Various new families of normal numbers

Jean-Marie De Koninck[1,2*] and Imre Kátai[2,3†]

[1*]Department of mathematics, Université Laval, 1045 Avenue de la médecine, Québec, G1V0A6, Québec, Canada.
[2]Computer Algebra Department, Eötvös Loránd University, Pázmány Péter Sétány I/C, Budapest, 1117, Hungary.

*Corresponding author(s). E-mail(s): jmdk@mat.ulaval.ca;
Contributing authors: katai@inf.elte.hu;
†These authors contributed equally to this work.

Abstract

After having stated several conjectures regarding potential families of normal numbers, we construct various new families of normal numbers using different concepts, in particular the whole set of partitions of the set $\{0, 1, \dots, k\}$, the smallest prime divisor of n which is larger than $\log n$, and finally a certain regroupment of the whole set of primes known as a disjoint classification of primes.

Keywords: normal numbers

MSC Classification: 11K16

1 Introduction

The concept of a normal number was introduced in 1909 by Émile Borel [1]. Given an integer $q \geq 2$, a *q-normal number*, or for short 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

$$\eta = \lfloor \eta \rfloor + \sum_{j=1}^{\infty} \frac{a_j}{q^j},$$

where each $a_j \in \{0, 1, \ldots, q-1\}$, we say that η is a *q-normal number* if for every integer $k \geq 1$ and $b_1 b_2 \ldots b_k \in \{0, 1, \ldots, q-1\}^k$, we have

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

A real number is said to be *simply normal* in a given base $q \geq 2$ if its base q digit expansion contains each of the digits $0, 1, \ldots, q-1$ at the same frequency. While all normal numbers are simply normal, the reverse is not true. For instance, the real number $0.101010\ldots$ is simply normal in base 2, but not normal.

Also, given an integer $q \geq 2$, it can be shown (see Theorem 8.1 of Chapter 1 in the book of Kuipers and Niederreiter [8]) that a real number η is normal in base q if and only if the sequence $(\{q^n \eta\})_{n \in \mathbb{N}}$ is uniformly distributed mod 1 (here, $\{y\}$ stands for the fractional part of y).

Let $q \geq 2$ be a fixed integer and set $\mathcal{A}_q := \{0, 1, 2, \ldots, q-1\}$. Given an integer $t \geq 1$, an expression of the form $i_1 i_2 \ldots i_t$, where each $i_j \in \mathcal{A}_q$, is called a *word* of length t. We shall use the symbol Λ to denote the *empty word*, which is a word of length 0.

Here, we construct various new families of normal numbers using different concepts, in particular the whole set of partitions of the set $\{0, 1, \ldots, k\}$, the smallest prime divisor of n which is larger than $\log n$, and finally the so-called disjoint classification of primes.

2 Sets of permutations

Given any two positive integers $a \neq b$, we set

$$u(a,b) := \begin{cases} 1 & \text{if } a < b, \\ 0 & \text{if } a > b. \end{cases}$$

Also, let us consider the whole set Π_{k+1} of permutations of the set $\{0, 1, 2, \ldots, k\}$. We will denote by $\wp$ the set of all primes and we will also be using the Euler totient function denoted by $\varphi(n)$, the function $P(n)$ which stands for the largest prime factor of n, setting $P(1) = 1$ for convenience, and finally the counting function $\pi(x)$, the number of primes not exceeding x.

We begin by stating four conjectures.

Conjecture 1 Let $k \in \mathbb{N}$. Given a permutation $\pi = \ell_0 \ell_1 \ldots \ell_k \in \Pi_{k+1}$, let

$$\mathcal{N}_k\left(x|\pi\right) := \#\{n \leq x : P(n + \ell_{i-1}) < P(n + \ell_i),\ i = 1, 2, \ldots, k\}.$$

Then,

$$\lim_{x \to \infty} \frac{1}{x} \mathcal{N}_k\left(x|\pi\right) = \frac{1}{(k+1)!}.$$

This is essentially a reformulation of Conjecture 37.6 from the book of De Koninck and Doyon [2]. Observe that this conjecture is open even in the case $k = 1$, which is

attributed to Erdős and Turan (see Erdős [7]) as they conjectured that

$$\lim_{x\to\infty}\frac{1}{x}\#\{n\le x : P(n)<P(n+1)\}=\frac{1}{2}.$$

Let us mention that some significant progress towards the proof of the Erdős-Turan conjecture has been made, in particular by Lü and Wang [9] as they proved in 2025 that, as $x\to\infty$,

$$\frac{1}{x}\#\{n\le x : P(n)<P(n+1)\}>0.2017.$$

The "normal number equivalence" of Conjecture 1 in the case $k=1$ is the following.

Conjecture 2 Let $\xi := 0.\mathrm{Concat}\big(u(P(n),P(n+1)) : n\in\mathbb{N}\big)$, that is, the concatenation of the binary digits $u(P(n),P(n+1))$ for $n=1,2,\ldots$. Then, the real number ξ is a simply normal number in base 2.

Conjecture 3 For each $K\in\mathbb{N}$, consider the corresponding set $\{\ell_1,\ell_2,\ldots,\ell_{\varphi(K)}\}$ of reduced residues modulo K and the function κ defined on primes p by

$$\kappa(p) := \begin{cases} a-1 & \text{if } \ p\equiv \ell_a \pmod K,\\ \Lambda & \text{if } \ p\mid K,\end{cases}$$

where Λ stands for the empty word, which we introduced in Section 1. Then, the real number $\xi_K := 0.\mathrm{Concat}\big(\kappa(P(n)) : n\in\mathbb{N}\big)$ is a normal number in base $\varphi(K)$.

Conjecture 4 Consider the sequence $(\lambda_p)_{p\in\wp}$ defined by

$$\lambda_p := \begin{cases} 1 & \text{if } \ P(p-1)<P(p+1),\\ 0 & \text{if } \ P(p-1)>P(p+1).\end{cases}$$

Then,

$$\lim_{x\to\infty}\frac{1}{\pi(x)}\#\{p\le x : \lambda_p=1\}=\frac{1}{2}$$

or equivalently, the real number $\xi := 0.\lambda_2\lambda_3\lambda_5\lambda_7\ldots$ is a normal number in base 2.

This last conjecture involving shifted primes is probably the most challenging of the above four conjectures.

Let us now introduce further notation which will be used in the next section. First, we set

$$E_k=\{0,1,\ldots,2^k-1\}. \tag{1}$$

Given distinct positive integers $a_0,a_1,\ldots,a_k$, further set

$$V(a_0,\ldots,a_k) := u(a_0,a_1)2^{k-1}+u(a_1,a_2)2^{k-2}+\cdots+u(a_{k-1},a_k). \tag{2}$$

Also, given $\pi\in\Pi_{k+1}$, we let

$$\alpha_k(n) := \sum_{\substack{\pi\in\Pi_{k+1}\\ V(\pi)=n}} 1$$

and

$$f_k(n) := \frac{\alpha_k(n)}{(k+1)!}. \tag{3}$$

In the case of $k = 3$, we obtain the following table:

n	$\alpha_k(n)$	$\{\pi \in \Pi_{k+1} : V(\pi) = n\}$
0	1	{3,2,1,0}
1	3	{2,1,0,3}, {3,1,0,2}, {3,2,0,1}
2	5	{1,0,3,2}, {2,0,3,1}, {2,1,3,0}, {3,0,2,1}, {3,1,2,0}
3	3	{1,0,2,3}, {2,0,1,3}, {3,0,1,2}
4	3	{0,3,2,1}, {1,3,2,0}, {2,3,1,0}
5	5	{0,2,1,3}, {0,3,1,2}, {1,2,0,3}, {1,3,0,2}, {2,3,0,1}
6	3	{0,1,3,2}, {0,2,3,1}, {1,2,3,0}
7	1	{0,1,2,3}

Moreover, for $k = 3, 4, 5$, we have the following data:

$$\begin{aligned}
\{\alpha_3(n) : 0 \le n \le 2^3 - 1\} &= \{1, 3, 5, 3, 3, 5, 3, 1\} \\
\{\alpha_4(n) : 0 \le n \le 2^4 - 1\} &= \{1, 4, 9, 6, 9, 16, 11, 4, 4, 11, 16, 9, 6, 9, 4, 1\} \\
\{\alpha_5(n) : 0 \le n \le 2^5 - 1\} &= \{1, 5, 14, 10, 19, 35, 26, 10, 14, 40, 61, 35, 26, 40, \\
&\qquad 19, 5, 5, 19, 40, 26, 35, 61, 40, 14, 10, 26, 35, 19, 10, 14, 5, 1\}
\end{aligned}$$

We now introduce a distribution function $G : [0, 1) \to [0, 1]$ which will be used in Theorem 2 (in the next section) and which we define as follows. Given $\lambda \in [0, 1)$, consider its binary expansion

$$\lambda = \frac{e_1}{2} + \frac{e_2}{2^2} + \cdots$$

with its truncated expansion

$$\lambda_T = \frac{e_1}{2} + \frac{e_2}{2^2} + \cdots + \frac{e_T}{2^T} \qquad (T = 1, 2, \ldots).$$

Also, recalling the definition of the set E_T introduced in (1), we set

$$A_T = 2^T \lambda_T = e_1 2^{T-1} + e_2 2^{T-2} + \cdots + e_T \qquad (\in E_T).$$

We now define the function G on the λ_T's as

$$G(\lambda_T) := \sum_{\substack{m \in E_T \\ m < A_T}} f_T(m),$$

where f_T is the function already defined in (3).

With the above definitions, observe that

$$\lambda_{T+1} = \lambda_T + \frac{e_{T+1}}{2^{T+1}} \quad \text{and} \quad A_{T+1} = 2A_T + e_{T+1}.$$

So, let us write $m \in E_{T+1}$ in the form

$$m = 2n + \ell, \quad \text{where } n \in E_T \text{ and } \ell \in \{0, 1\}.$$

Moreover, observe that, with the function V as in (2), we have

$$V(a_1, \ldots, a_{k+1}) = 2V(a_1, \ldots, a_k) + u(a_k, a_{k+1}).$$

Hence,

$$f_{k+1}(2n) + f_{k+1}(2n+1) = f_k(n),$$

and therefore,

$$\begin{aligned} G(\lambda_{T+1}) &= \sum_{2n+\ell < A_{T+1}} f_{T+1}(2n+\ell) \\ &= \sum_{n < A_T} (f_{T+1}(2n) + f_{T+1}(2n+1)) + S, \end{aligned}$$

where

$$S = \begin{cases} 0 & \text{if } e_{T+1} = 0, \\ f_{T+1}(2A_T) & \text{if } e_{T+1} = 1. \end{cases}$$

Hence it follows that

$$G(\lambda_{T+1}) \geq G(\lambda_T).$$

In light of the above, we can now define G as

$$G(\lambda) := \lim_{T \to \infty} G(\lambda_T).$$

One can now claim that G is a continuous function on the interval $[0, 1)$. Indeed, let us assume that $e_T = 0$ and let

$$\widetilde{\lambda_T} := \lambda_T + \frac{1}{2^T}.$$

With this set up, we have

$$G\left(\widetilde{\lambda_T}\right) = G(\lambda_T) + f_T(A_T).$$

On the other hand, it follows from the definition de f_T given in (3) that

$$f_T(A_T) \to 0 \quad \text{as} \quad T \to \infty. \tag{4}$$

Letting $\mu, \lambda \in (\lambda_T, \widetilde{\lambda_T})$, we have that

$$|G(\mu) - G(\lambda)| \leq f_T(A_T) + (G(\lambda) - G(\lambda_T)). \tag{5}$$

Now, given an arbitrarily small $\varepsilon > 0$ and taking into account (4), there exists a number T such that $e_T = 0$ and for which the right hand side of (5) does not exceed ε, thereby completing the proof that G is a continuous function in the interval $[0, 1)$.

3 The smallest prime factor of n larger than $\log n$

Given an integer $n \geq 2$, we shall denote by q_n the smallest prime divisor of n which is larger than $\log n$, and in the rare case where $P(n) < \log n$, we then set $q_n = 1$. Indeed, we may claim that the occurrence "$P(n) < \log n$" is negligible, because, using the well-known estimate

$$\Psi(x, y) := \#\{n \leq x : P(n) \leq y\} \ll x \exp\left\{-\frac{1}{2}\frac{\log x}{\log y}\right\} \qquad (2 \leq y \leq x)$$

(see for instance Theorem 9.5 in De Koninck and Luca [6]), we have that

$$\sum_{\substack{n \leq x \\ P(n) < \log n}} 1 \leq \sum_{\substack{n \leq x \\ P(n) < \log x}} 1 = \Psi(x, \log x) \ll x \exp\left\{-\frac{1}{2}\frac{\log x}{\log\log x}\right\} = o(x) \qquad (x \to \infty),$$

Our first result is the following.

Theorem 1 *Given $k \in \mathbb{N}$, consider an arbitrary partition $\pi = \ell_0\ell_1 \ldots \ell_k$ in Π_{k+1} and the corresponding counting function*

$$\mathcal{N}_k(x|\pi) := \#\left\{n \leq x : q_{n+\ell_0} < q_{n+\ell_1} < \cdots < q_{n+\ell_k}\right\}.$$

Then,

$$\lim_{x\to\infty} \frac{\mathcal{N}_k(x|\pi)}{x} = \frac{1}{(k+1)!}. \tag{6}$$

As we will see in the next section, Theorem 1 ends up being a special case of Theorem 3 stated in the next section, Hence, no proof of Theorem 1 will be presented.

Before moving on, we make the following remarks.

Remark 1 One can prove an analog of Theorem 1 for primes, in the following sense. Given $k \in \mathbb{N}$ and $p \in \wp$, let q_{p+j} stand for the smallest prime divisor of $p + j$ which is larger than $\log(p + j)$. Then, given any fixed permutation $\ell_1, \ldots, \ell_k$ of $\{1, 2, \ldots, k\}$, we have

$$\frac{1}{\pi(x)}\#\left\{p \leq x : q_{p+\ell_1} < \cdots < q_{p+\ell_k}\right\} = \frac{1}{k!}.$$

Remark 2 In the statement of Theorem 1, the condition $q_n > \log n$ can be relaxed. Indeed, this condition can be changed to $q_n > Y(n)$, where $Y(n)$ satisfies $\lim_{n\to\infty} Y(n) = \infty$ and $\lim_{n\to\infty} \dfrac{\log Y(n)}{\log n} = 0$.

The "normal number equivalence" of Theorem 1 is the following.

Theorem 2 *With q_n defined as above, for each integer $n \geq 2$, set $c_n := u(q_n, q_{n+1})$ and consider the real number $\gamma = 0.c_2c_3c_4\ldots$. Then,*

$$\lim_{x\to\infty} \frac{1}{x} \#\{n \leq x : \{2^n\gamma\} < \lambda\} = G(\lambda),$$

an estimate to which applying Theorem 1 with $k = 1$ ensures that the right hand side of (6) is equal to $1/2$ implying that γ is a simply normal number in base 2.

Proof Theorem 2 is a consequence of Theorem 1. To see this, it is enough to observe that for every

$$m = r_{k-1} \cdot 2^{k-1} + \cdots + r_1 \cdot 2 + r_0,$$

and considering the counting function

$$\mathcal{N}_k^*(x|m) := \#\{n \leq x : c_n = r_{k-1}, c_{n+1} = r_{k-2}, \ldots, c_{n+k-1} = r_0\},$$

we have, using estimate (6) of Theorem 1 and the notation introduced in Section 2,

$$\begin{aligned}\lim_{x\to\infty} \frac{\mathcal{N}_k^*(x|m)}{x} &= \lim_{x\to\infty} \sum_{\substack{\pi\in\Pi_{k+1}\\ V(\pi)=m}} \frac{\mathcal{N}_k(x|\pi)}{x} \\ &= \sum_{\substack{\pi\in\Pi_{k+1}\\ V(\pi)=m}} \frac{1}{(k+1)!} = \frac{\alpha_k(m)}{(k+1)!} = f_k(m).\end{aligned}$$

□

4 Disjoint classification of primes

In 1995, we introduced [3] the notion of a *disjoint classification of primes*, that is a collection of $T+1$ disjoint sets of primes $\mathcal{R}, \wp_1, \wp_2, \ldots, \wp_T$, whose union is $\wp$, where $\mathcal{R}$ is a finite set (perhaps empty) and where the other T sets are of positive densities $d_1, d_2, \ldots, d_T$ (with clearly $\sum_{i=1}^{T} d_i = 1$). For instance, the sets $\wp_1 = \{p : p \equiv 1 \pmod 4\}$, $\wp_2 = \{p : p \equiv 3 \pmod 4\}$ and $\mathcal{R} = \{2\}$ provide a disjoint classification of primes. We used this concept (see [4] and [5]) to create large families of normal numbers.

Here, setting $\pi([a,b] \cap \wp_i)$ as the number of primes belonging to both the interval $[a,b]$ and the set $\wp_i$, we will assume that for each $i \in \{1,\ldots,T\}$, we have

$$\pi([u, u+v] \cap \wp_i) = \frac{\pi([u,u+v])}{T} + O\left(\frac{u}{\log^5 u}\right) \quad \text{for all positive } v < u, \tag{7}$$

thus guaranteing that each set $\wp_i$ obtains its fair share of the primes belonging to the interval $[u, u+v]$ (as we did in Theorem B of our previous paper [5]). Let $\Delta = \delta_0\delta_1 \dots \delta_k$ be an arbitrary word over $\{1, \dots, T\}$ and let $\pi = \ell_0\ell_1 \dots \ell_k$ be a permutation in Π_{k+1}, and consider the counting function

$$\mathcal{N}_k(x|\Delta, \pi) := \#\{n \le x : q_{n+\ell_j} < q_{n+\ell_{j+1}} \text{ for } j = 0, \dots, k-1 \\ \text{and } q_{n+\nu} \in \wp_{\delta_\nu} \text{ for } \nu = 0, \dots, k\}.$$

With this set up, we can prove the following.

Theorem 3 *With T, k, Δ and π as above, we have that*

$$\lim_{x\to\infty} \frac{\mathcal{N}_k(x|\Delta, \pi)}{x} = \left(\frac{1}{T}\right)^{k+1} \cdot \frac{1}{(k+1)!}.$$

Proof First of all, for convenience, we set $\alpha := \left(\frac{1}{T}\right)^{k+1} \cdot \frac{1}{(k+1)!}$. Let x be a large number. Set $h := \lfloor \log x \rfloor + 1$. Also, let

$$L_{h-j} := \left[e^{h-(j+1)}, e^{h-j}\right) \qquad \text{for } j = 0, 1, \dots, h-1,$$

so that in particular, $L_h = \left[e^{h-1}, e^h\right)$. Now, given $n \in L_h$, we let q_n stand for the smallest prime factor of n which is larger than h. Moreover, for each $j = 0, 1, \dots, h-1$, let $\mathcal{N}(L_{h-j}) := \#\{n \in L_{h-j}\}$, and also, given $\Delta = \delta_0\delta_1 \dots \delta_k$, an arbitrary word over $\{1, \dots, T\}$, and $\pi = \ell_0\ell_1 \dots \ell_k$, a permutation in Π_{k+1}, we let $\mathcal{N}(L_{h-j}|\Delta, \pi)$ stand for the number of those $n \in L_{h-j}$ for which $q_{n+\ell_j} < q_{n+\ell_{j+1}}$ for $j = 0, 1, \dots, k-1$ and $q_{n+\nu} \in \wp_{\delta_\nu}$ for $\nu = 0, 1, \dots, k$. Clearly, it is enough to show that, for each $j = 0, 1, \dots, h$, we have

$$\lim_{h\to\infty} \frac{\mathcal{N}(L_{h-j}|\Delta, \pi)}{\mathcal{N}(L_h)} = \alpha. \tag{8}$$

We begin by making three key observations:

(I) In the case when $h+1$ is a prime number, then the number of those $n \in L_h$ for which

$$h+1 \mid n(n+1)\cdots(n+k)$$

does not exceed $\dfrac{(k+1)\mathcal{N}(L_h)}{h+1} + O(1)$, which is $o(\mathcal{N}(L_h))$ as $h \to \infty$.

(II) Let $\varepsilon_1 > 0$ be a small number. Then, using a simple sieve argument and Mertens' theorem in the form

$$\prod_{p\le x}\left(1 - \frac{1}{p}\right) = (1 + o(1))\frac{e^{-\gamma}}{\log x} \quad \text{as } x \to \infty,$$

where γ stands for the Euler-Mascheroni constant, we obtain that

$$\#\{n \in L_h : q_n \ge h^{1/\varepsilon_1}\} < c_1\mathcal{N}(L_h) \prod_{h<p<h^{1/\varepsilon_1}}\left(1 - \frac{1}{p}\right) < c_1\varepsilon_1\mathcal{N}(L_h).$$

(III) Let B_h stand for the set of those $n \in L_h$ for which

$$1 < \frac{\log q_{n+u}}{\log q_{n+v}} < 1 + \frac{1}{(\log\log q_{n+v})^2}$$

for integers $0 \le u < v \le k$, Now, observe that

$$S := \sum_{1 < \frac{\log q}{\log p} < 1 + \frac{1}{(\log\log p)^2}} \sum_{p<q} \frac{1}{pq} < \infty. \tag{9}$$

Indeed,

$$\begin{aligned} S &= \sum_p \frac{1}{p} \sum_{p<q<\exp\left\{\log p\left(1+\frac{1}{(\log\log p)^2}\right)\right\}} \frac{1}{q} \\ &= \sum_p \frac{1}{p}\left(\log\log p + \log\left(1 + \frac{1}{(\log\log p)^2}\right) - \log\log p\right) \\ &\ll \sum_p \frac{1}{p(\log\log p)^2} = \int_2^\infty \frac{dt}{t(\log\log t)^2}\, d\,\pi(t) \\ &\ll \int_2^\infty \frac{du}{u\log u(\log\log u)^2} = O(1), \end{aligned}$$

which proves (9). Now, it follows from (9) that

$$\frac{\#B_h}{\mathcal{N}(L_h)} \to 0 \qquad (h \to \infty).$$

In light of the above three observations, in order to prove (8), we only need to consider those primes $p_0, p_1, \dots, p_k$ satisfying

$$h + 1 < p_0 < p_1 < \cdots < p_k < h^{1/\varepsilon_1}. \tag{10}$$

So, for each fixed permutation $\pi = \ell_0\ell_1 \dots \ell_k \in \Pi_{k+1}$ and sequence of p_j's satisfying (10), we denote by $H(p_0, p_1, \dots, p_k|\pi)$ the number of those $n \in L_h$ for which $q_{n+\ell_j} = p_j$ for $j = 0, 1, \dots, k$. By a simple sieve argument, we have that, as $h \to \infty$,

$$\begin{aligned} H(p_0, p_1, \dots, p_k|\pi) &= (1 + o(1))\frac{\mathcal{N}(L_h)}{p_0p_1\cdots p_k} \\ &\quad \times \prod_{h+1<q<p_0}\left(1 - \frac{k+1}{q}\right)\prod_{p_0<q<p_1}\left(1 - \frac{k}{q}\right)\cdots\prod_{p_{k-1}<q<p_k}\left(1 - \frac{1}{q}\right) \\ &= (1 + o(1))\mathcal{N}(L_h)\prod_{j=0}^k \lambda_{p_j}, \end{aligned} \tag{11}$$

where we used the abbreviation $\lambda_p := \dfrac{\log h}{p\log p}$ and where we made repeated use of Mertens' theorem (mentioned above in (II)).

Let us fix $p_0, p_1, \dots, p_k$. Since the above asymptotic formula (11) does not depend on the particular choice of the partition $\pi(= \ell_0 \dots \ell_k)$ and since in light of (7), we have that

$$\sum_{\substack{u<p<u+v \\ p\in\wp_{\delta_j}}} \frac{1}{p\log p} = \frac{1+o(1)}{T} \sum_{u<p<u+v} \frac{1}{p\log p} \qquad \text{for each } j = 1, \dots T,$$

we may write that for any of the given $j = 0, 1, \dots, k-1$, we have that, as $h \to \infty$,

$$\frac{\mathcal{N}(L_{h-j}|\Delta, \pi)}{\mathcal{N}(L_h)} = \frac{1+o(1)}{T^{k+1}} \sum_{h+1<p_0<\dots<p_k<h^{1/\varepsilon_1}} \lambda_{p_0} \cdots \lambda_{p_k}. \tag{12}$$

Now, for convenience, let us set

$$H := \sum_{h+1<p<h^{1/\varepsilon_1}} \lambda_p \quad \text{and} \quad S := \sum_{h+1<p_0<\dots<p_k<h^{1/\varepsilon_1}} \lambda_{p_0} \cdots \lambda_{p_k}.$$

First observe that, as $h \to \infty$,

$$H = \int_h^{h^{1/\varepsilon_1}} \frac{\log h}{u \log u}\, d\,\pi(u) = (1+o(1)) \log h \int_{\log h}^{\frac{1}{\varepsilon_1}\log h} \frac{dv}{v^2} = (1+o(1))(1-\varepsilon_1). \tag{13}$$

Observe also that

$$0 \le \frac{H^{k+1}}{(k+1)!} - S = \sum\nolimits^{*} \lambda_{p_0} \cdots \lambda_{p_k}, \tag{14}$$

where the $*$ on this last sum indicates that the sum runs over those $(k+1)$-tuples $(p_0, p_1, \dots, p_k)$ which contain at least one couple $\{p_i, p_j\}$ with $i \neq j$ and $p_i = p_j$. This last sum is "small". Indeed,

$$\sum\nolimits^{*} \lambda_{p_0} \cdots \lambda_{p_k} \le \frac{k(k-1)}{2} \sum_{h<p<h^{1/\varepsilon_1}} \lambda_p^2 \cdot H^{k-1} \tag{15}$$

and, trivially,

$$\sum_{h<p<h^{1/\varepsilon_1}} \lambda_p^2 \le \frac{\log h}{h \log h} \sum_{h<p<h^{1/\varepsilon_1}} \lambda_p = \frac{1}{h} \cdot H \le \frac{2}{h}, \tag{16}$$

where we used (13). Hence, combining (15) and (16), and recalling (13), we obtain that

$$\sum\nolimits^{*} \lambda_{p_0} \cdots \lambda_{p_k} \le \frac{k(k-1)}{2} \cdot \frac{2}{h} \cdot (1-\varepsilon_1)^{k-1} \to 0 \quad \text{as } h \to \infty. \tag{17}$$

Thus, using (17) in (14), we may conclude from (14), again using (13), that

$$S = (1+o(1)) \frac{H^{k+1}}{(k+1)!} = (1+o(1)) \frac{(1-\varepsilon_1)^{k+1}}{(k+1)!}. \tag{18}$$

Substituting (18) in (12), we finally obtain that

$$\frac{\mathcal{N}(L_{h-j}|\Delta, \pi)}{\mathcal{N}(L_h)} = \frac{(1+o(1))}{T^{k+1}(k+1)!} \qquad (h \to \infty),$$

thus completing the proof of (8) and therefore of Theorem 3. □

5 Applications

Here are three examples of disjoint classification of primes to which one could apply Theorem 3.

Example 1 Let $D \geq 3$ be an integer and let $\ell_1, \ldots, \ell_{\varphi(D)}$ be the set of reduced residues modulo D. Then, by choosing $\wp_j := \{p \in \wp : p \equiv \ell_j \pmod{D}\}$ for $j = 1, \ldots, \varphi(D)$ and $\mathcal{R} = \{p \in \wp : p \mid D\}$, one can then see that the sets of primes $\mathcal{R}, \wp_1, \wp_2, \ldots, \wp_{\varphi(D)}$ represent a disjoint classification of primes. Then, for any given prime number p, we set

$$s(p) := \begin{cases} j & \text{if } \ p \in \wp_j, \\ \Lambda & \text{if } \ p \in \mathcal{R}. \end{cases}$$

Given $\Delta = \delta_0 \ldots \delta_k$, an arbitrary word over the set of digits $\{1, \ldots, \varphi(D)\}$ and a permutation $\pi = \ell_0 \ldots \ell_k \in \Pi_{k+1}$, we then have

$$\begin{aligned} \lim_{x\to\infty} \frac{\mathcal{N}_k(x|\Delta, \pi)}{x} &= \lim_{x\to\infty} \frac{1}{x} \#\big\{n \leq x : q_{n+\ell_0} < q_{n+\ell_1} < \cdots < q_{n+\ell_k} \\ &\qquad \text{and } s(q_{n+\ell_j}) = \delta_j \text{ for } j = 0, 1, \ldots, k\big\} \\ &= \frac{1}{(k+1)!\varphi(D)^{k+1}}. \end{aligned}$$

Example 2 Let $\mathcal{F}$ be the set of those irrational numbers α for which given any pair of co-prime integers D, ℓ for which given any set $I = [u, v) \subseteq [0, 1)$, we have (in what follows, $\{\alpha p\}$ stands for the fractional part of αp)

$$\#\{p \leq x : p \equiv \ell \pmod{D},\ \{\alpha p\} \in I\} = (v - u)\pi(x; D, \ell) + O\left(\frac{\pi(x)}{\log^2 x}\right).$$

By using a method due to Vinogradov (see Chapter 11 in his book [10]), one can show that almost all irrational numbers α belong to $\mathcal{F}$. Now, let $K \geq 3$ be an integer and consider the half-open intervals $I_j = [j/K, (j+1)/K)$, $j = 0, 1, \ldots, K-1$, and consider the sets $\wp_j = \{p \in \wp : \{\alpha p\} \in I_j\}$, where $j = 0, 1, \ldots, K-1$. Then, Theorem 3 holds for this choice $\wp = \wp_0 \cup \ldots \cup \wp_{K-1}$ of a disjoint classification of primes.

Example 3 Let $\alpha \in \mathcal{F}$, D, K and I_j be as in Examples 1 and 2, and consider the sets

$$\wp_{\nu,j} = \{p \in \wp : p \equiv \ell_\nu \pmod{D} \text{ and } \{\alpha p\} \in I_j\} \qquad (\nu = 1, \ldots, \varphi(D),\ j = 0, 1, \ldots, K-1).$$

Again, Theorem 3 applies to these sets, since, using the Vinogradov method mentioned in Example 2, one can show that

$$\lim_{x\to\infty} \frac{1}{\pi(x)} \#\{p \leq x : p \in \wp_{\nu,j}\} = \frac{1}{K\varphi(D)}.$$

6 Further remarks

Let $a_1 < a_2 < \cdots$ be an infinite sequence of positive integers and let q_n be the smallest prime divisor of a_n which is larger than $\log n$. We formulated and proved Theorem 1 in the particular case $a_n = n$. One could show that an analog theorem holds if $a_n = Q(n)$, where $Q \in \mathbb{Z}[x]$ is an irreducible polynomial satisfying $Q(n) > 0$ for all $n \in \mathbb{N}$. The familiar function $Q(n) = n^2 + 1$ is such a polynomial.

Acknowledgments. The authors are grateful to the referee for the careful reading of the paper and for helpful suggestions. The work of the first author was funded partially by a grant from the *Natural Sciences and Engineering Research Council of Canada*.

Data availability statement

Section 2 of the paper includes the data supporting the results proved in the paper. The data in this section was generated using the software system Mathematica.

Declarations

The authors have no conflict of interest.

References

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

[2] J.-M. De Koninck and N. Doyon, *The Life of Primes in 37 Episodes*, American Mathematical Society, Providence, Rhode Island, 2021.

[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, *Construction of normal numbers by classified prime divisors of integers II*, Functiones et Approximatio **49** (2013), No. 1, 7–27.

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

[7] P. Erdős, *Some unconventional problems in number theory* In *Journées arithmétiques de Luminy (Colloque Intern. CNRS, Centre univ. Luminy, Luminy, 1978)*, volume 61 of *Astérisque*, pages 73–82. Soc. Math. France, Paris, 1979.

[8] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, John Wiley & Sons, New York, 1974.

[9] Xiaodong Lü and Zhiwei Wang, *On the largest prime factors of consecutive integers*, Monatsh. Math. **206** (2025), no. 2, 403—418.

[10] I.M. Vinogradov, *The method of trigonometrical sums in the theory of numbers* (Russian) Trav. Inst. Math. Stekloff **23** (1947), 109 pp.