Using large prime divisors to construct normal numbers

Jean-Marie De Koninck[1] and Imre Kátai[2]

Dedicated to Professor Karl-Heinz Indlekofer on his seventieth anniversary

Édition du 24 septembre 2012

Abstract

Given an integer $q \geq 2$, a *q-normal number* is an irrational number ξ such that any preassigned sequence of ℓ digits occurs in the q-ary expansion of ξ at the expected frequency, namely $1/q^\ell$. Let $\eta(x)$ be a slowly increasing function such that $\frac{\log \eta(x)}{\log x} \to 0$ as $x \to \infty$. Then, letting $P(n)$ stand for the largest prime factor of n, set $Q(n)$ to be the smallest prime divisor of n which is larger than $\eta(n)$, while setting $Q(n) = 1$ if $P(n) > \eta(n)$. Then, we show that the real number $0.Q(1)Q(2)\ldots$ is a normal number in base 10. With various similar constructions, we create large families of normal numbers in any given base $q \geq 2$. Finally, we consider exponential sums involving the $Q(n)$ function.

AMS Subject Classification numbers: 11K16, 11N37, 11A41
Key words: normal numbers, largest prime factor, smallest prime factor

1 Introduction

Given an integer $q \geq 2$, a *q-normal number*, or simply a *normal number*, is an irrational number whose q-ary expansion is such that any preassigned sequence, of length $\ell \geq 1$, of base q digits from this expansion, occurs at the expected frequency, namely $1/q^\ell$.

Let $A_q := \{0, 1, \ldots, q-1\}$. Given an integer $\ell \geq 1$, an expression of the form $i_1 i_2 \ldots i_\ell$, where each $i_j \in A_q$ is called a *word* of length ℓ. We sometimes write $\lambda(\beta) = \ell$ to indicate that β is a *word* of length ℓ. The symbol Λ will denote the *empty word*. We let A_q^ℓ stand for the set of all words of length ℓ and A_q^* stand for the set of all the words regardless of their length.

Given a positive integer n, we write its q-ary expansion as

$$n = \varepsilon_0(n) + \varepsilon_1(n) q + \cdots + \varepsilon_t(n) q^t, \tag{1.1}$$

where $\varepsilon_i(n) \in A_q$ for $0 \leq i \leq t$ and $\varepsilon_t(n) \neq 0$. To this representation, we associate the word

$$\overline{n} = \varepsilon_0(n)\varepsilon_1(n)\ldots\varepsilon_t(n) = \varepsilon_0\varepsilon_1\ldots\varepsilon_t \in A_q^{t+1}. \tag{1.2}$$

[1]Research supported in part by a grant from NSERC.
[2]Research supported by ELTE IK.

Let $P(n)$ stand for the largest prime factor of $n \geq 2$, with $P(1) = 1$. In a recent paper [5], we showed that if $F \in \mathbb{Z}[x]$ is a polynomial of positive degree with $F(x) > 0$ for $x > 0$, then the real numbers

$$0.\overline{F(P(2))}\,\overline{F(P(3))}\dots\overline{F(P(n))}\dots$$

and

$$0.\overline{F(P(2+1))}\,\overline{F(P(3+1))}\dots\overline{F(P(p+1))}\dots,$$

where p runs through the sequence of primes, are q-normal numbers.

Let $\eta(x)$ be a slowly increasing function, that is an increasing function satisfying $\lim_{x\to\infty} \frac{\eta(cx)}{\eta(x)} = 1$ for any fixed constant $c > 0$. Being slowly increasing, it satisfies in particular the condition $\frac{\log \eta(x)}{\log x} \to 0$ as $x \to \infty$.

We then let $Q(n)$ be the smallest prime divisor of n which is larger than $\eta(n)$, while setting $Q(n) = 1$ if $P(n) > \eta(n)$. Then, we show that the real number $0.\overline{Q(1)}\,\overline{Q(2)}\,\overline{Q(3)}\dots$ is a q-normal number. With various similar constructions, we create large families of normal numbers in any given base $q \geq 2$.

Finally, we consider exponential sums involving the $Q(n)$ function.

2 Main results

Theorem 1. *Given an arbitrary basis $q \geq 2$ and for any integer n, let $\overline{n}$ be as in (1.2). Then the number*

$$\xi_1 = 0.\overline{Q(1)}\,\overline{Q(2)}\,\overline{Q(3)}\dots$$

is a q-normal number.

Let $\wp$ stand for the set of all primes. Given an integer $q \geq 2$, let $\mathcal{R}, \wp_0, \wp_1, \dots, \wp_{q-1}$ be disjoint sets of prime numbers such that

$$\wp = \mathcal{R} \cup \wp_0 \cup \wp_1 \cup \dots \cup \wp_{q-1},$$

and such that, uniformly for $2 \leq v \leq u$ as $u \to \infty$,

$$\pi([u, u+v] \cap \wp_j) = \frac{1}{q}\pi([u, u+v]) + O\left(\frac{u}{\log^5 u}\right) \qquad (j = 0, 1, \dots, q-1),$$

so that, in particular,

$$\pi([u, u+v] \cap \mathcal{R}) = O\left(\frac{u}{\log^5 u}\right).$$

Then, consider the function κ defined on $\wp$ as follows:

$$\kappa(p) = \begin{cases} \ell & \text{if } p \in \wp_\ell, \\ \Lambda & \text{if } p \in \mathcal{R}. \end{cases}$$

With this notation, we have

Theorem 2. *The number*

$$\xi_2 = 0.\kappa(Q(1))\kappa(Q(2))\kappa(Q(3))\dots$$

is a q-normal number.

Remark 1. *In an earlier paper [4], we used such classification of prime numbers to create normal numbers, but by simply concatenating the numbers* $\kappa(1)$, $\kappa(2)$, $\kappa(3)$, ...

Let a be a fixed non zero integer. Then we have the following result.

Theorem 3. *The number*

$$\xi_3 = 0.\kappa(Q(2+a))\kappa(Q(3+a))\kappa(Q(5+a))\dots\kappa(Q(p+a))\dots,$$

where p runs through the set of primes, is a q-normal number.

Define $\wp^*$ as the set of all the prime numbers $p \equiv 1 \pmod 4$. Then, let $\mathcal{R}^*, \wp_0^*, \wp_1^*, \dots, \wp_{q-1}^*$ be disjoint sets of prime numbers such that

$$\wp^* = \mathcal{R}^* \cup \wp_0^* \cup \wp_1^* \cup \cdots \cup \wp_{q-1}^*,$$

and such that, uniformly for $2 \le v \le u$ as $u \to \infty$,

$$\pi([u, u+v] \cap \wp_j^*) = \frac{1}{q}\pi([u, u+v] \cap \wp^*) + O\left(\frac{u}{\log^5 u}\right) \qquad (j = 0, 1, \dots, q-1),$$

so that, in particular,

$$\pi([u, u+v] \cap \mathcal{R}^*) = O\left(\frac{u}{\log^5 u}\right).$$

Then, consider the following function defined on $\wp$ as follows

$$\nu(p) = \begin{cases} \ell & \text{if } p \in \wp_\ell^*, \\ \Lambda & \text{if } p \notin \bigcup_{\ell=0}^{q-1} \wp_\ell^*. \end{cases}$$

With this notation, we have the following result.

Theorem 4. *The number*

$$\xi_4 = 0.\nu(Q(1))\nu((Q(2))\nu(Q(3))\dots$$

is a q-normal number.

Consider the arithmetic function $f(n) = n^2 + 1$. Then, we have the following result.

Theorem 5. *The two numbers*

$$\begin{aligned}\xi_5 &= 0.\kappa(Q(f(1)))\kappa(Q(f(2)))\kappa(Q(f(3)))\dots,\\ \xi_6 &= 0.\kappa(Q(f(2)))\kappa(Q(f(3)))\kappa(Q(f(5)))\dots\kappa(Q(f(p)))\dots,\end{aligned}$$

where p runs through the set of primes, are q-normal numbers.

Remark 2. *One can show that this last result remains true if $f(n)$ is replaced by another non constant irreducible polynomial.*

We now introduce the product function $F(n) = n(n+1)\cdots(n+q-1)$. Observe that if for some positive integer n, we have $p = Q(F(n)) > q$, then $p|n+\ell$ only for one $\ell \in \{0, 1, \dots, q-1\}$, implying that ℓ is uniquely determined for all positive integers n such that $Q(F(n)) > q$. Thus we may define the function

$$\tau(n) = \begin{cases} \ell & \text{if } p = Q(F(n)) > q \text{ and } p|n+\ell, \\ \Lambda & \text{otherwise.} \end{cases}$$

Using this notation, we have the following result.

Theorem 6. *The number*

$$\xi_7 = 0.\tau(q+1)\tau(q+2)\tau(q+3)\dots$$

is a q-normal number.

We now introduce the product function $G(n) = (n+1)(n+2)\cdots(n+q)$ and further define the function

$$\rho(n) = \begin{cases} \ell & \text{if } p = Q(G(n)) > q+1 \text{ and } p|n+\ell+1, \\ \Lambda & \text{otherwise.} \end{cases}$$

Moreover, let $(p_j)_{j\ge1}$ be the sequence of all primes larger than q, that is, $q < p_1 < p_2 < \cdots$ With this notation, we have the following result.

Theorem 7. *The number*

$$\xi_8 = 0.\rho(p_1)\rho(p_2)\rho(p_3)\dots$$

is a q-normal number.

Let α be an arbitrary irrational number. We will be using the standard notation $e(y) = \exp\{2\pi i y\}$. We then have the following.

Theorem 8. *Let*

$$A(x) := \sum_{n\le x} f(n) e(\alpha Q(n)),$$

where f is any given multiplicative function satisfying $|f(n)| = 1$ for all positive integers n. Then,

$$\lim_{x\to\infty} \frac{A(x)}{x} = 0. \tag{2.1}$$

3 Notation and preliminary lemmas

For each integer $n \geq 2$, let $L(n) = \left\lfloor \dfrac{\log n}{\log q} \right\rfloor$. Let $\beta \in A_q^\ell$ and n be written as in (1.1). We then let $\nu_\beta(\overline{n})$ stand for the number of occurrences of the word β in the q-ary expansion of the positive integer n, that is, the number of times that $\varepsilon_j(n)\ldots\varepsilon_{j+\ell-1}(n) = \beta$ as j varies from 0 to $t-(\ell-1)$.

The letters p and π will always denote prime numbers. The letter c with or without subscript always denotes a positive constant but not necessarily the same at each occurrence.

We will be using a key result obtained by Bassily and Kátai [1] and which we state here as two lemmas, a proof of which, in a more general context, can be found in our earlier paper [5].

Lemma 1. *Let κ_u be a function of u such that $\kappa_u > 1$ for all u. Given a word $\beta \in A_q^\ell$ and setting*

$$V_\beta(u) := \#\left\{p \in \wp : u \leq p \leq 2u \text{ such that } \left|\nu_\beta(\overline{p}) - \frac{L(u)}{q^\ell}\right| > \kappa_u\sqrt{L(u)}\right\},$$

then, there exists a positive constant c such that

$$V_\beta(u) \leq \frac{cu}{(\log u)\kappa_u^2}.$$

Lemma 2. *Let κ_u be as in Lemma 1. Given $\beta_1, \beta_2 \in A_q^\ell$ with $\beta_1 \neq \beta_2$, set*

$$\Delta_{\beta_1,\beta_2}(u) := \#\left\{p \in \wp : u \leq p \leq 2u \text{ such that } |\nu_{\beta_1}(\overline{p}) - \nu_{\beta_2}(\overline{p})| > \kappa_u\sqrt{L(u)}\right\}.$$

Then, for some positive constant c,

$$\Delta_{\beta_1,\beta_2}(u) \leq \frac{cu}{(\log u)\kappa_u^2}.$$

4 Proof of Theorem 2

We start by proving Theorem 2 since its content will be useful for the proof of Theorem 1.

Let $I_x = [x, 2x]$ and first observe that

$$\begin{aligned}
\#\{n \in I_x : \text{ there exists } p|n,\ p \in [\eta(x), \eta(2x)]\} &\leq \sum_{\eta(x)\leq p\leq\eta(2x)} \left(\left\lfloor\frac{2x}{p}\right\rfloor - \left\lfloor\frac{x}{p}\right\rfloor\right) \\
&\leq c\,x \sum_{\eta(x)\leq p\leq\eta(2x)} \frac{1}{p}
\end{aligned}$$

$$= \quad o(1) \qquad (x \to \infty).$$

This means that with the exception of $o(x)$ integers $n \in I_x$, the number $Q(n)$ is the smallest prime divisor of n bigger than $\eta(x)$.

Secondly, observe that we may assume that, given any fixed small $\varepsilon > 0$, we may assume that $Q(n) \le \eta(x)^{1/\varepsilon}$. Indeed,

$$\#\{n \in I_x : Q(n) > \eta(x)^{1/\varepsilon}\} \ll x \prod_{\eta(x) < p \le \eta(x)^{1/\varepsilon}} \left(1 - \frac{1}{p}\right) \ll \varepsilon x. \tag{4.1}$$

Now let $p_0, p_1, \ldots, p_{k-1}$ be any distinct primes belonging to the interval $(\eta(x), \eta(x)^{1/\varepsilon})$, and let $p_0^* < p_1^* < \cdots < p_{k-1}^*$ be the unique permutation of the primes $p_0, p_1, \ldots, p_{k-1}$, namely the one such that has all its members appear in increasing order, so that we have

$$\eta(x) < p_0^* < p_1^* < \cdots < p_{k-1}^* < \eta(x)^{1/\varepsilon}.$$

Our first goal will be to estimate the size of

$$N(x|p_0, p_1, \ldots, p_{k-1}) := \#\{n \le x : Q(n+j) = p_j,\ j = 0, 1, \ldots, k-1\}.$$

We must therefore estimate the number of those integers $n \in I_x$ for which $p_j | n + j$ $(j = 0, 1, \ldots, k-1)$, while at the same time $(\pi_j, n+j) = 1$ if $\eta(x) < \pi_j < p_j$ $(j = 0, 1, \ldots, k-1)$. Before moving on, let us set

$$Q_k = p_0 p_1 \cdots p_{k-1} \quad \text{and} \quad T_j = \prod_{\eta(x) < \pi < p_j} \pi \qquad (j = 0, 1, \ldots, k-1).$$

It is then easy to see that, say by using the Eratosthenian sieve (see for instance Chapter 12 in the book of De Koninck and Luca [2]), we obtain

$$N(x|p_0, p_1, \ldots, p_{k-1}) = (1 + o(1)) \frac{x}{Q_k} \Sigma_0 \qquad (x \to \infty), \tag{4.2}$$

where

$$\Sigma_0 = \sum_{\substack{\delta_0, \ldots, \delta_{k-1} \\ \delta_j | T_j \ (j=0,1,\ldots,k-1) \\ (\delta_i, \delta_j) = 1 \ \text{if} \ i \ne j}} \frac{\mu(\delta_0) \cdots \mu(\delta_{k-1})}{\delta_0 \cdots \delta_{k-1}}$$

(here μ stands for the Möbius function.) One can see that

$$\begin{aligned} \Sigma_0 &= \prod_{\eta(x) < \pi < p_0^*} \left(1 - \frac{k}{\pi}\right) \cdot \prod_{p_0^* < \pi < p_1^*} \left(1 - \frac{k-1}{\pi}\right) \cdots \prod_{p_{k-2}^* < \pi < p_{k-1}^*} \left(1 - \frac{1}{\pi}\right) \\ &= (1 + o(1)) \left(\frac{\log p_0^*}{\log \eta(x)}\right)^{-k} \left(\frac{\log p_1^*}{\log p_0^*}\right)^{-k+1} \cdots \left(\frac{\log p_{k-1}^*}{\log p_{k-2}^*}\right)^{-1}. \end{aligned} \tag{4.3}$$

Hence, if we set $\sigma(p) := \dfrac{\log \eta(x)}{\log p}$, it follows from (4.3) that

$$\Sigma_0 = (1+o(1))\sigma(p_0)\cdots\sigma(p_{k-1}) \qquad (x\to\infty). \tag{4.4}$$

Substituting (4.4) in (4.2), we obtain

$$N(x|p_0,p_1,\ldots,p_{k-1}) = (1+o(1))x\prod_{j=0}^{k-1}\frac{\sigma(p_j)}{p_j} \qquad (x\to\infty), \tag{4.5}$$

an estimate which holds uniformly for $\eta(x)\le p_j\le \eta(x)^{1/\varepsilon}$ $(j=0,1,\ldots,k-1)$.

We will now use a technique which we first used in [3] to study the distribution of subsets of primes in the prime factorization of integers. We first introduce the sequence

$$u_0=\eta(x), \qquad u_{j+1}=u_j+\frac{u_j}{\log^2 u_j} \quad \text{for each } j=0,1,2,\ldots$$

and then let T be the unique positive integer satisfying $u_{T-1}<\eta(x)^{1/\varepsilon}\le u_T$. Then, consider the intervals

$$J_0 := [u_0,u_1), \quad J_1 := [u_1,u_2), \ldots, \quad J_{T-1} := [u_{T-1},u_T).$$

Choose k arbitrary integers $j_0,\ldots,j_{k-1}\in\{0,1,\ldots,T-1\}$, as well as k arbitrary integers $i_0,\ldots,i_{k-1}$ from the set $\{0,1,\ldots,q-1\}$, and consider the quantity

$$M\left(x \,\middle|\, \begin{matrix} j_0,j_1,\ldots,j_{k-1}\\ i_0,i_1,\ldots,i_{k-1}\end{matrix}\right) = \sum_{p_\ell\in J_\ell\cap\wp_{i_\ell}} N(x|p_0,\ldots,p_{k-1}). \tag{4.6}$$

Observe that $\dfrac{\sigma(p_h)}{p_h}=(1+o(1))\dfrac{\sigma(u_h)}{u_h}$ as $x\to\infty$ if $p\in J_h$. It follows from this observation and using (4.5) and (4.6) that

$$M\left(x \,\middle|\, \begin{matrix} j_0,j_1,\ldots,j_{k-1}\\ i_0,i_1,\ldots,i_{k-1}\end{matrix}\right) = (1+o(1))x\sum_{p_\ell\in J_\ell\cap\wp_{i_\ell}}\prod_{j=0}^{k-1}\frac{\sigma(u_j)}{u_j}. \tag{4.7}$$

Using Theorem 1 of our 1995 paper [3] in combination with (4.7), we obtain that

$$M\left(x \,\middle|\, \begin{matrix} j_0,j_1,\ldots,j_{k-1}\\ i_0,i_1,\ldots,i_{k-1}\end{matrix}\right) = (1+o(1))M\left(x \,\middle|\, \begin{matrix} j_0,j_1,\ldots,j_{k-1}\\ i'_0,i'_1,\ldots,i'_{k-1}\end{matrix}\right) \qquad (x\to\infty), \tag{4.8}$$

where $(i'_0,i'_1,\ldots,i'_k)$ is any arbitrary sequence of length k composed of integers from the set $\{0,\ldots,q-1\}$.

Finally, consider the expression

$$A_x := \kappa(Q(\lfloor x\rfloor))\ldots\kappa(Q(\lfloor 2x\rfloor-1)).$$

It follows from (4.8) that, for any given word $\beta\in A_q^k$, the number of occurrences of β as a subword in the word A_x is equal to $(1+o(1))\dfrac{x}{q^k}$ as $x\to\infty$, thus completing the proof of Theorem 2.

5 Proof of Theorem 1

Let

$$B_x = \overline{Q(\lfloor x \rfloor)} \ldots \overline{Q(\lfloor 2x \rfloor - 1)}.$$

Also, let $Q^*(n) = \min\limits_{\substack{p|n \\ p>\eta(x)}} p$ and observe that $Q^*(n) \le Q(n)$, while if $Q^*(n) \ne Q(n)$, then $p|n$ if $\eta(x) < p < \eta(2x)$.

Moreover, let

$$B_x^* = \overline{Q^*(\lfloor x \rfloor)} \ldots \overline{Q^*(\lfloor 2x \rfloor - 1)}.$$

Clearly, we have, since $\eta(x)$ was chosen to be a slowly oscillating function,

$$0 \le \lambda(B_x) - \lambda(B_x^*) \le cx \sum_{\eta(x)<p<\eta(2x)} \frac{\log p}{\log q} \le c_1 x \log \frac{\eta(2x)}{\eta(x)} = o(x) \qquad (x \to \infty). \tag{5.1}$$

It follows from (5.1) that we now only need to estimate $\lambda(B_x^*)$. To do so, we first let δ_x be a function tending to 0 very slowly as $x \to \infty$, in a manner specified below. If $p < x^{\delta_x}$, we have

$$\begin{aligned} R_p(x) := \#\{n \in I_x : Q^*(n) = p\} &= (1+o(1)) \frac{x}{p} \prod_{\eta(x)<\pi<p} \left(1 - \frac{1}{\pi}\right) \\ &= (1+o(1)) \frac{x}{p} \frac{\log \eta(x)}{\log p} \qquad (x \to \infty), \end{aligned} \tag{5.2}$$

while on the other hand, if $x^{\delta_x} \le p \le 2x$, we have

$$R_p(x) < c \frac{x}{p} \frac{\log \eta(x)}{\log p}. \tag{5.3}$$

Now, observe that, as $x \to \infty$,

$$\begin{aligned} \lambda(B_x^*) &= \sum_{\eta(x)<p\le 2x} R_p(x)\lambda(p) = \sum_{\eta(x)<p\le 2x} R_p(x) \left\lfloor \frac{\log p}{\log q} \right\rfloor \\ &= (1+o(1)) \frac{x}{\log q} \sum_{\eta(x)<p\le 2x} \frac{\log \eta(x)}{p} + O\left(x \log \eta(x) \sum_{x^{\delta_x}<p\le x} \frac{1}{p} \right) \\ &= (1+o(1)) x \frac{\log \eta(x)}{\log q} \log \frac{\log x}{\log \eta(x)} + O\left(x \log \eta(x) \log \frac{1}{\delta_x} \right). \end{aligned} \tag{5.4}$$

Choosing the function δ_x in such a way that

$$\log \frac{1}{\delta_x} = o\left(\log \frac{\log x}{\log \eta(x)} \right)$$

allows us to replace (5.4) with

$$\lambda(B_x^*) = (1+o(1))x\frac{\log\eta(x)}{\log q}\log\frac{\log x}{\log\eta(x)} \qquad (x\to\infty). \tag{5.5}$$

Now, let $\beta_1, \beta_2 \in A_q^k$. We will now make use of Lemmas 1 and 2. For this, we first write

$$[\eta(x), x^{\delta_x}] = \bigcup_{j=0}^{T} I_{u_j},$$

where

$$I_{u_j} = [u_j, u_{j+1}), \qquad \text{with } u_0 = \eta(x), \qquad u_j = 2^j\eta(x) \qquad \text{for } j = 1, 2, \ldots, T+1,$$

where T is defined as the unique positive integer satisfying $u_T < x^{\delta_x} \le u_{T+1}$.

In the spirit of Lemma 1, we will say that the prime $p \in I_u$ is a *bad prime* if

$$\max_{\beta\in A_q^\ell}\left|\nu_\beta(\overline{p}) - \frac{L(u)}{q^\ell}\right| > \kappa_u\sqrt{L(u)}$$

and a *good prime* if

$$\left|\nu_\beta(\overline{p}) - \frac{L(u)}{q^\ell}\right| \le \kappa_u\sqrt{L(u)}.$$

We will now separate the sum $\sum R_p(x)\lambda(p)$ running over the primes p located in the intervals $[u_j, u_{j+1})$ into two categories, namely the bad primes and the good primes.

First, using (5.2) and (5.3), we have

$$\sum_{\substack{p\in[u_j,u_{j+1})\\ p \text{ bad}}} R_p(x)\lambda(p) \le c\kappa(u_j)\sum_{p\in[u_j,u_{j+1})}\frac{x\log\eta(x)}{p\log p} \ll x\frac{\log\eta(x)}{\log\eta(x)+j\log 2}. \tag{5.6}$$

On the other hand, if p is a good prime, one can easily establish that the number of occurrences of the words β_1 and β_2 in the word B_x^* are close to each other, in the sense that

$$\nu_{\beta_1}(B_x^*) - \nu_{\beta_2}(B_x^*) = o(\lambda(B_x^*)). \tag{5.7}$$

Hence, proceeding as in [5], it follows, considering the true size of $\lambda(B_x^*)$ given by (5.5) and in light of (5.1), (5.6) and (5.7), that the number of words $\beta \in A_q^k$ appearing in B_x is equal to $(1+o(1))\dfrac{\lambda(B_x)}{q^k}$ as $x\to\infty$.

We then proceed in a same manner to obtain similar estimates successively for the intervals $I_{x/2}, I_{x/2^2}, \ldots$ Thus, repeating the argument used in [5], Theorem 1 follows immediately.

The proofs of Theorems 3 through 7 can be obtained along the same lines and will therefore be omitted.

6 Proof of Theorem 8

To prove this theorem, we will consider two cases separately.

Let us first assume that

$$\sum_p \frac{\Re(1-f(p)p^{-i\tau})}{p} < \infty \quad \text{for some real number } \tau. \tag{6.1}$$

It can be proved (as we did in [6]) that one can assume that $\tau = 0$.

For a start, define the additive function u implicitly on prime powers by $f(p^\beta) = e^{iu(p^\beta)}$. Then, for each large number D, define the multiplicative function f_D on prime powers by

$$f_D(p^\beta) = \begin{cases} f(p^\beta) & \text{if } p \le D, \\ 1 & \text{if } p > D. \end{cases}$$

In light of (6.1), we have that

$$\sum_p \frac{u^2(p)}{p} < \infty. \tag{6.2}$$

Further set

$$a_D(x) = \sum_{D<p\le x} \frac{u(p)}{p-1}, \qquad b_D^2(x) = \sum_{D<p\le x} \frac{u^2(p)}{p}.$$

Since

$$f(n) = f_D(n) \exp\left\{ i \sum_{p^\beta \| n} u(p^\beta) \right\} = f_D(n) \exp\left\{ iu_D(n) \right\},$$

say, then, by using the Turán-Kubilius inequality, we obtain that

$$A(x) - A_D(x) = O\left(x b_D(x) \right),$$

where

$$A_D(x) = \eta_D(x) \sum_{n\le x} f_D(n) e(\alpha Q(n)),$$

where $\eta_D(x) = e^{ia_D(x)}$.

Further define the function τ_D implicitly by the equation $f_D(n) = \sum_{d|n} \tau_D(d)$. It is clear that $\tau_D(d) = 0$ if $(d, D) > 1$, while $|\tau_D(p^\beta)| \le 2$ for all prime powers p^β.

We clearly have

$$A_D(x) = \eta_D(x) \sum_{P(d)\le D} \tau_D(d) \sum_{md\le x} e(\alpha Q(md)) = \eta_D(x) \sum_{P(d)\le D} \tau_D(d) \Sigma_d, \tag{6.3}$$

say. On the other hand,

$$\frac{1}{x} \sum_{P(d)\le D} |\tau_D(d)|\, |\Sigma_d| \le \sum_{P(d)\le D} \frac{|\tau_D(d)|}{d} \le \prod_{p\le D} \left(1 + \frac{2}{p-1} \right).$$

Therefore, for some k_D, we have

$$\frac{1}{x}\sum_{d>k_D}|\tau_D(d)||\Sigma_d| \le \rho_D,$$

where $\rho_D \to 0$ as $D\to\infty$.

Let us now consider the sum

$$T_Y = \sum_{Y\le m\le 2Y} e(\alpha Q(m)). \tag{6.4}$$

Recall that $Q(m)$ is the smallest prime divisor of m which is larger than $\eta(m)$. Now, consider the somewhat similar function $Q_1(m)$, which stands for the smallest prime divisor of m which is larger than $\eta(x)$. Recalling the argument used at the beginning of the proof of Theorem 2, we easily see that

$$\#\{m\in[Y,2Y] : Q_1(m)\ne Q(m)\} = cY\log\frac{\eta(2Y)}{\eta(Y)} = o(Y) \quad \text{as } Y\to\infty. \tag{6.5}$$

Therefore, setting

$$T_Y^{(1)} = \sum_{Y\le m\le 2Y} e(\alpha Q_1(m)),$$

it is clear that

$$\left|T_Y - T_Y^{(1)}\right| = o(Y) \qquad (Y\to\infty).$$

Moreover, as $Y\to\infty$, we have

$$\begin{aligned}\#\{m\in[Y,2Y] : Q_1(m)=p\} &= (1+o(1))\frac{Y}{p}\prod_{\eta(Y)<\pi<p}\left(1-\frac{1}{\pi}\right)\\ &= (1+o(1))\frac{Y}{p}\frac{\log\eta(Y)}{\log p}.\end{aligned} \tag{6.6}$$

Similarly as we obtained (4.1), we easily prove that

$$\#\{m\in[Y,2Y] : Q(m) > \eta(Y)^{1/\varepsilon}\} \ll \varepsilon Y. \tag{6.7}$$

On the other hand, using (6.4), (6.6) and (4.1), we have

$$T_Y = Y\sum_{\eta(Y)<p<\eta(Y)^{1/\varepsilon}}\frac{e(\alpha p)\log\eta(Y)}{p\log p} + O(\varepsilon Y). \tag{6.8}$$

By using the well known I.M. Vinogradov theorem [10] asserting that

$$\lim_{x\to\infty}\frac{1}{\pi(x)}\sum_{p\le x} e(\alpha p) = 0,$$

we obtain from (6.8) that

$$\left| \frac{T_Y}{Y} \right| \le \varepsilon + o(1) \qquad (Y \to \infty). \tag{6.9}$$

Using this, we can estimate Σ_d. Indeed, we have

$$|\Sigma_d| \le \left| \sum_{\frac{x}{2^L d} < m < \frac{x}{d}} e(\alpha Q(dm)) \right| + \frac{x}{2^L d}. \tag{6.10}$$

Let ℓ_D be an arbitrary large number and choose L so that

$$\eta\left(\frac{x}{2^L d}\right) > \ell_D.$$

Note that for an arbitrary large L, this inequality will hold provided x is large enough. Applying (6.9), it follows from (6.10) that

$$|\Sigma_d| \le \frac{x}{2^L d} + c\varepsilon \frac{x}{d}. \tag{6.11}$$

Using (6.11) in (6.3), we obtain that

$$|A_D(x)| \le x\left(c\varepsilon + \frac{1}{2^L}\right) \prod_{p \le D} \left(1 + \frac{2}{p-1}\right) + x \sum_{d > \ell_D} \frac{|\tau_D(d)|}{d}. \tag{6.12}$$

Since D and L were chosen to be arbitrary numbers, it follows from (6.12) that

$$\lim_{x \to \infty} \frac{A_D(x)}{x} = 0. \tag{6.13}$$

Since

$$\frac{A(x)}{x} = \frac{A_D(x)}{x} + O\left(b_D(x)\right)$$

and recalling the definition of $b_D(x)$ and estimate (6.2), it follows from (6.13) that

$$\limsup_{x \to \infty} \frac{A(x)}{x} \le c b_D(x) = o(1),$$

so that if $D \to \infty$, we immediately obtain (2.1) for the first case, that is when (6.1) holds.

It remains to consider the case

$$\sum_p \frac{\Re(1 - f(p)p^{-i\tau})}{p} = \infty \qquad \text{for all real numbers } \tau. \tag{6.14}$$

First, it is clear that, using (6.5), we have

$$\begin{aligned} E(x) &:= \sum_{x<n\le 2x} f(n)e(\alpha Q(n)) \\ &= \sum_{x<n\le 2x} f(n)e(\alpha Q_1(n)) + \sum_{\substack{x<n\le 2x\\ Q_1(n)\ne Q(n)}} f(n)e(\alpha Q(n)) \\ &= \sum_{x<n\le 2x} f(n)e(\alpha Q_1(n)) + o(x) \\ &= E_1(x) + o(x), \end{aligned} \tag{6.15}$$

say.

In light of (6.7), we may ignore those $n \in (x, 2x]$ for which $Q_1(n) > \eta(x)^{1/\varepsilon}$, that is,

$$\sum_{\substack{x<n\le 2x\\ Q_1(n)>\eta(x)^{1/\varepsilon}}} f(n)e(\alpha Q_1(n)) \ll \varepsilon x. \tag{6.16}$$

Combining (6.15) and (6.16), we can write that

$$E(x) = \sum_{\eta(x)<p<\eta(x)^{1/\varepsilon}} f(p)e(\alpha p)\Sigma_p + O(\varepsilon x), \tag{6.17}$$

where, setting $\Pi_p := \prod_{\eta(x)<\pi<p} \pi$,

$$\Sigma_p = \sum_{\substack{\frac{x}{p}<m\le \frac{2x}{p}\\ (m,\Pi_p)=1}} f(m). \tag{6.18}$$

Now, consider the summation

$$S(x) = \sum_{n\le x} f(n).$$

In light of (6.14), it follows from a classical theorem of Halász (see [9]) that there exists a function $\varepsilon(x)$ which tends to 0 monotonically as $x \to \infty$ such that

$$\frac{|S(x)|}{x} \le \varepsilon(x),$$

which in turn implies that

$$\frac{|S(2x) - S(x)|}{x} \le \varepsilon(x). \tag{6.19}$$

From (6.18), we get that

$$\begin{aligned}\Sigma_p &= \sum_{\frac{x}{p}<m\leq\frac{2x}{p}} f(m) \sum_{\delta|(\Pi_p,m)} \mu(\delta)\\ &= \sum_{\delta|\Pi_p} \mu(\delta) \sum_{x<m\delta p\leq 2x} f(m\delta)\\ &= \sum_{\delta|\Pi_p} \mu(\delta) f(\delta) \left(S\left(\frac{2x}{\delta p}\right) - S\left(\frac{x}{\delta p}\right)\right) + Er_p, \end{aligned} \tag{6.20}$$

where Er_p is the error term coming from those terms for which $(m,\delta)>1$.

Thus, it follows from (6.19) and (6.20) that

$$|\Sigma_p| \leq \sum_{\delta|\Pi_p} \mu^2(\delta)\varepsilon\left(\frac{x}{\delta p}\right) + |Er_p| \leq \frac{x}{p}\sum_{\delta|\Pi_p}\frac{\mu^2(\delta)}{\delta} + |Er_p|, \tag{6.21}$$

where we used the fact that since $\max\limits_{\substack{\eta(x)<p<\eta(x)^{1/\varepsilon}\\ \delta|\Pi_p}} \dfrac{p\delta}{x} \to 0$ as $x\to\infty$, then $\varepsilon(x/\delta p) = o(x/\delta p)$ uniformly for $\eta(x)<p<\eta(x)^{1/\varepsilon}$ and $\delta|\Pi_p$.

Now, since

$$\sum_{\delta|\Pi_p}\frac{\mu^2(\delta)}{\delta} \leq \prod_{\eta(x)<\pi<\eta(n)^{1/\varepsilon}} \left(1+\frac{1}{\pi}\right) \leq c\frac{1}{\varepsilon},$$

it follows from (6.21) that, as $x\to\infty$,

$$|\Sigma_p| \leq \frac{cx}{p\varepsilon}\cdot o(1) + |Er_p|. \tag{6.22}$$

Using (6.22) in (6.17),we obtain that, as $x\to\infty$,

$$|E(x)| \leq \frac{cx}{\varepsilon}\left(\sum_{\eta(x)<p<\eta(x)^{1/\varepsilon}} \frac{1}{p}\right)\cdot o(1) + V(x) + O(\varepsilon x). \tag{6.23}$$

where

$$V(x) = \sum_{\eta(x)<p<\eta(x)^{1/\varepsilon}} |Er_p|.$$

We will now show that

$$V(x) = o(x) \qquad (x\to\infty). \tag{6.24}$$

Setting $J=J(x)=(\eta(x),\eta(x)^{1/\varepsilon})$ and writing those $m\delta p$ such that $(m,\delta)>1$ as $m\delta p = \ell\kappa^2\delta_1 p$, where κ and δ_1 are squarefree numbers whose prime factors all belong to J, we have that

$$V(x) \leq \sum_{\kappa\geq\eta(x)} \mu^2(\kappa) \sum_{\substack{x<\ell\kappa^2\delta_1 p\leq 2x\\ p\in J\\ \pi|\delta_1\Rightarrow \pi\in J}} \mu^2(\delta_1)$$

$$
\begin{aligned}
&= \sum_{\kappa\geq\eta(x)} \mu^2(\kappa) \sum_{\substack{p\in J\\ \pi|\delta_1\Rightarrow\pi\in J}} \mu^2(\delta_1) \sum_{\frac{x}{\kappa^2\delta_1 p}<\ell\leq\frac{2x}{\kappa^2\delta_1 p}} 1\\
&\leq cx \sum_{\kappa\geq\eta(x)} \frac{\mu^2(\kappa)}{\kappa^2} \sum_{\substack{p\in J\\ \pi|\delta_1\Rightarrow\pi\in J}} \frac{\mu^2(\delta_1)}{\delta_1 p}. \qquad (6.25)
\end{aligned}
$$

Since it is easily checked that

$$
\begin{aligned}
\sum_{p\in J}\frac{1}{p} &\leq c_1\log\frac{1}{\varepsilon},\\
\sum_{\pi|\delta_1\Rightarrow\pi\in J}\frac{\mu^2(\delta_1)}{\delta_1} &\leq \prod_{\pi\in J}\left(1+\frac{1}{\pi}\right)\leq\frac{c_2}{\varepsilon}\log\eta(x),\\
\sum_{\kappa\geq\eta(x)}\frac{1}{\kappa^2} &\leq \frac{c_3}{\eta(x)},
\end{aligned}
$$

then using these estimates in (6.25), we obtain that

$$
V(x)\leq c_4 x\frac{\log\eta(x)}{\eta(x)}\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}=o(x)\qquad (x\to\infty),
$$

thus proving our claim (6.24).

Substituting (6.24) in (6.23), we obtain that

$$
|E(x)|\leq cx\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}\cdot o(1)+o(x)+O(\varepsilon x)=o(x)\qquad (x\to\infty),
$$

from which it follows that given any arbitrarily small number $\xi>0$, there is some $x_0=x_0(\xi)$ such that

$$
|E(X)|\leq\xi X\qquad \text{for all } X>x_0. \tag{6.26}
$$

Therefore, given any fixed large number x and letting L be the smallest integer such that $2^L>x/2$, we have that, using (6.26) repetitively,

$$
|A(x)|=\left|\sum_{a=1}^{L}E\left(\frac{x}{2^a}\right)\right|\leq c\xi\sum_{a=1}^{L}\frac{x}{2^a}<c\xi x,
$$

thus proving (2.1) in the second case, as requested.

This completes the proof of Theorem 8.

References

[1] N.L. Bassily and I. Kátai, *Distribution of consecutive digits in the q-ary expansions of some sequences of integers*, *Journal of Mathematical Sciences* **78**, no. 1 (1996), 11-17.

[2] J.M. De Koninck and F. Luca, *Analytic Number Theory: Exploring the Anatomy of Integers*, Graduate Studies in Mathematics, Vol. 134, American Mathematical Society, Providence, Rhode Island, 2012.

[3] J.M. De Koninck and I. Kátai, *On the distribution of subsets of primes in the prime factorization of integers*, Acta Arithmetica **72**, no.2 (1995), 169-200.

[4] J.M. De Koninck and I. Kátai, *Construction of normal numbers by classified prime divisors of integers*, Functiones et Approximatio **45.2** (2011), 231-253.

[5] J.M. De Koninck and I. Kátai, *On a problem on normal numbers raised by Igor Shparlinski*, *Bulletin of the Austr. Math. Soc.* **84** (2011), 337-349.

[6] J.M. De Koninck and I. Kátai, *Exponential sums involving the largest prime factor function*, Acta Arithmetica **146** (2011), 233-245.

[7] J.M. De Koninck and I. Kátai, *Construction of normal numbers using the distribution of the k-th largest prime factor*, Bull. Australian Mathematical Society, to appear.

[8] G. Halász, *Über die Mittelwerte multiplikativen zahlentheoretischer Funktionen*, Acta Math. Acad. Scient. Hungaricae **19** (1968), 365–404.

[9] H.H. Halberstam and H.E. Richert, *Sieve Methods*, Academic Press, London, 1974.

[10] I.M. Vinogradov, *The method of trigonometric sums in the theory of numbers*, translated, revised and annotated by A. Davenport and K.F.Roth, Interscience, New York, 1954.

Jean-Marie De Koninck
Dép. de mathématiques et de statistique
Université Laval
Québec
Québec G1V 0A6
Canada
jmdk@mat.ulaval.ca

Imre Kátai
Computer Algebra Department
Eötvös Loránd University
1117 Budapest
Pázmány Péter Sétány I/C
Hungary
katai@compalg.inf.elte.hu

JMDK, le 24 septembre 2012; fichier: normal-numbers-smallest-prime-divisor-2012.tex