Some new methods for constructing normal numbers

JEAN-MARIE DE KONINCK[1] and IMRE KÁTAI[2]

Édition du 18 décembre 2011

Abstract

Given an integer $q \geq 2$, a *q-normal number*, or simply a *normal number*, is a real number whose q-ary expansion is such that any preassigned sequence, of length $k \geq 1$, of base q digits from this expansion, occurs at the expected frequency, namely $1/q^k$. We expose two new methods which allow for the construction of large families of normal numbers.

AMS Subject Classification numbers: 11K16, 11N37, 11A41
Key words: normal numbers, primes, arithmetic function

1 Introduction

Given an integer $q \geq 2$, a *q-normal number*, or simply a *normal number*, is a real number whose q-ary expansion is such that any preassigned sequence, of length $k \geq 1$, of base q digits from this expansion, occurs at the expected frequency, namely $1/q^k$. Equivalently, given a positive real number $\eta < 1$ whose expansion is $\eta = 0, a_1 a_2 \ldots$ with each $a_i \in \{0, 1, \ldots, q-1\}$, that is, $\eta = \sum_{j=1}^{\infty} \frac{a_j}{q^j}$, we say that η is a normal number if the sequence $\{q^m \eta\}$, $m = 1, 2, \ldots$ (here $\{y\}$ stands for the fractional part of y), is uniformly distributed in the interval $[0, 1[$.

It is easily seen that η is a q-normal number if and only if

$$\lim_{N \to \infty} \frac{1}{N} \#\{j < N : a_{j+1} \ldots a_{j+k} = b_1 \ldots b_k\} = \frac{1}{q^k}$$

for every $b_1 \ldots b_k \in \{0, 1, \ldots, q-1\}^k$.

The problem of determining if a given number is normal is unresolved. For instance, fundamental constants such as π, e, $\sqrt{2}$, $\log 2$ as well as the famous Apéry constant $\zeta(3)$, have not yet been proven to be normal numbers, although numerical evidence tends to indicate that they are. Interestingly, Borel [2] has shown that almost all real numbers are normal, that is that the set of those real numbers which are not normal has Lebesgue measure 0.

In this paper, we expose two new methods which allow the construction of large families of normal numbers.

[1]Research supported in part by a grant from NSERC.
[2]Research supported by a grant from the European Union and the European Social Fund.

2 Notations

Given an integer $t \geq 1$, an expression of the form $i_1 i_2 \dots i_t$, where each i_j is one of the numbers $0, 1, \dots, q-1$, is called a *word of length* t. Given a word α, we shall write $\lambda(\alpha) = t$ to indicate that α is a *word* of length t. We shall also use the symbol Λ to denote the *empty word*.

Let $q \geq 2$ be a fixed integer and let $A = A_q = \{0, 1, 2, \dots, q-1\}$. Then, $A^t = A_q^t$ will stand for the set of words of length t over A, while $A^* = A_q^*$ will stand for the set of words over A, including the empty word Λ, that is

$$A^* = A_q^* = \bigcup_{t=0}^{\infty} A^t, \qquad \text{where } A^0 = \{\Lambda\}.$$

Moreover, the concatenation of two words $\alpha, \beta \in A^*$, written $\alpha\beta$, also belongs to A^*. It is clear that $\lambda(\alpha\beta) = \lambda(\alpha) + \lambda(\beta)$.

Given a fixed integer $q \geq 2$, we will write positive integers n as

$$n = \sum_{j \geq 0} \varepsilon_j(n) q^j, \text{ with each } \varepsilon_j(n) \in A_q, \tag{2.1}$$

where the above sum is clearly finite, and use the notation

$$\overline{n} = \varepsilon_0(n)\varepsilon_1(n)\dots\varepsilon_t(n), \tag{2.2}$$

where $\varepsilon_t(n) \neq 0$.

Let k be a fixed positive integer. For each word $\beta = b_1 \dots b_k \in A^k$, we let $\nu_\beta(\overline{n})$ stand for the number of occurrences of β in the representation (2.2) of the positive integer n, that is, the number of times that $\varepsilon_j(n)\dots\varepsilon_{j+k-1}(n) = \beta$ as j varies from 0 to $t-(k-1)$.

Let $\eta_\infty = \varepsilon_1\varepsilon_2\varepsilon_3\dots$, where each ε_i is an element of A_q. Let $\eta_N = \varepsilon_1\varepsilon_2\dots\varepsilon_N$. Moreover, for each $\beta = \delta_1\dots\delta_k \in A_q^*$ and integer $N \geq 2$, let $M(N, \beta)$ stand for the number of occurrences of β as a subsequence of the consecutive digits of η_N, that is

$$M(N, \beta) = \#\{(\alpha, \gamma) : \eta_N = \alpha\beta\gamma,\ \alpha\gamma \in A_q^*\}.$$

We will say that η_∞ is a *normal sequence* if

$$\lim_{N\to\infty} \frac{M(N, \beta)}{N} = \frac{1}{q^{\lambda(\beta)}} \quad \text{for all } \beta \in A_q^*. \tag{2.3}$$

Let $\wp$ stand for the set of all prime numbers and let $\widetilde{\wp}$ stand for an infinite subset of $\wp$. We shall denote by $\mathcal{N}(\widetilde{\wp})$ the multiplicative semigroup generated by $\widetilde{\wp}$.

Let ξ be the real number belonging to the interval $[0, 1]$ whose q-ary expansion is

$$\xi = 0.\varepsilon_1\varepsilon_2\varepsilon_3\dots$$

and, for each integer $N \geq 1$, set

$$\xi_N = 0.\varepsilon_1\varepsilon_2 \ldots \varepsilon_N.$$

With β and $M(N,\beta)$ as above, we will say that ξ is normal if (2.3) holds.

We shall let $\omega(n)$ stand for the number of distinct prime factors of n. Throughout this paper, the letter p, with or without subscripts, always denotes a prime number. Finally, the letters c and C always denote positive constants, but not necessarily the same at each occurrence.

3 First method

Let $\mathcal{B}$ be an infinite set of positive integers and let $B(x) = \#\{b \leq x : b \in \mathcal{B}\}$. Further, let $F : \mathcal{B} \to \mathbb{N}$ be a function for which, for some positive integer r and constants $0 < c_1 < c_2 < +\infty$,

$$c_1 \leq \frac{F(b)}{b^r} \leq c_2 \qquad \text{for all } b \in \mathcal{B}.$$

Let x be a large number and set $N = \left\lfloor \dfrac{\log x}{\log q} \right\rfloor$.

Let $0 \leq \ell_1 < \cdots < \ell_h$ $(\leq rN)$ be integers and let $a_1, \ldots, a_h \in A_q$. Using the notation given in (2.1) and (2.2), we further let

$$\mathcal{B}_F\left(x \,\middle|\, \begin{matrix} \ell_1, \ldots, \ell_h \\ a_1, \ldots, a_h \end{matrix}\right) = \{b \leq x : b \in \mathcal{B},\ \varepsilon_{\ell_j}(F(b)) = a_j, j = 1, \ldots, h\}$$

and

$$B_F\left(x \,\middle|\, \begin{matrix} \ell_1, \ldots, \ell_h \\ a_1, \ldots, a_h \end{matrix}\right) = \#\mathcal{B}_F\left(x \,\middle|\, \begin{matrix} \ell_1, \ldots, \ell_h \\ a_1, \ldots, a_h \end{matrix}\right).$$

We say that $F(\mathcal{B})$ is a *q-ary smooth sequence* if there exists a positive constant $\alpha < 1$ and a function $\varepsilon(x)$, which tends to 0 as x tends to infinity, such that for every fixed integer $h \geq 1$,

$$\sup_{N^\alpha \leq \ell_1 < \cdots < \ell_h \leq rN - N^\alpha} \left| \frac{q^h B_F\left(x \,\middle|\, \begin{matrix} \ell_1, \ldots, \ell_h \\ a_1, \ldots, a_h \end{matrix}\right)}{B(x)} - 1 \right| \leq c(h)\varepsilon(x) \tag{3.1}$$

(where $c(h)$ is a positive constant depending only on h) and also such that $B(x) \gg \dfrac{x}{\log x}$.

Now, let $F(\mathcal{B})$ be a q-ary smooth sequence. Let $b_1 < b_2 < \cdots$ stand for the list of all elements of $\mathcal{B}$. Let also

$$\xi_n = \overline{F(b_n)} = \varepsilon_0(F(b_n)) \ldots \varepsilon_t(F(b_n)).$$

Since $c_1 b_n^r \le F(b_n) \le c_2 b_n^r$, it follows that $t = \dfrac{r \log\log b_n}{\log q} + O(1)$.

Let $n \in I := [x, 2x]$, $\beta \in A_q^k$, $\nu_\beta(\xi_n)$ be the number of occurrences of β as a subword in ξ_n. Let us write

$$\nu_\beta(\xi_n) = \nu_\beta^{(1)}(\xi_n) + \nu_\beta^{(2)}(\xi_n) + \nu_\beta^{(3)}(\xi_n),$$

where $\nu_\beta^{(1)}(\xi_n)$ is the number of those occurrences of β in ξ_n as $\xi_n = \delta\beta\gamma$, with $\lambda(\delta) \le N^\alpha$, while $\nu_\beta^{(3)}(\xi_n)$ is the number of those occurrences of β in ξ_n for which $\lambda(\gamma) \le N^\alpha$.

Now, from (3.1), we obtain that

$$\begin{aligned} \Sigma_1 &:= \sum_{n\in I} \nu_\beta^{(2)}(\xi_n) = \frac{1}{q^k} \sum_{n \in I} (\lambda(\xi_n) + O(N^\alpha)) \\ &= \frac{1}{q^k} rNx + O(x(\log x)^\alpha). \end{aligned} \tag{3.2}$$

Similarly, one can estimate the expression

$$\Sigma_2 := \sum_{n\in I} (\nu_\beta^{(2)}(\xi_n))^2.$$

Indeed, $(\nu_\beta^{(2)}(\xi_n))^2$ stands for the number of solutions of

$$\xi_n = \delta_1\beta\gamma_1, \quad \xi_n = \delta_2\beta\gamma_2 \qquad \text{such that } \lambda(\delta_j) > N^\alpha,\ \lambda(\gamma_j) \le N^\alpha \ (j = 1, 2).$$

Observe that the number of those occurrences of β which are not disjoint, say

$$\xi_n = \boxed{\quad \delta_1 \quad | \quad \beta \quad | \quad \gamma_1 \quad}$$
$$\xi_n = \boxed{\quad \delta_2 \qquad | \quad \beta \quad |\gamma_2 }$$

is less than $(2k+1)\nu_\beta^{(2)}(\xi_n)$.

Thus,

$$\Sigma_2 = \frac{1}{q^{2k}} (rN)^2 x + O(xN). \tag{3.3}$$

It follows from (3.2) and (3.3) that

$$\sum_{n\in I} \left(\nu_\beta^{(2)}(\xi_n) - \frac{1}{q^k} rN \right)^2 \ll xN$$

so that

$$\sum_{n\in I} \left(\nu_\beta(\xi_n) - \frac{1}{q^k} rN \right)^2 \ll xN + O\left(xN^{2\alpha}\right).$$

We have thus established the following result.

Theorem 1. *Let $F(\mathcal{B})$ be a q-ary smooth sequence. Let $b_1 < b_2 < \cdots$ stand for the list of all elements of $\mathcal{B}$. Let also*

$$\xi_n = \overline{F(b_n)} = \varepsilon_0(F(b_n)) \ldots \varepsilon_t(F(b_n))$$

and set

$$\eta = 0.\xi_1 \xi_2 \ldots .$$

Consider η as the real number whose q-ary expansion is the concatenation of the numbers $\xi_1, \xi_2, \ldots$. Then η is a q-normal number.

Since one can easily construct plenty of q-ary smooth sequences, Theorem 1 allows for the construction of many q-ary normal numbers. In this context, we formulate the following two remarks.

Remark 1. *Assume that $P \in \mathbb{Z}[x]$ with $r = deg(P) \geq 1$ and that $\lim_{x\to\infty} P(x) = +\infty$. Bassily and Kátai [1] proved, using theorems of I.M. Vinogradov [8] and of L.K. Hua [5] concerning the estimation of trigonometric sums, that $P(\mathbb{N})$ and $P(\wp)$ are both smooth sequences. Nakai and Shiokawa [6] explicitly proved that, in this case, η is a normal number. They even estimated the discrepancy of the sequence $\{q^m \eta\}_{m\geq 1}$.*

Remark 2. *Recently, German and Kátai [4] showed the following:*

> *Let $n_1 < n_2 < \cdots$ stand for the list of all elements of $\mathcal{N}(\widetilde{\wp})$. Moreover, let $N(x) := \#\{n \leq x : n \in \mathcal{N}(\widetilde{\wp})\}$ and let $P \in \mathbb{Z}[x]$. Assume also that $P(n)$ takes only non negative values. Then $\eta = 0.\overline{P(n_1)}\,\overline{P(n_2)} \ldots$ is a normal number.*

Much more is true. Indeed, we can also obtain the following result.

Theorem 2. *Let $n_1 < n_2 < \cdots$ be a sequence of integers such that $\#\{n_j \leq x\} > \rho x$ provided $x > x_0$, for some positive constant ρ. Then, using the notations of Theorem 1, let*

$$\mu = 0.\xi_{n_1} \xi_{n_2} \ldots$$

Then μ is a q-normal number.

4 Second method

Theorem 3. *Let $q \geq 2$ be a fixed integer. Given a positive integer $n = p_1^{e_1} \cdots p_{k+1}^{e_{k+1}}$ with primes $p_1 < \cdots < p_{k+1}$ and positive exponents $e_1, \ldots, e_{k+1}$, we introduce the numbers $c_1(n), \ldots, c_k(n)$ defined by*

$$c_j(n) := \left\lfloor \frac{q \log p_j}{\log p_{j+1}} \right\rfloor \in A_q \qquad (j = 1, \ldots, k)$$

and define the arithmetic function

$$H(n) = \begin{cases} c_1(n)\dots c_k(n) & \text{if } \omega(n) \ge 2, \\ \Lambda & \text{if } \omega(n) \le 1. \end{cases}$$

Consider the number

$$\xi = 0.H(1)H(2)H(3)\dots$$

Then ξ is a q-normal number.

Proof. As we will see, this theorem is an easy consequence of a variant of the Turán-Kubilius inequality.

Let $b_1, \dots, b_k$ be fixed elements of A_q. Then, for each sequence of $k+1$ primes $p_1 < \dots < p_{k+1}$, define the function

$$f(p_1, \dots, p_{k+1}) = \begin{cases} 1 & \text{if } \left\lfloor \frac{q \log p_j}{\log p_{j+1}} \right\rfloor = b_j \text{ for each } j \in \{1, \dots, k\}, \\ 0 & \text{otherwise.} \end{cases}$$

From this, we define the arithmetic function F as follows. If $n = q_1^{\alpha_1} \cdots q_\mu^{\alpha_\mu}$, where $q_1 < \dots < q_\mu$ are prime numbers and $\alpha_1, \dots, \alpha_\mu \in \mathbb{N}$, let

$$F(n) = F(n|b_1, \dots, b_k) = \sum_{j=0}^{\mu-k-1} f(q_{j+1}, \dots, q_{j+k+1}).$$

We will now show that $F(n)$ is close to $\frac{1}{q^k}\omega(n)$ for almost all positive integers n.

Let $Y_x = \exp\exp\{\sqrt{\log\log x}\}$ and $Z_x = x/Y_x$ and further set

$$\begin{aligned} F_0(n) &= \sum_{q_{j+1} \le Y_x} f(q_{j+1}, \dots, q_{j+k+1}), \\ F_1(n) &= \sum_{Y_x < q_{j+1} \le Z_x} f(q_{j+1}, \dots, q_{j+k+1}), \\ F_2(n) &= \sum_{q_{j+1} > Z_x} f(q_{j+1}, \dots, q_{j+k+1}), \end{aligned}$$

so that

$$F(n) = F_0(n) + F_1(n) + F_2(n). \tag{4.1}$$

Now, it is clear that

$$F_0(n) \le \omega_{Y_x}(n) := \sum_{\substack{p|n \\ p \le Y_x}} 1$$

and that

$$F_2(n) \le \sum_{\substack{p|n \\ p > Z_x}} 1.$$

Therefore,

$$\sum_{n\le x} F_0(n) \le x \sum_{p\le Y_x} \frac{1}{p} \le cx\sqrt{\log\log x} \tag{4.2}$$

and

$$\sum_{n\le x} F_2(n) \le x \sum_{Z_z<p\le x} \frac{1}{p} \le cx \log\left(\frac{\log x}{\log Z_x}\right) \ll cx\frac{e^{\sqrt{\log\log x}}}{\log x}. \tag{4.3}$$

We now move on to estimate $\sum_{n\le x}(F_1(n)-A)^2$ for a suitable expression A, which shall be given explicitly later.

We first write this sum as follows:

$$\begin{aligned}\sum_{n\le x}(F_1(n)-A)^2 &= \sum_{n\le x} F_1(n)^2 - 2A\sum_{n\le x} F_1(n) + A^2\lfloor x\rfloor \\ &= S_1 - 2AS_2 + A^2\lfloor x\rfloor, \end{aligned} \tag{4.4}$$

say.

Let $Y_x < p_1 < \cdots < p_{k+1}$. We say that $p_1,\ldots,p_{k+1}$ is a *chain of prime divisors* of n, which we note as $p_1 \mapsto p_2 \mapsto \ldots \mapsto p_{k+1}|n$, if $\gcd\left(\frac{n}{p_1\cdots p_{k+1}}, p\right) = 1$ for all primes p in the interval $[p_1, p_{k+1}]$ with the possible exception of the primes p belonging to the set $\{p_1,\ldots,p_{k+1}\}$.

Observe that the contribution to the sums S_1 and S_2 of those positive integers $n \le x$ for which $p^2|n$ for some prime p is small, since the contribution of those particular integers $n\le x$ is less than

$$cxk\sum_{p>Y_x}\frac{1}{p^2} \le \frac{cxk}{Y_x} = o(x).$$

Hence we can assume that the sums S_1 and S_2 run only over squarefree integers n.

We now introduce the function

$$\Gamma(u,v) := \prod_{\substack{p\in\wp \\ u\le p<v}}\left(1-\frac{1}{p}\right)$$

and observe that it follows from Theorem 5.3 of Prachar [7] that

$$\Gamma(u,v) = \frac{\log u}{\log v}\left(1+O\left(\exp\{-\sqrt{\log u}\}\right)\right). \tag{4.5}$$

Now, using Lemma 2.1 from Elliott [3], one can establish that

$$\#\left\{\nu \le \frac{x}{p_1\cdots p_{k+1}} : \gcd\left(\nu, \prod_{p_1\le p\le p_{k+1}} p\right) = 1\right\}$$

$$= \frac{x}{p_1 \cdots p_{k+1}} \Gamma(p_1, p_{k+1}) \left(1 + O\left(\log^{-C} p_1\right)\right), \tag{4.6}$$

where C is an arbitrary but fixed positive constant.

It follows from (4.6) using (4.5) that

$$\begin{aligned} S_2 &= x \sum_{\substack{p_1 < \cdots < p_{k+1} \le x \\ Y_x < p_1 < Z_x}} \frac{f(p_1, \ldots, p_{k+1})}{p_1 \cdots p_{k+1}} \Gamma(p_1, p_{k+1}) \\ &\qquad + O\left(\sum_{\substack{p_1 < \cdots < p_{k+1} \le x \\ Y_x < p_1 < Z_x}} \frac{f(p_1, \ldots, p_{k+1})}{p_1 \cdots p_{k+1}} \frac{\Gamma(p_1, p_{k+1})}{\log^C p_1} \right) \\ &= x \sum_{\substack{p_1 < \cdots < p_{k+1} \le x \\ Y_x < p_1 < Z_x}} \frac{f(p_1, \ldots, p_{k+1})}{p_1 \cdots p_{k+1}} \frac{\log p_1}{\log p_{k+1}} \\ &\qquad + O\left(\sum_{\substack{p_1 < \cdots < p_{k+1} \le x \\ Y_x < p_1 < Z_x}} \frac{f(p_1, \ldots, p_{k+1})}{p_1 \cdots p_{k+1}} \frac{1}{\log^C p_1} \right). \end{aligned} \tag{4.7}$$

In order to estimate the main term on the right hand side of (4.7), we let

$$L(x) = \sum_{\substack{p_1 < \cdots < p_{k+1} \le x \\ Y_x < p_1 < Z_x}} \frac{f(p_1, \ldots, p_{k+1})}{p_1 \cdots p_{k+1}} \frac{\log p_1}{\log p_{k+1}}$$

and we also consider the sum $L_0(x)$, that is essentially the same sum as the sum $L(x)$ but where we drop the condition $Y_x < p_1 < Z_x$ in the summation.

Note that, in light of (4.2), the error $L_0(x) - L(x)$ satisfies

$$0 \le L_0(x) - L(x) \le c \frac{1}{x} \sum_{n \le x} \omega_{Y_x}(n) \ll \sqrt{\log \log x}. \tag{4.8}$$

Now, since, for each $j \in \{1, 2, \ldots, k\}$, we have

$$\sum_{\left\lfloor \frac{q \log p_j}{\log p_{j+1}} \right\rfloor = b_j} \frac{\log p_j}{p_j} = \frac{1}{q} \log p_{j+1} + O(1),$$

it follows that, after iteration, we easily obtain that

$$L_0(x) = \frac{1}{q^k} \sum_{p_{k+1} \le x} \frac{1}{p_{k+1}} + O(1) = \frac{1}{q^k} \log \log x + O(1). \tag{4.9}$$

Now, because of (4.8), we have that $L(x) - L_0(x) = o(\log \log x)$, so that it follows from (4.9) that

$$L(x) = \frac{1}{q^k} \log \log x + O(1). \tag{4.10}$$

Substituting (4.10) in (4.7), we get

$$S_2 = \frac{1}{q^k} x \log\log x + O(x). \tag{4.11}$$

In order to estimate S_1, we proceed as follows. We have

$$\begin{aligned} S_1 &= \sum_{n\le x} F_1(n)^2 \\ &= 2\sum_{n\le x} \sum_{\substack{p_1\mapsto\cdots\mapsto p_{k+1}|n \\ q_1\mapsto\cdots\mapsto q_{k+1}|n \\ p_{k+1}<q_1}} f(p_1,\ldots,p_{k+1}) f(q_1,\ldots,q_{k+1}) + E(x), \end{aligned} \tag{4.12}$$

where the error term $E(x)$ arises from those $k+1$ tuples $\{p_1,\ldots,p_{k+1}\}$ and $\{q_1,\ldots,q_{k+1}\}$ which have common elements. One can see that the sum of $f(p_1,\ldots,p_{k+1})f(q_1,\ldots,q_{k+1})$ on such $k+1$ tuples is less than $k\omega(n)$, implying that

$$E(x) \ll x\log\log x. \tag{4.13}$$

Using the fact that

$$\begin{aligned} &\#\left\{\nu \le \frac{x}{p_1\cdots p_{k+1}q_1\cdots q_{k+1}} : \left(\nu, \prod_{p_1<p<p_{k+1}} p \times \prod_{q_1<p<q_{k+1}} p\right) = 1\right\} \\ &\qquad = x\frac{\Gamma(p_1,p_{k+1})\Gamma(q_1,q_{k+1})}{p_1\cdots p_{k+1}q_1\cdots q_{k+1}}\left(1+O\left(\frac{1}{\log^C p_1}\right)\right), \end{aligned}$$

for some positive constant C. It follows from (4.12) and (4.13), while arguing as we did for the estimation of S_2, that

$$S_1 = x\left(\sum_{\substack{p_1<\cdots<p_{k+1}\le x \\ Y_x<p_1<Z_x}} \frac{f(p_1,\ldots,p_{k+1})\Gamma(p_1,p_{k+1})}{p_1\cdots p_{k+1}}\right)^2 + O(x\log\log x). \tag{4.14}$$

Hence, in light of (4.11) and (4.14), we get that

$$S_1 = x\left(\frac{\log\log x}{q^k} + O(1)\right)^2 = x\left(\frac{\log\log x}{q^k}\right)^2 + O(x\log\log x).$$

Hence, choosing $A = \frac{1}{q^k}\log\log x$, it follows that the left hand side of (4.4) satisfies

$$\sum_{n\le x}\left(F_1(n) - \frac{1}{q^k}\log\log x\right)^2 \ll \frac{1}{q^k} x\log\log x. \tag{4.15}$$

Recall that $F_1(n)$, as well as $F(n)$, depends on $b_1, \ldots, b_k$, while A does not. Hence, setting

$$G(n) = \sum_{\{b_1,\ldots,b_k\} \in A_q^k} F(n|b_1, \ldots, b_k),$$

a sum containing q^k terms, we get that

$$\sum_{n \leq x} \left(F(n|b_1, \ldots, b_k) - \frac{G(n)}{q^k} \right)^2 \ll x \log\log x,$$

so that $\dfrac{G(n)}{q^k}$ does not depend on the choice of $(b_1, \ldots, b_k) \in A_q^k$.

Now, by using the Cauchy-Schwarz inequality along with (4.2) and (4.3), we obtain, in light of (4.1), that

$$\begin{aligned} \sum_{n \leq x} \left| F(n) - \frac{1}{q^k} x_2 \right| &\leq \sum_{n \leq x} \left| F_1(n) - \frac{1}{q^k} x_2 \right| + \sum_{n \leq x} |F_0(n)| + \sum_{n \leq x} |F_2(n)| \\ &\leq \sqrt{x} \left(\sum_{n \leq x} \left| F_1(n) - \frac{1}{q^k} x_2 \right|^2 \right)^{1/2} + O(x\sqrt{\log\log x}). \end{aligned} \tag{4.16}$$

Hence, it follows from (4.15) and (4.16) that

$$\sum_{n \leq x} \left| F(n) - \frac{1}{q^k} x_2 \right| \leq Cx\sqrt{\log\log x}. \tag{4.17}$$

Hence, given any two k-tuples $(b_1, \ldots, b_k)$ and $(b'_1, \ldots, b'_k)$ both belonging to A_q^k, it follows from (4.17) that

$$\sum_{n \leq x} |F(n|b_1, \ldots, b_k) - F(n|b'_1, \ldots, b'_k)| \leq 2Cx\sqrt{\log\log x},$$

thus implying that the probability of the occurrence of $b_1, \ldots, b_k$ in the chain of prime divisors $p_1 \mapsto \ldots \mapsto p_{k+1}|n$ is almost the same as that of the occurrence of $b'_1, \ldots, b'_k$ for any $(b'_1, \ldots, b'_k) \in A_q^k$. This final remark proves that ξ is a normal number and thus completes the proof of Theorem 3.

□

5 Final remarks

This last method can easily be applied to prove the following more general theorem.

Theorem 4. *Let $R[x] \in \mathbb{Z}[x]$, the leading coefficient of which is positive. Let m_0 be a positive integer such that $R(m) \geq 0$ for all $m \geq m_0$. Moreover, let $H(n)$ be defined as in Theorem 3 and set*

$$\xi = 0.H(R(m_0))H(R(m_0+1))H(R(m_0+2))\dots$$

Also, let $m_0 \leq p_1 < p_2 < \cdots$ be the sequence of all primes no smaller than m_0 and set

$$\eta = 0.H(R(p_1))H(R(p_2))H(R(p_3))\dots$$

Then ξ and η are q-normal numbers.

Even more is true, namely the following.

Theorem 5. *Let $(m_0 <)n_1 < n_2 < \cdots$ be a sequence of integers for which $\#\{n_j \leq x\} > \rho x$ provided $x > x_0$, for some positive constant ρ. Then, using the notations of Theorem 4, let*

$$\tau = 0.H(R(n_1))H(R(n_2))\dots$$

Then τ is a q-normal number.

Moreover, let $(m_0 <)\pi_1 < \pi_2 < \cdots$ be a sequence of primes for which $\#\{\pi_j \leq x\} > \delta\pi(x)$ provided $x > x_0$, for some positive constant δ. Let

$$\kappa = 0.H(R(\pi_1))H(R(\pi_2))\dots$$

Then κ is a q-normal number.

References

[1] N.L. Bassily and I. Kátai, *Distribution of the values of q-additive functions on polynomial sequences*, Acta Math. Hungar. **68 (4)** (1995), 353-361.

[2] E. Borel, *Les probabilités dénombrables et leurs applications arithmétiques*, Rend. Circ. Mat. Palermo **27** (1909), 247–271.

[3] P.D.T.A. Elliott, *Probabilistic Number Theory I, Mean Value Theorems*, Springer-Verlag, Berlin, 1979.

[4] L. Germán and I. Kátai, *Distribution of the values of q-additive functions on some multiplicative semigroups II* (preprint).

[5] L.K. Hua, *Additive Theory of Prime Numbers*, AMS (Providence, Rhode Island, 1965). Translations of Mathematical Monographs, Vol. 13.

[6] Y. Nakai and I. Shiokawa, *Normality of numbers generated by the values of polynomials at primes*, Acta Arith. **81** (1997), no. 4, 345-356.

[7] K. Prachar, *Primzahlverteilung*, Springer, Berlin, 1957.

[8] I.M. Vinogradov, *Method of Trigonometric Sums in the Theory of Numbers* [in Russian], Nauka, Moscow (1980).

Jean-Marie De Koninck
Département 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