).$$

Theorem 1. *For each real number c ,*

$$G(c) = \Phi(\sqrt{3}c).$$

Remark. The fact that the function $G(c)$ is well defined is in itself an interesting result.

The following theorem reveals the interval in which the function $F_q(n)$ takes its values.

Theorem 2. *For each integer $q \geq 2$ and each integer $n \geq 2$,*

$$\left\lfloor \frac{\log \log (n^{1/\omega(n)})}{\log q} \right\rfloor + \omega(n) \leq F_q(n) \leq \left\lfloor \frac{\log n}{\log q} \right\rfloor + 2\omega(n).$$

3. Preliminary results

The first lemma contains classical estimates on powerful numbers. Recall that a positive integer is said to be a *powerful number* if $p|n$ implies that $p^2|n$.

But first, some notation. Given a positive integer n , we shall write $n = uv$, where

$$u = u(n) := \prod_{p \parallel n} p$$

and

$$v = v(n) := \frac{n}{u},$$

so that u is the square free part of n and v its powerful part.

Lemma 1. *As $y \rightarrow \infty$,*

$$(i) \quad \sum_{\substack{n > y \\ p|n \Rightarrow p^2|n}} \frac{1}{n} \ll \frac{1}{\sqrt{y}},$$

$$(ii) \quad \#\{n < x : v(n) > y\} \ll \frac{x}{\sqrt{y}},$$

where the implicit constant does not depend on x .

Proof of Lemma 1. For (i), see De Koninck and Kátai [2].

To establish (ii), we simply observe that it follows from (i) that

$$\#\{n < x : v(n) > y\} \leq \sum_{\substack{v > y \\ p|v \Rightarrow p^2|v}} \frac{x}{v} \ll \frac{x}{\sqrt{y}}.$$

Lemma 2. *There exist two positive constants c_1 and c_2 such that, as $x \rightarrow \infty$,*

$$\sum_{p \leq x} \frac{1}{p} = \log \log x + c_1 + O\left(\exp\{-c_2(\log x)^{3/5}\}\right).$$

Proof of Lemma 2. It is known (see Vinogradov [8]) that

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}}{\log x} \left(1 + O\left(\exp\{-c_2(\log x)^{3/5}\}\right)\right),$$

where γ is Euler’s constant. Taking logarithms on both sides, we easily see that

$$\sum_{p \leq x} \frac{1}{p} = \log \log x + \gamma - \sum_{p \leq x, \nu \geq 2} \frac{1}{\nu p^\nu} + \log \left(1 + O\left(\exp\{-c_2(\log x)^{3/5}\}\right)\right) =$$

$$\begin{aligned}
 &= \log \log x + \gamma - \sum_{p, \nu \geq 2} \frac{1}{\nu p^\nu} + O\left(\frac{1}{x}\right) + O\left(\exp\{-c_2(\log x)^{3/5}\}\right) = \\
 &= \log \log x + c_1 + O\left(\exp\{-c_2(\log x)^{3/5}\}\right),
 \end{aligned}$$

as required.

Lemma 3. (Central limit theorem) *Let X_1, X_2, \dots be independent random variables and let*

$$\begin{aligned}
 \mu_i &= E[X_i], \\
 \sigma_i^2 &= E[(X_i - \mu_i)^2], \\
 r_i^3 &= E[(X_i - \mu_i)^3].
 \end{aligned}$$

If

$$\lim_{n \rightarrow \infty} \frac{\left(\sum_{i=1}^n r_i^3\right)^{1/3}}{\sqrt{\sum_{i=1}^n \sigma_i^2}} = 0,$$

then

$$\lim_{n \rightarrow \infty} P\left(\frac{\sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}} < y\right) = \Phi(y).$$

Proof of Lemma 3. This is Lyapunov’s condition in the Central Limit Theorem. For a proof this classical result, see Bernstein [1].

Lemma 4. *For each fixed integer $q \geq 2$ and each fixed integer $r \geq 1$, we have, as $x \rightarrow \infty$,*

$$\begin{aligned}
 \sum_{p < x} \frac{1}{p} \left\{ \frac{\log p}{\log q} \right\} &= \frac{\log \log x}{2} + O(1), \\
 \sum_{p < x} \frac{1}{p} \left\{ \frac{\log p}{\log q} \right\}^r &= \frac{\log \log x}{r+1} + O\left(\sqrt{\frac{\log \log x}{r}}\right).
 \end{aligned}$$

Proof of Lemma 4. We first establish the second relation. To do so, we call upon the following inequality which is valid for all positive integers k and r :

$$(2) \quad \sum_{p < x} \frac{1}{p} \left\{ \frac{\log p}{\log q} \right\}^r = \sum_{j=0}^{k-1} \sum_{\substack{p < x \\ \frac{j}{k} \leq \left\{ \frac{\log p}{\log q} \right\} < \frac{j+1}{k}}} \frac{1}{p} \left\{ \frac{\log p}{\log q} \right\}^r.$$

The sum on the right hand side can be written as

$$(3) \quad \sum_{\substack{p < x \\ \frac{j}{k} \leq \left\{ \frac{\log p}{\log q} \right\} < \frac{j+1}{k}}} \frac{1}{p} \left\{ \frac{\log p}{\log q} \right\}^r = \sum_{\ell=0}^{\lfloor \frac{\log x}{\log q} \rfloor} \sum_{q^{\ell + \frac{j}{k}} \leq p < \min(q^{\ell + \frac{j+1}{k}}, x)} \frac{1}{p} \left\{ \frac{\log p}{\log q} \right\}^r.$$

On the other hand, observe that

$$(4) \quad \sum_{q^{\ell + \frac{j}{k}} \leq p < q^{\ell + \frac{j+1}{k}}} \frac{1}{p} \left\{ \frac{\log p}{\log q} \right\}^r = \left(\frac{j}{k} + \frac{\xi}{k} \right)^r \sum_{q^{\ell + \frac{j}{k}} \leq p < q^{\ell + \frac{j+1}{k}}} \frac{1}{p}$$

for some real ξ such that $|\xi| < 1$. Using Lemma 2 (replacing the error term by $O(1/\log^2 x)$, say), we obtain

$$(5) \quad \begin{aligned} \sum_{q^{\ell + \frac{j}{k}} \leq p < q^{\ell + \frac{j+1}{k}}} \frac{1}{p} &= \log \log \left(q^{\ell + \frac{j+1}{k}} \right) - \log \log \left(q^{\ell + \frac{j}{k}} \right) + O \left(\frac{1}{\ell^2 \log^2 q} \right) = \\ &= \log \left(\ell + \frac{j+1}{k} \right) - \log \left(\ell + \frac{j}{k} \right) + O \left(\frac{1}{\ell^2 \log^2 q} \right) = \\ &= \frac{1}{k\ell} + O \left(\frac{1}{k\ell^2} \right) + O \left(\frac{1}{\ell^2 \log^2 q} \right). \end{aligned}$$

Combining relations (3), (4) and (5), we obtain that

$$(6) \quad \begin{aligned} \sum_{\substack{p < x \\ \frac{j}{k} \leq \left\{ \frac{\log p}{\log q} \right\} < \frac{j+1}{k}}} \frac{1}{p} \left\{ \frac{\log p}{\log q} \right\}^r &= \\ &= \left(\frac{j}{k} + \frac{\xi}{k} \right)^r \sum_{\ell=1}^{\lfloor \frac{\log x}{\log q} \rfloor} \left(\frac{1}{k\ell} + O \left(\frac{1}{k\ell^2} \right) + O \left(\frac{1}{\ell^2 \log^2 q} \right) \right). \end{aligned}$$

Observe also that

$$(7) \quad \sum_{\ell=1}^{\lfloor \frac{\log x}{\log q} \rfloor} \left(\frac{1}{k\ell} + O\left(\frac{1}{k\ell^2}\right) + O\left(\frac{1}{\ell^2 \log^2 q}\right) \right) = \frac{1}{k} \log \log x + O(1).$$

Combining relations (2), (6) and (7) with the identity

$$\left(\frac{j}{k} + \frac{\xi}{k}\right)^r = \frac{j^r}{k^r} + O\left(\frac{r(j+1)^{r-1}}{k^r}\right),$$

we obtain

$$(8) \quad \sum_{p < x} \frac{1}{p} \left\{ \frac{\log p}{\log q} \right\}^r = \\ = \sum_{j=0}^{k-1} \left(\frac{j^r}{k^{r+1}} \log \log x + O\left(\frac{r(j+1)^{r-1}}{k^{r+1}} \log \log x\right) + O\left(\frac{j^r}{k^r}\right) \right) + O(1).$$

The right hand side member of (8) is equal to

$$(9) \quad \frac{1}{r+1} \log \log x + O\left(\frac{(k+1)^r}{k^{r+1}} \log \log x\right) + O\left(\frac{k}{r}\right) + O(1).$$

Choosing $k = \lfloor \sqrt{r \log \log x} \rfloor$, the proof of the second relation of Lemma 4 then follows from relations (8) and (9).

In order to prove the first relation of Lemma 4, we first observe that, using the Prime Number Theorem in the form $\sum_{p < x} \frac{\log p}{p} = \log x + O(1)$, we have

$$\sum_{p < x} \frac{1}{p} \left\{ \frac{\log p}{\log q} \right\} = \sum_{p < x} \frac{1}{p} \frac{\log p}{\log q} - \sum_{p < x} \frac{1}{p} \left[\frac{\log p}{\log q} \right] = \\ = \frac{\log x}{\log q} + O(1) - \sum_{p < x} \frac{1}{p} \left[\frac{\log p}{\log q} \right].$$

Moreover,

$$\sum_{p < x} \frac{1}{p} \left[\frac{\log p}{\log q} \right] = \sum_{j=0}^{\lfloor \frac{\log x}{\log q} \rfloor} j \sum_{q^j < p \leq \min(x, q^{j+1})} \frac{1}{p}.$$

Using Lemma 2 (replacing the error term by $O(1/\log^3 x)$, say) we obtain

$$\begin{aligned} \sum_{q^j < p \leq q^{j+1}} \frac{1}{p} &= \log \log q^{j+1} - \log \log q^j + O\left(\frac{1}{j^3 \log^3 q}\right) = \\ &= \frac{1}{j} - \frac{1}{2j^2} + O\left(\frac{1}{j^3}\right) \quad (j \geq 1). \end{aligned}$$

We may therefore conclude that

$$\begin{aligned} \sum_{p < x} \frac{1}{p} \left\lfloor \frac{\log p}{\log q} \right\rfloor &= \sum_{j=1}^{\lfloor \frac{\log x}{\log q} \rfloor - 1} \left(1 - \frac{1}{2j} + O\left(\frac{1}{j^2}\right)\right) + O(1) = \\ &= \frac{\log x}{\log q} - \frac{\log \log x}{2} + O(\log \log q) + O(1), \end{aligned}$$

which proves the first equation of Lemma 4 and thus completes the proof of the lemma.

Let x be a large fixed positive integer and set

$$R := x \prod_{p < x} p.$$

We consider the set $U = \{n < R\}$ with the probability measure

$$P(S) = \frac{\#S}{R}, \quad \text{for each } S \subseteq U.$$

For each prime number $p < x$, we introduce the random variables

$$\xi_p(n) := \begin{cases} 1 - \left\lfloor \frac{\log p}{\log q} \right\rfloor & \text{if } p|n, \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 5. *For each prime number $p < x$, the following equalities hold:*

$$\begin{aligned} \mu_p &:= E[\xi_p] = \frac{1}{p} \left(1 - \left\lfloor \frac{\log p}{\log q} \right\rfloor\right), \\ \sigma_p^2 &:= E[(\xi_p - \mu_p)^2] = \left(\frac{1}{p} - \frac{1}{p^2}\right) \left(1 - \left\lfloor \frac{\log p}{\log q} \right\rfloor\right)^2, \\ E[(\xi_p - \mu_p)^3] &= \left(\frac{1}{p} - \frac{3}{p^2} + \frac{2}{p^3}\right) \left(1 - \left\lfloor \frac{\log p}{\log q} \right\rfloor\right)^3. \end{aligned}$$

Proof of Lemma 5. Since for each prime number $p < x$, we have $p|R$, the random variables ξ_p are independent. Moreover, one can easily verify the following equalities:

$$P\left(\xi_p = 1 - \left\{\frac{\log p}{\log q}\right\}\right) = \frac{1}{p},$$

$$P(\xi_p = 0) = \frac{p-1}{p}.$$

From these, it follows immediately that

$$(10) \quad E[\xi_p] = \frac{1}{p} \left(1 - \left\{\frac{\log p}{\log q}\right\}\right),$$

$$(11) \quad E[\xi_p^2] = \frac{1}{p} \left(1 - \left\{\frac{\log p}{\log q}\right\}\right)^2,$$

$$(12) \quad E[\xi_p^3] = \frac{1}{p} \left(1 - \left\{\frac{\log p}{\log q}\right\}\right)^3$$

All three equalities of Lemma 5 then easily follow from (10), (11) and (12).

Lemma 6. For each real number y ,

$$\lim_{x \rightarrow \infty} P\left(\frac{\sum_{p < x} \xi_p - \frac{1}{2} \log \log x}{\sqrt{\frac{1}{3} \log \log x}} < y\right) = \Phi(y).$$

Proof of Lemma 6. This result follows from Lemmas 4, 5 and 3 (Central Limit Theorem).

On the same probability space $\{n < R\}$, we define the random variables

$$\chi_p(n) := \begin{cases} 1 - \left\{\frac{\log p}{\log q}\right\} & \text{if } p|a, \\ 0 & \text{otherwise.} \end{cases}$$

where a is the smallest positive integer such that $a \equiv n \pmod{x}$.

Lemma 7. *As $x \rightarrow \infty$,*

$$E \left[\left| \sum_p \xi_p - \sum_p \chi_p \right| \right] < \frac{\pi(x)}{x} = \frac{1 + o(1)}{\log x}.$$

Proof of Lemma 7. We need to observe that

$$P \left(\chi_p = 1 - \left\{ \frac{\log p}{\log q} \right\} \right) = \frac{1}{R} \#\{n < R : p|a\} = \frac{1}{R} \frac{R}{x} \left\lfloor \frac{x}{p} \right\rfloor = \frac{1}{x} \left\lfloor \frac{x}{p} \right\rfloor.$$

Indeed, it then follows that

$$E[\chi_p] = \left(1 - \left\{ \frac{\log p}{\log q} \right\} \right) \frac{1}{x} \left\lfloor \frac{x}{p} \right\rfloor = \frac{1}{p} \left(1 - \left\{ \frac{\log p}{\log q} \right\} \right) + \frac{\xi}{x}$$

for some $|\xi| < 1$. Hence

$$|E[\chi_p] - E[\xi_p]| < \frac{1}{x}$$

and therefore

$$E \left[\left| \sum_p \xi_p - \sum_p \chi_p \right| \right] \leq \sum_{p < x} |E[\chi_p] - E[\xi_p]| < \frac{\pi(x)}{x},$$

which completes the proof of Lemma 7.

Lemma 8. *As $x \rightarrow \infty$,*

$$P \left(\left| \sum_{p < x} \xi_p - \sum_{p < x} \chi_p \right| > 1 \right) < \frac{1 + o(1)}{\log x}.$$

Proof of Lemma 8. This result is an immediate consequence of Lemma 7 and the Markov inequality (see for instance Galambos [6], p.150).

Lemma 9. *Given a fixed integer $N \geq 2$, let $\alpha_i \geq t \geq N^{1/(N-1)}$ for $i = 1, \dots, N$. Then*

$$\sum_{i=1}^N \alpha_i \leq \frac{1}{c} \prod_{i=1}^N \alpha_i,$$

where $c = \frac{t^{N-1}}{N}$.

Proof of Lemma 9. Assume that $\alpha_i \geq t \geq N^{1/(N-1)}$ for $i = 1, \dots, N$ and that

$$\sum_{i=1}^N \alpha_i > \frac{1}{c} \prod_{i=1}^N \alpha_i.$$

We then have

$$\sum_{i=1}^N \alpha_i > \left(\frac{N \prod_{i=1}^N \alpha_i}{t^{N-1} \alpha_j} \right) \alpha_j \quad (j = 1, \dots, N).$$

Observe that, using the fact that $\alpha_i \geq t$,

$$\frac{N \prod_{i=1}^N \alpha_i}{t^{N-1} \alpha_j} \geq N \quad (j = 1, \dots, N).$$

We therefore obtain that for each integer $j = 1, \dots, N$,

$$\sum_{i=1}^N \alpha_i > N \alpha_j,$$

which contradicts the fact that

$$\sum_{i=1}^N \alpha_i \geq N \max_i \alpha_i,$$

thus completing the proof of Lemma 9.

4. The proofs of the main results

Proof of Theorem 1. Assume that $n \leq x$ satisfies the inequality

$$v(n) < \log \log n.$$

By Lemma 1 (ii), we thus omit at most $\frac{x}{\sqrt{\log \log x}}$ integers $n \leq x$. By the definition of the function $h_1(n)$, we then obtain

$$(13) \quad h_1(n) = O(\log \log \log n).$$

Moreover, by definition, we have

$$h_2(n) \leq \omega(v(n)) \left(\left\lfloor \frac{\log \left(\frac{\log v(n)}{\log 2} \right)}{\log q} \right\rfloor + 1 \right).$$

It follows from this that

$$(14) \quad h_2(n) \ll \frac{\log v(n)}{\log \log v(n)} \log \log v(n) \ll \log \log \log n.$$

Hence, combining (13) and (14), we have

$$(15) \quad h_1(n) + h_2(n) = O(\log \log \log n).$$

Assume also that $\frac{x}{\log \log x} < n < x$, so that

$$(16) \quad \frac{\log n}{\log q} = \frac{\log x}{\log q} + O(\log \log \log x).$$

Combining (1), (13), (15) and (16), we obtain

$$(17) \quad \# \left\{ n < x : F_q(n) < \frac{\log x}{\log q} + w \right\} = \\ = \# \{ n < x : h_3(n) < w + O(\log \log \log x) \} + O \left(\frac{x}{\sqrt{\log \log x}} \right).$$

Calling upon the identity

$$(18) \quad h_3(n) = \sum_{p < x} \chi_p(n),$$

it follows from (17) and (18) that

$$(19) \quad \# \left\{ n < x : F_q(n) < \frac{\log x}{\log q} + w \right\} = \\ = \# \left\{ n < x : \sum_{p < x} \chi_p(n) < w + O(\log \log \log x) \right\} + O \left(\frac{x}{\sqrt{\log \log x}} \right).$$

By the definition of the $\chi_p(n)$, we have that

$$(20) \quad \# \left\{ n < x : \sum_{p < x} \chi_p(n) < w + O(\log \log \log x) \right\} = \\ = \frac{x}{R} \# \left\{ n < R : \sum_{p < x} \chi_p(n) < w + O(\log \log \log x) \right\}.$$

On the other hand, by Lemma 8, we have that

$$(21) \quad \# \left\{ n < R : \sum_{p < x} \chi_p(n) < w + O(\log \log \log x) \right\} = \\ = \# \left\{ n < R : \sum_{p < x} \xi_p(n) < w + O(\log \log \log x) \right\} + O\left(\frac{R}{\log x}\right).$$

From (21) and Lemma 7, it then follows that

$$(22) \quad \# \left\{ n < R : \sum_{p < x} \chi_p(n) < \frac{1}{2} \log \log x + c\sqrt{\log \log x} + O(\log \log \log x) \right\} = \\ = R(1 + o(1))\Phi(\sqrt{3}c).$$

Combining (19), (20) and (22), we finally obtain

$$\# \left\{ n < x : F_q(n) < \frac{\log x}{\log q} + \frac{1}{2} \log \log x + c\sqrt{\log \log x} \right\} = x(1 + o(1))\Phi(\sqrt{3}c),$$

thus completing the proof of Theorem 1.

Proof of Theorem 2. We first proof the upper bound. We have

$$F_q(n) = \sum_{p|n} \left(\left\lfloor \frac{\log p}{\log q} \right\rfloor + 1 \right) + \sum_{\substack{p^a || n \\ a \geq 2}} \left(\left\lfloor \frac{\log a}{\log q} \right\rfloor + 1 \right) \leq \\ \leq \frac{\log \left(\prod_{p^a || n} ap \right)}{\log q} + 2\omega(n).$$

Since $a^{\frac{1}{a-1}} \leq 2$ for each $a \geq 2$, we have that $ap \leq p^a$ for each prime $p \geq 2$. Hence,

$$F_q(n) \leq \frac{\log n}{\log q} + 2\omega(n),$$

thus establishing the upper bound.

We now prove the lower bound. As before, we write $n = u(n)v(n)$. Since $(u(n), v(n)) = 1$, we have

$$\begin{aligned} F_q(n) &\geq \sum_{p|u(n)} \max\left(1, \frac{\log p}{\log q}\right) + \sum_{p^a \| v(n)} \max\left(2, \frac{\log \log p^a}{\log q}\right) = \\ &= \frac{1}{\log q} \left(\sum_{p|u(n)} \max(\log q, \log p) + \sum_{p^a \| v(n)} \max(\log q^2, \log \log p^a) \right) \geq \\ &\geq \frac{1}{\log q} \sum_{p^a \| n} \max(\log q, \log \log p^a) = \\ &= \frac{1}{\log q} \log \left(\prod_{p^a \| n} \max(q, \log p^a) \right). \end{aligned}$$

Using Lemma 9, we then have

$$\begin{aligned} F_q(n) &\geq \frac{1}{\log q} \log \left(\frac{q^{\omega(n)-1}}{\omega(n)} \sum_{p^a \| n} \max(q, \log p^a) \right) \geq \\ &\geq \frac{1}{\log q} \log \left(\frac{q^{\omega(n)-1}}{\omega(n)} \log n \right) = \\ &= \frac{\log \log n}{\log q} + \omega(n) - 1 - \frac{\log \omega(n)}{\log q}. \end{aligned}$$

Moreover, since $\frac{\log \log n^{\frac{1}{\omega(n)}}}{\log q}$ is not an integer for $n, q \geq 2$, it follows that

$$\begin{aligned} F_q(n) &\geq \left\lceil \frac{\log \log n^{\frac{1}{\omega(n)}}}{\log q} + \omega(n) - 1 \right\rceil = \\ &= \left\lfloor \frac{\log \log n^{\frac{1}{\omega(n)}}}{\log q} \right\rfloor + \omega(n) - 1 = \\ &= \left\lfloor \frac{\log \log n^{\frac{1}{\omega(n)}}}{\log q} \right\rfloor + \omega(n), \end{aligned}$$

thus establishing the lower bound and completing the proof of **Theorem 2**.

5. Final remarks

The study of the behavior of the function $H_q(x, y)$ is still very much uncharted. For instance, for any fixed value of y , **Theorem 2** only reveals that $H_q(\infty, y) < \infty$. Hence obtaining a general fairly good estimate for $H_q(x, y)$ is certainly an interesting challenge. On the other hand, we believe that the result for economical numbers could be generalized to yield

$$H_q\left(x, \frac{\log x}{\log q} + c \log \log x\right) = \frac{x}{(\log x)^{R(q,c)+o(1)}} \quad \left(x \rightarrow \infty, -\infty < c < \frac{1}{2}\right).$$

To prove or disprove this claim and moreover to describe the behavior of the function $R(q, c)$ in the eventuality that the claim is true would also be very interesting.

Acknowledgement. The first author was supported in part by a grant from NSERC.

References

- [1] **Бернштейн С.Н.**, О работе П.Л. Чебышева в теории вероятностей, *Научное наследие П.Л. Чебышева. Выпуск первый: Математика*, под ред. С.Н. Бернштейна, Академия наук СССР, Москва-Ленинград, 1945. (*Bernstein S.N.*, On the work of P.L. Chebyshev in probability theory, *The scientific legacy of P.L. Chebyshev. Part I: Mathematics*, ed. S.N. Bernstein, USSR Academy of Sciences, Moscow-Leningrad, 1945. (in Russian))
- [2] **De Koninck J.-M. and Kátai I.**, On the mean value of the index of composition, *Monatshefte für Mathematik*, **145** (2) (2005), 131-144.
- [3] **De Koninck J.-M. and Luca F.**, On strings of consecutive economical numbers of arbitrary length, *Integers*, **5** (2005), #A5.
- [4] **De Koninck J.-M. and Luca F.**, Counting the number of economical numbers, *Publ. Math. Debrecen*, **68** (2006), 97-113.
- [5] **De Koninck J.-M., Doyon N. et Luca F.**, Sur la quantité de nombres économiques, *Acta Arithmetica* (à paraître)

- [6] **Galambos J.**, *Introductory probability theory*, Marcel Dekker, New York, 1984.
- [7] **Santos B.R.**, Problem 2204. Equidigital representation, *J. Recreational Mathematics*, **27** (1995), 58-59.
- [8] **Виноградов А.И.**, Об остатке в формуле Мертенса, *Доклады АН СССР*, **148** (2) (1963), 262-263. (*Vinogradov A.I.*, On the remainder in Merten's formula, *Dokl. Akad. Nauk SSSR* **148** (1963), 262-263. (in Russian))

J.-M. DeKoninck, N. Doyon and P. Letendre

Département de mathématiques et de statistique

Université Laval

Québec G1K 7P4, Canada

`jmdk@mat.ulaval.ca`

`nicolas.doyon@mat.ulaval.ca`

`patrick.letendre.1@ulaval.ca`