TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 372, Number 6, 15 September 2019, Pages 4425–4446
https://doi.org/10.1090/tran/7706
Article electronically published on November 26, 2018

NORMAL NUMBERS WITH DIGIT DEPENDENCIES

CHRISTOPH AISTLEITNER, VERÓNICA BECHER, AND OLIVIER CARTON

ABSTRACT. We give metric theorems for the property of Borel normality for real numbers under the assumption of digit dependencies in their expansion in a given integer base. We quantify precisely how much digit dependence can be allowed such that almost all real numbers are normal. Our theorem states that almost all real numbers are normal when at least slightly more than $\log\log n$ consecutive digits with indices starting at position n are independent. As the main application, we consider the Toeplitz set T_P, which is the set of all sequences $a_1a_2\dots$ of symbols from $\{0,\dots,b-1\}$ such that a_n is equal to a_{pn} for every p in P and $n = 1, 2, \dots$. Here b is an integer base and P is a finite set of prime numbers. We show that almost every real number whose base b expansion is in T_P is normal to base b. In the case when P is the singleton set $\{2\}$ we prove that more is true: almost every real number whose base b expansion is in T_P is normal to all integer bases. We also consider the Toeplitz transform which maps the set of all sequences to the set T_P, and we characterize the normal sequences whose Toeplitz transform is normal as well.

1. Introduction and statement of results

For a real number x in the unit interval, its expansion in an integer base $b \geq 2$ is a sequence of integers $a_1, a_2 \dots$, where $0 \leq a_j < b$ for every j, such that

$$x = \sum_{j=1}^{\infty} a_j b^{-j}. \tag{1}$$

We require that $a_j < b-1$ infinitely often to ensure that every number has a unique representation. When the base is understood we write $x = 0.a_1a_2\dots$.

The concept of normality of numbers was introduced by Borel [7] in 1909, and there are several equivalent formulations (see [4, 8]). The most convenient for our purposes was given by Pillai [15]: A real number x is simply normal to a given base b if every possible digit in $\{0, 1, \dots, b-1\}$ occurs in the b-ary expansion of x with the same asymptotic frequency (that is, with frequency $1/b$). A real number x is normal to base b if it is simply normal to all the bases $b, b^2, b^3, \dots$. Absolute normality is defined as normality to every integer base $b \geq 2$. Borel proved that almost all numbers (with respect to Lebesgue measure) are absolutely normal.

Received by the editors April 9, 2018, and, in revised form, June 26, 2018 and September 4, 2018.

2010 *Mathematics Subject Classification.* Primary 11K16; Secondary 68R15.

The first author is supported by the Austrian Science Fund (FWF), projects Y-901 and F-5512-N26.

The second and third authors are members of the Laboratoire International Associé INFINIS, CONICET/Universidad de Buenos Aires–CNRS/Université Paris Diderot, and they are supported by the ECOS project PA17C04.

The third author is also partially funded by the DeLTA project (ANR-16-CE40-0007).

In this paper we consider normality under the assumption of additional dependencies between the digits of a number. Let a base b be fixed, and consider the set of real numbers $x = 0.a_1a_2a_3a_4\ldots$ in the unit interval where the digits $a_1, a_2, a_3, a_4\ldots$ can be divided into *free* or *independent* digits on the one hand, and *dependent* digits on the other hand. The free digits can be chosen at will, while for the dependent digits there is a restriction which prescribes their values deterministically from the values of a certain set of digits with smaller indices. For example, consider the restriction that the equality

$$a_{2n} = a_n$$

must hold for all $n \geq 1$; then $a_1, a_3, a_5, \ldots$ are independent and can be freely chosen, while $a_2, a_4, a_6, \ldots$ are dependent since they are completely determined by earlier digits.

A special form of such digit dependencies has been formalized by Jacobs and Keane in [12] by considering *Toeplitz sequences* and the *Toeplitz transform*, which we present now. Fix an integer $b \geq 2$. Let A denote the alphabet $A = \{0, \ldots, b-1\}$, and write A^ω for the set of all infinite sequences of symbols from A. For a positive integer r and a set $P = \{p_1, \ldots, p_r\}$ of r prime numbers, let T_P be the set of all Toeplitz sequences; that is, the set of all sequences $t_1t_2t_3\cdots$ in A^ω such that for every $n \geq 1$ and for every $i = 1, \ldots, r$,

$$t_n = t_{np_i}. \tag{2}$$

The Toeplitz transform maps sequences in A^ω to the Toeplitz set T_P. Let $j_1, j_2, j_3, \ldots$ be the enumeration in increasing order of all positive integers which are not divisible by any of the primes $p_1, \ldots, p_r$. Then, every positive integer n has a unique decomposition $n = j_k p_1^{e_1} \cdots p_r^{e_r}$, where each integer e_i is the p_i-adic valuation of n, for $i = 1, \ldots, r$. The Toeplitz transform $\tau_P : A^\omega \to A^\omega$ is defined as

$$\tau_P(a_1a_2a_3\ldots) = t_1t_2t_3\ldots$$

where

$$t_n = a_k \quad \text{when } n \text{ has the decomposition} \quad n = j_k p_1^{e_1} \cdots p_r^{e_r}. \tag{3}$$

Thus, the image of A^ω under the transform τ_P is the set T_P. Since elements of A^ω can be identified with real numbers in $[0, 1]$ in a natural way via (1), the transform τ_P induces a transform $[0, 1] \mapsto T_P$, which we denote by τ_P as well. We endow T_P with a probability measure μ, which is the forward push by τ_P of the uniform probability measure λ on A^ω (which, in turn, is the infinite product measure generated by the uniform measure on $\{0, \ldots, b-1\}$). Note that μ can also be seen as a measure on the set of all sequences, that is on A^ω, since T_P is embedded in A^ω. Again, as already noted above for τ_P, by identifying infinite sequences with real numbers, the measure μ on A^ω also induces a measure on $[0, 1]$, which we denote by μ as well. For any measurable set $X \subseteq T_P$, $\mu(X) = \lambda(\tau_P^{-1}(X))$. Informally speaking, μ is the natural uniform measure on the set of all sequences (resp., real numbers) which respect the digit dependencies imposed by (2).

The Toeplitz transform τ_P also induces a function $\delta : \mathbb{N} \mapsto \mathbb{N}$ on the index set, by defining $\delta(n) = k$ for k and n as in (3). Hence, $\delta(n) = k$ means that the nth symbol t_n of the image of $a_1a_2a_3\ldots$ under the Toeplitz transform is a_k. This can also be written as

$$t_1t_2t_3\cdots = \tau_P(a_1a_2a_3\cdots) = a_{\delta(1)}a_{\delta(2)}a_{\delta(3)}\cdots.$$

The nth symbol $t_n(x)$ of $\tau_P(x)$ is a measurable function $[0,1] \mapsto \{0,\dots,b-1\}$. Thus, $t_n(x)$ is random variable on the space $([0,1], \mathcal{B}(0,1), \lambda)$. Since $t_n(x) = a_{\delta(n)}$ for all n, it is easy to see that two random variables t_m and t_n are independent (with respect to both measures λ and μ) if and only if $\delta(m) \neq \delta(n)$; that is, if they do not originate in the same digit of x by means of the Toeplitz transform.

We say that an infinite sequence of symbols from $\{0,\dots,b-1\}$ is normal if it is the expansion of a real number which is normal to base b. Our first result is the following theorem. It shows that "typical" elements of T_P are normal, just as by Borel's theorem typical real numbers are normal. Thus, imposing additional digit dependencies does not destroy the fact that almost all numbers are normal.

Theorem 1. *Let $b \geq 2$ be an integer, and let P be a finite set of primes. Let μ be the "uniform" probability measure on the set T_P, defined above. Then, μ-almost all elements of T_P are the expansion in base b of a normal number.*

The proof of Theorem 1 generalizes the one given by Alexander Shen (personal communication, June 2016) for the special case $P = \{2\}$. It relies on the fact that the sequence of all integers which are generated by a finite set P of primes, sorted in increasing order, grows very quickly. More precisely, there is a very strong gap condition which gives a lower bound for the minimal size of the gap $k - k'$ whenever $k > k'$ are two numbers generated by P, as a function of k'. In the case when $P = \{p\}$ is a singleton, this is a trivial observation, since then the integers generated by p form a geometric progression, but when P has cardinality at least 2 this is a subtle property, which can be established using Alan Baker's celebrated theory of linear forms of logarithms (see for example [3]). Note that the set of numbers generated by finitely many primes forms a semigroup; it can be extended to a group of what is called S-units, which are well-studied objects in algebraic number theory due to their connection with the theory of Diophantine equations. It is a well-known fact in probabilistic number theory that parametric lacunary sequences, which are sequences of the form $(m_n x)_{n\geq 1}$ where there exists $c > 1$ such that $m_{n+1}/m_n > 1$ for all n, often lead to weakly dependent random systems that behave asymptotically like truly independent random systems; see [1] for a survey.

The classical lacunary systems originate from geometric progressions, but it turned out that often it is possible to adapt the machinery to sequences of integers generated by finitely many primes. Classical results in that direction are denseness properties which Furstenberg [11] deduced from the disjointness of corresponding measure-preserving transforms, and Philipp's [14] law of the iterated logarithm. However, the setting in the present paper is very different from these earlier results. By Weyl's criterion, normality of a number x to base b is equivalent to the fact that the sequence of fractional parts of $x, bx, b^2x, \dots$ is uniformly distributed modulo 1. Now our purpose is not to replace $(b^n)_{n\geq 1}$ by a sequence of integers generated by finitely many primes (as it is in [11, 14]), but instead we impose a restriction on the digits of x for special sets of indices, and try to prove that for such x we have uniform distribution mod 1 of the fractional parts of $x, bx, b^2x, \dots$. These are very different problems and completely different methods are required.

Theorem 1 shows that μ-almost all numbers are normal in a given base b. However, it turns out that much more is true, at least in the case when $P = \{2\}$. For real numbers satisfying $a_{2n} = a_n$ for every $n \geq 1$ in their expansion in base b, we prove that almost surely they are actually absolutely normal. This is the same as saying that μ-almost all real numbers in $[0, 1]$ are absolutely normal. This result is in the following theorem.

Theorem 2. *Let $b \geq 2$ be an integer, let $P = \{2\}$, and let μ be the "uniform" probability measure on T_P. Then, μ-almost all elements of T_P are the expansion in base b of an absolutely normal number.*

To prove Theorem 2 we adapt the work of Cassels [9] and Schmidt [16]. Our argument is also based on the idea of giving upper bounds for certain Riesz products, although the setting is quite different. For example, Cassels worked on a Cantor-type set of real numbers whose ternary expansion avoids the digit 2 (and which therefore cannot be normal to base 3), and he had to establish certain regularity properties of the uniform measure supported on this fractal set. In contrast, we clearly have to deal with the measure μ which is the uniform measure on the set of real numbers which respect the digit restriction (2). This property on digits is more delicate than that of avoiding certain digits altogether, but still it turns out that it is possible to use similar techniques. Our proof of Theorem 2 can in principle be generalized to Toeplitz sequences for arbitrary finite sets of primes P instead of $P = \{2\}$. However, to keep the proof reasonably simple, we do not deal with this general setting in the present paper. The proof works by partitioning the set of all possible positions of symbols into equivalence classes such that at all these positions of the Toeplitz sequence the same symbol occurs; using notation introduced above, each equivalence class collects all those indices for which the function σ gives the same value. In case $P = \{2\}$, all Toeplitz sequences satisfy that all the positions of the form 2^n, for $n = 0, 1, \ldots$ have the same symbol, all the positions of the form $3\ 2^n$, for $n = 0, 1, \ldots$ have the same symbol, all the positions of the form $5\ 2^n$, for $n = 0, 1, \ldots$ have the same symbol, and so on. This determines, for each odd number, one equivalence class of positions. In case P is a finite set of prime numbers, $P = \{p_1, p_2, \ldots, p_r\}$, the definition of the equivalence classes is subtler. For each positive integer s that is not a multiple of any p in P, all the positions of the form $sp_1^{n_1} p_2^{n_2} \ldots p_r^{n_r}$ where each $n_1, n_2 \ldots$ take the values $0, 1, 2, \ldots$ form an equivalence class. The proof of Theorem 2 generalized to a set P with finitely many primes requires a bound exponential sums over all the equivalence classes.

We are also interested in a general framework for results of the form of Theorem 1, where, however, the digit dependencies can be much more general than those imposed for Toeplitz sequences. How much digit dependence can be allowed in some given base such that, still, almost all real numbers are normal to that base? Our Theorem 3 below quantifies how many consecutive digits have to be independent, in order to keep the usual property that almost all numbers are normal. Quite surprisingly, it turns out that only a very low degree of independence is necessary. The theorem says, roughly speaking, that as long as we can assure that slightly more than $\log \log n$ consecutive digits with indices starting at n are independent for all sufficiently large n, then almost all real numbers are normal. On the other hand, assuming independence of blocks of $\log \log n$ consecutive digits is not sufficient.

For the statement of the following theorem, let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Let $X_1, X_2, \ldots$ be a sequence of random variables (that is, measurable functions) from $(\Omega, \mathcal{A}, \mathbb{P})$ into $\{0, \ldots, b-1\}$.

Theorem 3. *Assume that for every $n \geq 1$ the random variable X_n is uniformly distributed on $\{0, \ldots, b-1\}$. Assume furthermore that there exists a function $g : \mathbb{N} \mapsto \mathbb{R}$ which is monotonically increasing to ∞ such that for all sufficiently large n the random variables*

$$X_n, \ X_{n+1}, \ \ldots, \ X_{n+\lceil g(n) \log\log n \rceil} \tag{4}$$

are mutually independent. Let x be the real number whose expansion in base b is given by $x = 0.X_1X_2X_3\ldots$. Then $\mathbb{P}$-almost surely the number x is normal to base b.

On the other hand, for every base b and every constant $K > 0$ there is an example where for every $n \geq 1$ the random variable X_n is uniformly distributed on $\{0, \ldots, b-1\}$, and where for all sufficiently large n the random variables

$$X_n, \ X_{n+1}, \ \ldots, \ X_{n+\lceil K \log\log n \rceil} \tag{5}$$

are mutually independent, but $\mathbb{P}$-almost surely the number $x = 0.X_1X_2X_3\ldots$ even fails to be simply normal.

Note that the theorem gives an optimal condition for the degree of independence which is necessary to have normality for "typical" numbers. From the proof of Theorem 3 it is visible that for the correctness of the conclusion of the first part of the theorem, it is not necessary that (4) holds for *all* (sufficiently large) n, but that it is possible to allow a set of exceptional indices which has to be assumed to have small density in some appropriate quantitative sense. However, to keep the presentation short we do not state the theorem in such generality.

Theorem 1 says that μ-almost all sequences are mapped by the Toeplitz transform τ_P to normal ones. A natural question is whether all the normal sequences are mapped by τ_P to normal ones. The following example shows that this is not the case. Let $P = \{2\}$, and let $x = a_1a_2a_3\cdots$ be a normal sequence such that $a_{2n} = a_n$ for each $n \geq 1$. Then, $\tau_P(x) = d_1d_2d_3\cdots$ satisfies for each $n \geq 1$, $d_{2n} = d_n$, and $d_{2n-1} = a_n$. Combining these relations we obtain $d_{4n-2} = d_{2n-1} = a_n = a_{2n} = d_{4n-1}$, which proves that $\tau_P(x)$ is not normal. This example actually shows that applying twice the transform τ_P to a normal sequence never yields a normal sequence. The last theorem of this paper, Theorem 4, gives a characterization of those normal sequences whose Toeplitz transform is also normal.

We refer to finite sequences of symbols in $\{0, \ldots, b-1\}$ as words. If u is a word we write $|u|$ to denote its length.

For simplicity, the next theorem is only stated and proved when the cardinality of the set P of prime numbers is 2, but it can be generalized to any finite set $P = \{p_1, \ldots, p_r\}$ of prime numbers.

Theorem 4. *Let $P = \{p_1, p_2\}$ where p_1 and p_2 are two prime numbers. For a sequence x in $\{0, \ldots, b-1\}^\omega$, the following conditions are equivalent.*

(I) *The Toeplitz transform $\tau_P(x)$ of x is normal to base b.*

(II) *For every integer $k \geq 0$ and every family $(u_{i_1,i_2})_{0 \leq i_1, i_2 \leq k}$ of words of arbitrary length, the limit*

$$\lim_{N\to\infty} \frac{1}{N} \# \left\{ n : \begin{array}{l} 1 \leq n \leq N, \ \text{and for all } 0 \leq i_1, i_2 \leq k, \\ u_{i_1,i_2} \ \text{occurs in } x \text{ at position } (p_1-1)(p_2-1)p_1^{i_1}p_2^{i_2}n \end{array} \right\}$$

is equal to $b^{-\sum_{0 \leq i_1, i_2 \leq k} |u_{i_1,i_2}|}$.

Note that condition (II) with $k = 0$ just states that the sequence x itself must be normal. This is indeed required because the symbols in $\tau_P(x)$ at positions not divisible by p_1 and p_2 are exactly those in x, in the same order. Condition (II) can be viewed as a sort of asymptotic probabilistic independence between words occurring at positions of the form $(p_1 - 1)(p_2 - 1)p_1^{i_1} p_2^{i_2} n$. It states that the asymptotic frequency of a family of words is exactly the product of their frequencies.

We hope that Theorem 4 will help to find a construction of a normal sequence in T_P for some general finite set P of primes. A construction of one explicit normal sequence in T_P for $b = 2$ and the special case $P = \{2\}$ appears in [4] and [5]. This construction can be generalized to any integer base b and any singleton P.

The remainder of the paper is devoted to the proofs of Theorems 1 to 4.

2. Proof of Theorem 1

We prove Theorem 1 by showing that it is a consequence of Theorem 3, together with number-theoretic results of Tijdeman [18]. Let r be a positive integer, and let $P = \{p_1, \dots, p_r\}$ be a set of r primes. We define the sets K and L as

$$K = \{p_1^{e_1} \cdots p_r^{e_r} : e_i \geq 0\} \quad \text{and} \quad L = \{\ell : p_i \nmid \ell, i = 1, \dots, r\}.$$

Thus, every positive integer n can be written in a unique way as $n = k\ell$ for some $k \in K$ and $\ell \in L$. We define a equivalence relation $\sim$ on the set of positive integers by writing $n \sim n'$ whenever there are $k, k' \in K$ and $\ell \in L$ such that $n = k\ell$ and $n' = k'\ell$.

Lemma 5. *There exists an integer n_0 such that if $n' \sim n$ and $n' > n > n_0$, then $n' - n > 2\sqrt{n}$.*

Proof of Lemma 5. In [18] Tijdeman proved that there exists a positive constant C such that for all $k, k' \in K$ satisfying $k < k'$, we have

$$k' - k > k/(\log k)^C.$$

There also exists an integer k_0 such that

$$k/(\log k)^C > 2\sqrt{k}$$

for all $k \geq k_0$. Let $n_0 = 4k_0^2$. Suppose that $n' = k'\ell$ and $n = k\ell$ for $k, k' \in K$ and $\ell \in L$. If $k \geq k_0$, then

$$n' - n = (k' - k)\ell > 2\sqrt{k}\ell > 2\sqrt{n}.$$

If $k < k_0$ and $n = k\ell > 4k_0^2$, then $\ell > 4k_0$. It follows that

$$n' - n = (k' - k)\ell > \ell > 2\sqrt{\ell}\sqrt{k_0} > 2\sqrt{n}. \qquad \square$$

Proof of Theorem 1. Let $x = 0.a_1a_2a_3\dots$ be a real number, and let $\tau_P(x) = 0.t_1t_2t_3\dots$ be its Toeplitz transform. As noted in the introduction, $t_n(x)$ is a measurable function from $([0,1], \mathcal{B}(0,1), \lambda)$ to $\{0, \dots, b-1\}$ for all n. Clearly t_n has uniform distribution on $\{0, \dots, b-1\}$, since $t_n(x) = a_{\delta(n)}$ and the digit $a_{\delta(n)}$ takes all possible values with equal probability with respect to Lebesgue measure. Lemma 5 can be rephrased as saying that for all sufficiently large n all the numbers

$$\delta(n), \delta(n+1), \dots, \delta(n + \lfloor 2\sqrt{n} \rfloor)$$

are different, since $n \sim n'$ holds if and only if $\delta(n) = \delta(n')$. Thus, for all sufficiently large n, the random variables

$$a_{\delta(n)}, a_{\delta(n+1)}, \ldots, a_{\delta(n+\lfloor 2\sqrt{n} \rfloor)}$$

are mutually independent with respect to λ, since different digits of a real number are mutually independent with respect to Lebesgue measure (the digits are Rademacher random variables; their independence with respect to λ was first observed by Steinhaus, see [1]). However, this is the same as saying that

$$t_n, t_{n+1}, \ldots, t_{n+\lfloor 2\sqrt{n} \rfloor}$$

are mutually independent for sufficiently large n, with respect to λ. Thus, we see that the assumptions of Theorem 3 are satisfied, and thus for λ-almost all input values x, the number $0.t_1t_2t_3\cdots = \tau_P(x)$ is normal to base b. From the way μ is obtained from λ, this is equivalent to saying that μ-almost all sequence in T_P are normal to base b, which proves Theorem 1. □

3. Proof of Theorem 2

Fix the integer $b \geq 2$, and let $P = \{2\}$. We need to show that for all integers $r \geq 2$, μ-almost all elements of T_P are the expansion of a number that is normal to base r. As usual, we say that two positive integers are *multiplicatively dependent* if one is a rational power of the other. In case b and r are multiplicatively dependent, Theorem 2 follows immediately from Theorem 1 because normality to base b is equivalent to normality to any multiplicatively dependent base r.

In case r is multiplicatively independent to b the main structure of our proof follows the work of Cassels in [9], but adapted to the uniform measure on the real numbers whose expansion is in T_P. We need two lemmas. The first one, Lemma 6, is similar to Schmidt's [16, Hilfssatz 5], except that in our case the product is taken only over the odd integers. The second one, Lemma 7, bounds the $L^2(\mu)$ norm of the appropriate exponential sums.

We start by introducing some notation. For $v = v_1v_2\ldots$ in T_P, let x_v be the real number in the unit interval defined by

$$x_v = \sum_{j\geq 1} b^{-j} v_j. \tag{6}$$

We write $T_P(\ell)$ for the set of sequences of length ℓ that are initial segments of elements in the Toeplitz set T_P for $P = \{2\}$; that is,

$$T_P(\ell) = \Big\{a_1a_2\ldots a_\ell \in \{0,\ldots,b-1\}^\ell : a_n = a_{2n} \text{ for each } 1 \leq n \leq \ell/2\Big\}.$$

Similar to (6), for $v = v_1v_2\ldots v_\ell$ in $T_P(\ell)$, we let x_v be

$$x_v = \sum_{j=1}^{\ell} b^{-j} v_j.$$

Lemma 6. *Let r and b be multiplicatively independent positive integers. There is a constant $c > 0$, depending only on r and b, such that for all positive integers J and L with $L \geq b^J$, and for every positive integer N,*

$$\sum_{n=0}^{N-1} \prod_{\substack{q=J+1\\ q \text{ odd}}}^{\infty} \left(\frac{1}{b} + \frac{b-1}{b} \left| \cos\left(\pi r^n L b^{-q}\right) \right| \right) \leq 2N^{1-c}.$$

Proof. Schmidt's [16, Hilfssatz 5] states that for all multiplicatively independent integers $r \geq 2$ and $s \geq 2$ there is a constant $c_1 > 0$, depending only on r and s, such that for all positive integers K and L with $L \geq s^K$, and for every positive integer N,[†]

$$\sum_{n=0}^{N-1} \prod_{k=K+1}^{\infty} |\cos(\pi r^n L s^{-k})| \leq 2N^{1-c_1}.$$

When examining the proof of this *Hilfssatz*, one sees that the only properties of the function $|\cos(\pi x)|$ that are used in the proof are the periodicity, the fact that $|\cos(\pi x)| \leq 1$, and finally the fact that $|\cos(\pi/s^2)| < 1$. However, all these properties also hold for the function $\frac{1}{p} + \frac{p-1}{p} |\cos(\pi x)|$ for any integer $p > 1$, so Schmidt's proof can also be used without any further changes to show that

$$\sum_{n=0}^{N-1} \prod_{k=K+1}^{\infty} \left(\frac{1}{p} + \frac{p-1}{p} \left| \cos\left(\pi r^n L s^{-k}\right) \right| \right) \leq 2N^{1-c_2} \tag{7}$$

for a constant $c_2 > 0$ depending only on p, r, and s.

Our lemma assumes that r and b are multiplicatively independent, so r and b^2 are multiplicatively independent as well. Replacing p by b and s by b^2 in (7), we obtain:

$$\sum_{n=0}^{N-1} \prod_{k=K+1}^{\infty} \left(\frac{1}{b} + \frac{b-1}{b} \left| \cos\left(\pi r^n L b^{-2k}\right) \right| \right) \leq 2N^{1-c_3}$$

for all K, L satisfying $L \geq (b^2)^K$, where the constant $c_3 > 0$ depends on r and b^2 (which is equivalent to saying that c_3 depends on r and b). In particular this holds for all L which are multiples of b, satisfying $L \geq b^{2K}$. So let us assume that L is a multiple of b, and that accordingly $L = bm$. Then we have

$$\sum_{n=0}^{N-1} \prod_{k=K+1}^{\infty} \left(\frac{1}{b} + \frac{b-1}{b} \left| \cos\left(\pi r^n m b^{-(2k-1)}\right) \right| \right) \leq 2N^{1-c_3},$$

provided that $bm \geq b^{2K}$. Now, writing $2k - 1 = q$, this is

$$\sum_{n=0}^{N-1} \prod_{\substack{q=2K+1,\\ q \text{ odd}}}^{\infty} \left(\frac{1}{b} + \frac{b-1}{b} \left| \cos\left(\pi r^n m b^{-q}\right) \right| \right) \leq 2N^{1-c_3}.$$

[†]In Schmidt's paper the sum is written as $\sum_{r=0}^{N-1}$, but actually the sum is $\sum_{n=0}^{N-1}$.

Finally writing $J+1 = 2K+1$ this is

$$\sum_{n=0}^{N-1} \prod_{\substack{q=J+1,\\ q \text{ odd}}}^{\infty} \left(\frac{1}{b} + \frac{b-1}{b} \left| \cos\left(\pi r^n m b^{-q}\right) \right| \right) \le 2N^{1-c_3},$$

which holds for $bm \ge b^{2K}$; that is (since $K = J/2$) for $bm \ge (b^2)^{J/2}$, or equivalently for $m \ge b^{J-1}$. We can relax the final restriction to $m \ge b^J$. This proves the lemma. □

As usual, we write $e(x)$ to denote $\mathrm{e}^{2\pi i x}$.

Lemma 7. *Let $b \ge 2$ be a integer, and assume that $P = \{2\}$. Let T_P be the corresponding Toeplitz transform in base b, and let μ be the associated measure, as introduced in section 1. Let $r \ge 2$ be an integer multiplicatively independent to b. Then for all integers $h \ge 1$ there exist constants $c > 0$ and $k_0 > 0$, depending only on $b, r,$ and h, such that for all positive integers k, m satisfying $m \ge k+1+2\log_r b \ge k_0$,*

$$\int_0^1 \left| \sum_{j=m+1}^{m+k} e(r^j h x) \right|^2 d\mu(x) \le k^{2-c}.$$

Proof. We write ℓ for the smallest even integer which is larger than

$$((m+k+1)\log_b r) + \log_b h.$$

Let $x \in [0,1]$ be given, and let $\bar{x}$ be the same number as x, but with the digits in positions $\ell+1, \ell+2, \dots$ after the decimal point all being replaced by zeros. Then $|x - \bar{x}| \le b^{-\ell} \le h^{-1} r^{-m-k-1}$. Thus, using derivatives and the inequality $|y^2 - z^2| \le |y+z| \cdot |y-z|$, we have:

$$\begin{aligned}
&\left| \left| \sum_{j=m+1}^{m+k} e(r^j h x) \right|^2 - \left| \sum_{j=m+1}^{m+k} e(r^j h \bar{x}) \right|^2 \right| \\
\le\ & \left(\left| \sum_{j=m+1}^{m+k} e(r^j h x) \right| + \left| \sum_{j=m+1}^{m+k} e(r^j h \bar{x}) \right| \right) \cdot \sum_{j=m+1}^{m+k} \left| e(r^j h x) - e(r^j h \bar{x}) \right| \\
\le\ & 2k \sum_{j=m+1}^{m+k} 2\pi r^j h |x - \bar{x}| \\
\le\ & 2k \sum_{j=m+1}^{m+k} 2\pi r^j r^{-m-k-1} \\
\le\ & c_1 k \qquad (8)
\end{aligned}$$

for some constant c_1 depending on r and h. Let $x_v \in T_p$. Then by construction of the measure μ we have

$$\mu\left(\left[x_v, x_v + b^{-\ell}\right)\right) = b^{-\ell/2},$$

since those $x \in [0,1]$ for which $\tau_p(x) \in [x_v, x_v + b^{-\ell})$ form an interval of length $b^{-\ell/2}$ (recall that we assumed that ℓ is even). Together with (8), this implies

$$\begin{aligned}
&\int_{[x_v, x_v+b^{-\ell})} \left| \sum_{j=m+1}^{m+k} e(r^j h x) \right|^2 d\mu(x) \\
\leq\ & \int_{[x_v, x_v+b^{-\ell})} \left| \sum_{j=m+1}^{m+k} e(r^j h x_v) \right|^2 + c_1 k \; d\mu(x) \\
=\ & \left(\left| \sum_{j=m+1}^{m+k} e(r^j h x_v) \right|^2 + c_1 k \right) \int_{[x_v, x_v+b^{-\ell})} d\mu(x) \\
=\ & b^{-\ell/2} \left(\left| \sum_{j=m+1}^{m+k} e(r^j h x_v) \right|^2 + c_1 k \right).
\end{aligned}$$

Since

$$\int_0^1 \left| \sum_{j=m+1}^{m+k} e(r^j h x) \right|^2 d\mu(x) = \sum_{v \in T_P(\ell)} \int_{[x_v, x_v+b^{-\ell})} \left| \sum_{j=m+1}^{m+k} e(r^j h x) \right|^2 d\mu(x),$$

we obtain

$$\int_0^1 \left| \sum_{j=m+1}^{m+k} e(r^j h x) \right|^2 d\mu(x) \leq c_1 k + b^{-\ell/2} \sum_{v \in T_P(\ell)} \left| \sum_{j=m+1}^{m+k} e(r^j h x_v) \right|^2, \tag{9}$$

and the main task for the proof of the lemma will be to estimate the sum on the right-hand side of (9).

Let

$$A(x,h,r,m,k) = \left| \sum_{j=m+1}^{m+k} e(r^j h x) \right|^2.$$

Since, in the sequel, the values r and h are fixed, and we will always use the expression with variables m and k, we abbreviate $A(x,h,r,m,k)$ by $A(x)$. We can rewrite the sum on the right-hand side of (9) as

$$\begin{aligned}
\sum_{v \in T_P(\ell)} A(x_v) &= \sum_{v \in T_P(\ell)} \sum_{j_1=m+1}^{m+k} \sum_{j_2=m+1}^{m+k} e((r^{j_2} - r^{j_1}) h x_v) \\
&\leq k b^{\ell/2} + 2 \sum_{i_1=1}^{k-1} \left| \sum_{v \in T_P(\ell)} \sum_{i_2=0}^{k-i_1-1} e\left(r^{m+1+i_2} (r^{i_1} - 1) h x_v \right) \right|, \qquad (10)
\end{aligned}$$

where the term $kb^{\ell/2}$ comes from the contribution of the diagonal $j_1 = j_2$, and where the summations in line (10) are obtained from those in the line above by substituting $i_1 = |j_2 - j_1|$, and then first summing over all j_1, j_2 for which $|j_2 - j_1|$ is fixed.

Since each sequence $v = v_1 v_2 \ldots v_\ell$ in $T_P(\ell)$ satisfies $v_q = v_{2q}$ for $q = 1, \ldots, \ell/2$, for every integer w we have

$$\sum_{v\in T_P(\ell)} e(wx_v) = \sum_{v\in T_P(\ell)} e\left(w\sum_{j=1}^{\ell} v_j b^{-j}\right) \tag{11}$$

$$= \sum_{v\in T_P(\ell)} e\left(w\sum_{\substack{q=1\\ q \text{ odd}}}^{\ell} v_q \left(\sum_{k=0}^{\lfloor \log_2(\ell/q)\rfloor} b^{-q2^k}\right)\right)$$

$$= \sum_{v\in T_P(\ell)} \prod_{\substack{q=1\\ q \text{ odd}}}^{\ell} \prod_{k=0}^{\lfloor \log_2(\ell/q)\rfloor} e\left(wv_q b^{-q2^k}\right)$$

$$= \prod_{\substack{q=1\\ q \text{ odd}}}^{\ell} \sum_{u=0}^{b-1} \prod_{k=0}^{\lfloor \log_2(\ell/q)\rfloor} e\left(uwb^{-q2^k}\right)$$

$$= \prod_{\substack{q=1\\ q \text{ odd}}}^{\ell} \sum_{u=0}^{b-1} e(uwM_q), \tag{12}$$

where

$$M_q = \sum_{k=0}^{\lfloor \log_2(\ell/q)\rfloor} b^{-q2^k}.$$

The internal $\sum_{u=0}^{b-1} e(uy)$ for some real y can be bounded by

$$\left|\sum_{u=0}^{b-1} e(uy)\right| = \left|\sum_{\substack{0\le u\le b-2\\ u \text{ even}}} e(uy)(1+e(y))\right|$$

$$\le \frac{b}{2}\Big|1+e(y)\Big| \qquad \text{if } b \text{ is even,} \tag{13}$$

and

$$\left|\sum_{u=0}^{b-1} e(uy)\right| \le 1 + \left|\sum_{\substack{1\le u\le b-2\\ u \text{ odd}}} e(uy)(1+e(y))\right|$$

$$\le 1 + \frac{b-1}{2}\Big|1+e(y)\Big| \qquad \text{if } b \text{ is odd.} \tag{14}$$

Since $|1+e(y)| = 2|\cos(\pi y)|$, and since the term in (14) is larger than that in (13), for all b we have the upper bound

$$\left|\sum_{u=0}^{b-1} e(uwM_q)\right| \le 1 + (b-1)\big|\cos(\pi w M_q)\big|$$

for the sums appearing in line (12).

Thus for (11) we have the estimate

$$\left|\sum_{v\in T_P(\ell)} e(wx_v)\right| \leq b^{\ell/2} \prod_{\substack{q=1\\ q \text{ odd}}}^{\ell} \left(\frac{1}{b}+\frac{b-1}{b}\left|\cos\left(\pi w M_q\right)\right|\right)$$

$$\leq b^{\ell/2} \prod_{\substack{\ell/2\leq q\leq \ell,\\ q \text{ odd}}} \left(\frac{1}{b}+\frac{b-1}{b}\left|\cos\left(\pi w M_q\right)\right|\right)$$

$$= b^{\ell/2} \prod_{\substack{\ell/2< q\leq \ell,\\ q \text{ odd}}} \left(\frac{1}{b}+\frac{b-1}{b}\left|\cos\left(\pi w b^{-q}\right)\right|\right), \tag{15}$$

where we used the crucial fact that for all q satisfying $\ell/q < 2$, we have $\lfloor \log_2(\ell/q)\rfloor = 0$, and thus $M_q = b^{-q}$. Note that we were allowed to simply remove some of the factors when changing from the first to the second line of the displayed formula, since all factors are trivially bounded by 1.

We will use (15) with $w = r^{m+1+i_2}(r^{i_1}-1)h$, with the ranges of i_1 and i_2 specified in (10). By our choice of ℓ for such w we have

$$\begin{aligned} wb^{-\ell} &\leq r^{m+k} h b^{-\ell} \\ &\leq r^{m+k} h b^{-((m+k+1)\log_b r)-\log_b h} \\ &\leq r^{-1} \leq \frac{1}{2}. \end{aligned}$$

Thus for such w we have

$$\prod_{\substack{q>\ell,\\ q \text{ odd}}} \left(\frac{1}{b}+\frac{b-1}{b}\left|\cos\left(\pi w b^{-q}\right)\right|\right) \geq \prod_{z=1}^{\infty} \left(\frac{1}{b}+\frac{b-1}{b}\left|\cos\left(\pi 2^{-z}\right)\right|\right) \geq c_2$$

for some constant $c_2 > 0$, and thus for such w the expression in line (15) is bounded by

$$c_2^{-1} \prod_{\substack{q>\ell/2,\\ q \text{ odd}}} \left(\frac{1}{b}+\frac{1}{b-1}\left|\cos\left(\pi w b^{-q}\right)\right|\right).$$

When we plug this estimate into (10) we obtain

$$\left|\sum_{v\in T_P(\ell)} A(x_v)\right| \leq kb^{\ell/2} +$$

$$+2c_2^{-1} b^{\ell/2} \sum_{i_1=1}^{k-1} \sum_{i_2=0}^{k-i_1-1} \prod_{\substack{q>\ell/2,\\ q \text{ odd}}} \left(\frac{1}{b}+\frac{b-1}{b}\left|\cos\left(\pi r^{m+1+i_2}(r^{i_1}-1)hb^{-q}\right)\right|\right). \tag{16}$$

We apply Lemma 6 to estimate the sums of products in this formula, and using the lemma with $L = r^{m+1}(r^{i_1}-1)h$ we obtain

$$\sum_{i_2=0}^{k-i_1-1} \prod_{\substack{q>\ell/2,\\ q \text{ odd}}} \left(\frac{1}{b}+\frac{1}{b-1}\left|\cos\left(\pi r^{m+1+i_2}(r^{i_1}-1)hb^{-q}\right)\right|\right) \leq 2k^{1-c_3} \tag{17}$$

for a constant $c_3 > 0$. Note that for the application of the lemma it was essential to ensure that $L \geq b^{\ell/2}$, which with our choice of L is $r^{m+1}(r^{i_1}-1)h \geq b^{\ell/2}$. However, this is true, since by the assumption $m \geq k+1+2\log_r b$ and our choice of $\ell \leq (((m+k+1)\log_b r) + \log_b h) + 2$ we have

$$\begin{aligned} r^{m+1}(r^{i_1}-1)h &\geq hr^m \\ &\geq hr^{(m+k+1+2\log_r b)/2} \\ &= hb^{((m+k+1)\log_b r)/2} b^1 \\ &\geq b^{((m+k+1)\log_b r)/2+(\log_b h)+1} \\ &\geq b^{\ell/2}. \end{aligned} \tag{18}$$

These formulas show where the difficulties come from in our setting, as compared to Cassels' and Schmidt's work. Since we cannot control those terms in the product where M_q is complicated, we have to restrict the product to relatively large values of q, where we have the simple situation that $M_q = b^{-q}$. However, since for this reason it is necessary that our product starts at a large value of q, in order to be able to apply Lemma 6 we have to make sure that the frequencies (denoted by L in the lemma) are large, which in turn requires that the summation in Lemma 7 cannot start at 1, but only at a value of j which is relatively large in comparison with the summation range k.

Using (17) for (16), we obtain

$$\left| \sum_{v \in T_P(\ell)} A(x_v) \right| \leq b^{\ell/2}(k + 4c_2^{-1}k^{2-c_3}).$$

Combining this with (9) we finally obtain

$$\int_0^1 \left| \sum_{j=m+1}^{m+k} e(r^j hx) \right|^2 d\mu(x) \leq (c_1 k + k + 4c_2^{-1}k^{2-c_3}) \leq k^{2-c_4}$$

for a constant $c_4 > 0$ and all sufficiently large k. □

Proof of Theorem 2. Let $\mathcal{R}$ be the set of real numbers in the unit interval whose expansion in base b is in T_P,

$$\mathcal{R} = \{x_v : v \in T_P\}.$$

Note that by construction μ is supported on $\mathcal{R}$, so $\mu([0,1]\backslash\mathcal{R}) = 0$. Consider an integer r multiplicatively independent to b. To prove that μ-almost all elements of $\mathcal{R}$ are normal to base r, by Weyl's criterion we have to show that for μ-almost all $x \in [0,1]$ we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e(r^n hx) = 0$$

for all integers $h > 0$.

For $k \geq 1$, set $m_k = \left\lceil e^{\sqrt{k}} \right\rceil$. Furthermore, we define $M_0 = 0$ and

$$M_k = m_1 + \cdots + m_k, \qquad k \geq 1.$$

For a fixed positive integer h, we define sets

$$E_k = \left\{ x \in [0,1] : \frac{1}{m_k} \left| \sum_{n=M_{k-1}+1}^{M_k} e(r^n h x) \right| > 1/k \right\}.$$

The summation has m_k terms and, for sufficiently large k, we have

$$m_k \leq M_{k-1} + 1 + 2\log_r b.$$

To see this, notice that $m_k \approx e^{\sqrt{k}}$ and $\sqrt{k}e^{\sqrt{k}} < M_k$. Thus, for all sufficiently large k we can apply Lemma 7, and by an application of Chebyshev's inequality we have

$$\mu(E_k) \leq \frac{k^2}{m_k^c},$$

where $c > 0$ is the constant from the conclusion of Lemma 7. By the rapid growth of the sequence $(m_k)_{k\geq 1}$, this implies

$$\sum_{k=1}^{\infty} \mu(E_k) < \infty,$$

and thus by the first Borel–Cantelli lemma, μ-almost surely only finitely many events E_k occur, so that in particular μ-almost surely we have

$$\frac{1}{m_k} \left| \sum_{n=M_{k-1}+1}^{M_k} e(r^n h x) \right| \to 0 \qquad \text{as } k \to \infty.$$

It is easily seen that this also implies that μ-almost surely

$$\lim_{k\to\infty} \frac{1}{M_k} \left| \sum_{n=1}^{M_k} e(r^n h x) \right| = 0.$$

Finally, by the subexponential growth of $(M_k)_{k\geq 1}$ for all sufficiently large N, there is a value of k such that $|N - M_k| = o(N)$. This implies

$$\lim_{N\to\infty} \frac{1}{N} \left| \sum_{n=1}^{N} e(r^n h x) \right| = 0, \qquad \mu\text{-almost surely.} \tag{19}$$

Clearly, there are only countably many possible values of h and r. Thus μ-almost all numbers $x \in [0,1]$ have the property that (19) is true for all positive integers h and for all integers $r \geq 2$ which are multiplicatively independent of b. This proves the theorem. □

4. Proof of Theorem 3

We start with the first part of the theorem. It turns out that it is sufficient to relax the conclusion of the theorem to simple normality.

Lemma 8. *Assume that for every $n \geq 1$, the random variable X_n is uniformly distributed on $\{0, \ldots, b-1\}$. Assume furthermore that there exists a function $g : \mathbb{N} \mapsto \mathbb{R}$ which is monotonically increasing to ∞ such that for all sufficiently large n, the random variables*

$$X_n,\ X_{n+1},\ \ldots,\ X_{n+\lceil g(n)\log\log n\rceil} \tag{20}$$

are mutually independent. Let x be the real number whose expansion in base b is given by $x = 0.X_1X_2X_3\dots$. Then $\mathbb{P}$-almost surely the number x is simply normal to base b.

We take Lemma 8 for granted, and show that it implies the first part of Theorem 3. Normality of a number to base b is equivalent to simple normality to all the bases $b, b^2, b^3, \dots$. Two consecutive random variables such as X_1, X_2 with values in $\{0, \dots, b-1\}$ define one random variable Y_1 with value in $\{0, \dots, b^2-1\}$ by setting $Y_1 = bX_1 + X_2$. Similarly X_3 and X_4 define Y_2, etc. Then the number $(0.X_1X_2\dots)_b$ has base-b^2 expansion $(0.Y_1Y_2\dots)_{b^2}$. Furthermore, since by the assumption of Theorem 3 the random variables

$$X_{2n-1}, \dots, X_{2n-1+\lceil g(2n-1)\log\log(2n-1)\rceil}$$

are independent for sufficiently large n, this implies that there is a function $\hat{g}$ which is monotonically increasing to ∞ such that the random variables

$$Y_n, \dots, Y_{n+\lceil \hat{g}(n)\log\log n\rceil}$$

are independent as well for all sufficiently large n. So the sequence $Y_1, Y_2, \dots$ also satisfies the assumptions of Lemma 8. Accordingly, if Lemma 8 is true, then almost surely the number $(0.Y_1Y_2\dots)_{b^2} = (0.X_1X_2\dots)_b$ is simply normal to base b^2. In the same way we can show that Lemma 8 implies that almost surely $(0.X_1X_2\dots)_b$ is simply normal in bases b^3, that almost surely it is simply normal to base b^4, and so on. Since there are only countably many bases $b, b^2, b^3, b^4, \dots$, it means that almost surely $(0.X_1X_2\dots)_b$ is simply normal to all bases $b, b^2, b^3, \dots$, which is equivalent to saying that almost surely $(0.X_1X_2\dots)_b$ is normal to base b. Thus, to establish the first part of Theorem 3, it is sufficient to prove Lemma 8.

Proof of Lemma 8. Let $u \in \{0, \dots, b-1\}$ be a digit. Assume that $\varepsilon > 0$ is fixed. Let $\theta = 1 + \varepsilon$, and for $j \geq 1$, define

$$N_j = \{n \geq 1 : \ \theta^{j-1} \leq n < \theta^j\}.$$

We partition N_j into disjoint sets of $m_j = \lceil \varepsilon^{-2} \log j \rceil$ consecutive integers, where for simplicity of writing we assume that m_j divides $\#N_j$ (so that we do not need to use one set of smaller cardinality at the end). We denote these sets by $M_1^{(j)}, \dots, M_{r(j)}^{(j)}$, where $r(j) = \#N_j/m_j$. It is easily verified that condition (20) implies that all the random variables $\{X_n : \ n \in M_i^{(j)}\}$ are mutually independent for all $i \in \{1, \dots, r(j)\}$, provided that j is sufficiently large; this follows from the fact that the block length m_j is of order roughly $\varepsilon^{-2}\log\log n$ for $n \in N_j$, while by assumption independence holds for random variables whose indices are within distance $g(n)\log\log n$ of each other, where $g(n) \to \infty$.

By Hoeffding's inequality (see for example [6, Theorem 2.16]) we have

$$\begin{aligned} \mathbb{P}\left(\left|\frac{1}{m_j}\sum_{n \in M_i^{(j)}} \mathbf{1}(X_n = u) - \frac{1}{b}\right| > \varepsilon\right) &\leq 2e^{-2\varepsilon^2 m_j} \\ &\leq \frac{2}{j^2} \qquad (21) \end{aligned}$$

for all $i \in \{1, \dots, r(j)\}$ and all sufficiently large j, where we used $m_j \geq \varepsilon^{-2}\log j$. Here, and in the sequel, we write $\mathbf{1}(E)$ for the indicator function of an event E. Let

$Z_{i,j}$ denote the random variable $\frac{1}{m_j}\sum_{n\in M_i^{(j)}} \mathbf{1}(X_n = u) - \frac{1}{b}$, for $i \in \{1,\dots,r(j)\}$. Then trivially $|Z_{i,j}| \leq 1$. Thus we have

$$\begin{aligned}\mathbb{E}\big(|Z_{i,j}| \cdot \mathbf{1}\left(|Z_{i,j}| > \varepsilon\right)\big) &\leq \mathbb{E}\left(\mathbf{1}\left(|Z_{i,j}| > \varepsilon\right)\right) \\ &= \mathbb{P}(|Z_{i,j}| > \varepsilon) \\ &\leq \frac{2}{j^2}\end{aligned}$$

for sufficiently large j, as calculated in (21). By linearity of the expectation, this implies

$$\mathbb{E}\left(\sum_{i=1}^{r(j)} \Big(|Z_{i,j}| \cdot \mathbf{1}\left(|Z_{i,j}| > \varepsilon\right)\Big)\right) \leq \frac{2r(j)}{j^2},$$

and thus by Markov's inequality

$$\mathbb{P}\left(\sum_{i=1}^{r(j)} |Z_{i,j}| \cdot \mathbf{1}\left(|Z_{i,j}| > \varepsilon\right) > \varepsilon r(j)\right) \leq \frac{2}{\varepsilon j^2}. \tag{22}$$

Note that

$$|Z_{i,j}| \leq \varepsilon + |Z_{i,j}| \cdot \mathbf{1}\left(|Z_{i,j}| > \varepsilon\right).$$

Thus

$$\sum_{i=1}^{r(j)} |Z_{i,j}| \leq \varepsilon r(j) + \sum_{i=1}^{r(j)} |Z_{i,j}| \cdot \mathbf{1}\left(|Z_{i,j}| > \varepsilon\right),$$

and (22) implies that

$$\mathbb{P}\left(\sum_{i=1}^{r(j)} |Z_{i,j}| > 2\varepsilon r(j)\right) \leq \frac{2}{\varepsilon j^2}. \tag{23}$$

Note that if $\sum_{i=1}^{r(j)} |Z_{i,j}| \leq 2\varepsilon r(j)$, then

$$\left|\frac{1}{\# N_j} \#\{n \in N_j :\ X_n = u\} - \frac{1}{b}\right| \leq 2\varepsilon. \tag{24}$$

The exceptional probabilities in (23) form a convergent series when summing over j, so by the first Borel–Cantelli lemma with probability one only finitely many of the corresponding events occur. Accordingly, $\mathbb{P}$-almost surely we have (24) for all sufficiently large j. From this it is easy to see that

$$\left|\frac{1}{N} \#\{n \leq N :\ X_n = u\} - \frac{1}{b}\right| \leq 4\varepsilon$$

almost surely for all sufficiently large N, where it is important that $N_{j+1} \approx (1+\varepsilon)N_j$. Since we can choose ε arbitrarily close to zero, this proves Lemma 8. □

Proof of the second part of Theorem 3. For the second part of the theorem, let $(Z_{j,m})_{j\geq 1, m\geq 0}$ be an array of independent, identically distributed (i.i.d.) random variables having uniform distribution on $A = \{0,\dots b-1\}$. Clearly it is sufficient to prove the second part of Theorem 3 for those values of K which are a positive (integral) powers of 2, so we will assume that K is of this form.

We define the digits X_n of a number $x = (0.X_1X_2X_3\dots)_b$ by setting $X_n = Z_{j,m}$, where $j = \lfloor \log_2 n \rfloor$, and where m is the unique integer for which $n \equiv m \mod 2Kr$

and r is defined as the largest positive integer which is a power of 2 and for which $2^{(2^r)} \le n$ (this definition only works for $n \ge 4$, so we may set $X_1X_2X_3 = 000$). So, for example, when $K = 1$, then for $n \in \{16, \dots, 31\}$ we have $j = 4$ and $r = 2$, and thus $X_{16}, X_{17}, \dots, X_{31}$ is the pattern $Z_{4,0}Z_{4,1}Z_{4,2}Z_{4,3}$ being repeated four times. Or when $n \in \{2^{16}, \dots, 2^{17} - 1\}$, then $j = 16$ and $r = 4$, and so $X_{2^{16}}, \dots, X_{2^{17}-1}$ is the pattern $Z_{16,0}, Z_{16,1}, \dots, Z_{16,8}$ being repeated $2^j/(2rK) = 2^{16}/8$ times. This example is constructed in such a way that the digits $X_n, \dots, X_{n+2Kr-1}$ are mutually independent, where by definition $2^{(2^r)} \ge n$, and thus

$$r \ge (\log_2 \log_2 n)/2 \ge 2(\log \log n)/3,$$

which shows that (5) indeed holds for all sufficiently large n.

We will now show that the random number x almost surely is not simply normal, which will follow from the fact that there is a digit $u \in A$ for which the ratio

$$\frac{1}{N}\Big|\{n \le N : \ X_n = u\}\Big|$$

does not converge to $1/b$. For simplicity, assume that $u = 0$. For $j \ge 1$, set $N_j = \{2^j, \dots, 2^{j+1} - 1\}$. From the construction of our sequence $(X_n)_{n \ge 1}$, it is easily seen that (for sufficiently large j) the block of digits $X_{2^j}, \dots, X_{2^{j+1}-1}$ consists of the block $Z_{j,0} \dots Z_{j,2Kr-1}$ for some appropriate value of $r = r(j)$, which is repeated $2^j/(2Kr)$ times. Thus we have

$$\#\Big\{n \in N_j : \ X_n = 0\Big\} = \frac{2^j}{2Kr} \ \#\Big\{0 \le m < 2Kr : \ Z_{j,m} = 0\Big\}. \tag{25}$$

We choose a "small" fixed value of $\varepsilon > 0$, and define events

$$E_j = \left\{ \left| \frac{1}{2^j} \cdot \# \{n \in N_j : \ X_n = 0\} - \frac{1}{b} \right| \ge \varepsilon \right\}.$$

By (25) we have

$$\mathbb{P}(E_j) = \mathbb{P}\left\{ \left| \frac{1}{2Kr} \cdot \#\{0 \le m < 2Kr : \ Z_{j,m} = 0\} - \frac{1}{b} \right| \ge \varepsilon \right\}.$$

To estimate $\mathbb{P}(E_j)$, note that we have

$$\#\Big\{0 \le m < 2Kr : \ Z_{j,m} = 0\Big\} = \sum_{m=0}^{2Kr-1} \mathbf{1}(Z_{j,m} = 0),$$

where $\mathbf{1}(Z_{j,m} = 0)$ is the indicator function of the event $Z_{j,m} = 0$, and where accordingly $(\mathbf{1}(Z_{j,m} = 0))_{0 \le m \le r(j)-1}$ is a sequence of i.i.d. random variables with mean $1/b$ and variance $\sigma_b^2 = \frac{1}{b}\left(1 - \frac{1}{b}\right)$ (where we use the fact that the independence property of the indicators $\mathbf{1}(Z_{j,m} = 0)$ is inherited from the independence assumption on the $Z_{j,m}$'s). Accordingly, $\sum_{m=0}^{2Kr-1} \mathbf{1}(Z_{j,m} = 0)$ has binomial distribution $B(r, 1/b)$. Using standard estimates for the tail probabilities of the binomial distribution, we can show that $\mathbb{P}(E_j)$ is of order roughly $e^{-c(\varepsilon)2Kr}$ for large r, where $c(\varepsilon) = (1 + o(1))\varepsilon^2\sigma_b^2/2$ as $\varepsilon \to 0$; this can be deduced either from lower bounds for the tail of the binomial distribution, together with a linearization of the Kullback–Leibler distance (see [2, Lemma 4.7.2] or [10, Theorem 11.1.3]), or from a comparison of the tail of a binomial distribution with the tail of the normal distribution (see [17]). Note in particular that $c(\varepsilon)$ goes to zero as a function of ε. Thus, if ε was chosen so small that $2Kc(\varepsilon) < 1/8$, then we certainly have

$$\mathbb{P}(E_j) \geq e^{-r/4}$$

for all sufficiently large j. By the definition of r we have $r = r(j) \geq \frac{\log_2(j+1)}{2} \geq \frac{\log j}{2}$ for all sufficiently large j. Thus,

$$\mathbb{P}(E_j) \geq e^{-r/4} \geq e^{-(\log j)/2} = \frac{1}{\sqrt{j}}$$

for all sufficiently large j. This allows us to deduce that

$$\sum_{j=1}^{\infty} \mathbb{P}(E_j) = +\infty.$$

Note that the events $(E_j)_{j\geq 1}$ are mutually independent, since E_j only depends on random variables $Z_{j,m}$ whose first index is j. Thus, by the second Borel–Cantelli lemma with probability one, infinitely many events E_j occur. However, this means that with probability one the digit 0 does not have the correct asymptotic frequency within the blocks of digits with indices in N_j. Note that the length of the blocks N_j grows so quickly that

$$\lim_{j\to\infty} \frac{\#N_j}{\#N_1 + \cdots + \#N_{j-1}} = 1 > 0,$$

so that the contribution of digits contained in N_j is not (asymptotically) negligible in comparison with the contribution of the digits from all the previous blocks. From this, it is easy to deduce that

$$\frac{1}{N} \cdot \#\Big\{n \leq N : \ X_n = 0\Big\} \not\to \frac{1}{b}$$

for $\mathbb{P}$-almost all x, which proves the second part of the theorem. □

5. Proof of Theorem 4

In this section, finite sequences of digits are called *words*. If $w = a_1 \cdots a_n$ is a word of length n and i, j are two integers such that $1 \leq i \leq j \leq n$, the word $a_i \cdots a_j$ is called either the *block* of w from position i to position j or the *block* of length $j-i+1$ at position i in w. Let us fix an integer k, and let us consider the finite set I of integers $I = \{i : 1 \leq i \leq (p_1p_2)^{k+1}\}$ and its subset $J = \{j \in I : p_1^{k+1}|j \text{ or } p_2^{k+1}|j\}$. The set J can be decomposed as $J = J_1 \cup J_2$, where $J_1 = \{p_1^{k+1}m : 1 \leq m \leq p_2^{k+1}\}$, $J_2 = \{p_2^{k+1}m : 1 \leq m \leq p_1^{k+1}\}$, and $J_1 \cap J_2 = \{(p_1p_2)^{k+1}\}$. It follows that J has cardinality $p_1^{k+1} + p_2^{k+1} - 1$. Let ρ_J be the function which maps each word w of length $(p_1p_2)^{k+1}$ to the word of length $(p_1p_2)^{k+1} - p_1^{k+1} - p_2^{k+1} + 1$ obtained by removing from w symbols at positions in J. Formally, if the word w is equal to $a_1 \cdots a_{(p_1p_2)^{k+1}}$, which is written $\prod_{i\in I} a_i$, the word $\rho_J(w)$ is equal to $\prod_{i\in I\setminus J} a_i$. Note that

$$(p_1p_2)^{k+1} - p_1^{k+1} - p_2^{k+1} + 1 = (p_1^{k+1} - 1)(p_2^{k+1} - 1).$$

We introduce a last notation. Let σ be a permutation of $\{1, \ldots, n\}$. It induces a permutation, also denoted by σ, of A^n defined by $\sigma(a_1 \cdots a_n) = a_{\sigma(1)} \cdots a_{\sigma(n)}$ for each word $a_1 \cdots a_n$ of length n.

Lemma 9. *Suppose that $\tau_P(x)$ is decomposed $\tau_P(x) = w_1w_2w_3\cdots$, where each word w_i has length $(p_1p_2)^{k+1}$. There is a permutation σ of*

$$\{1,\ldots,(p_1^{k+1}-1)(p_2^{k+1}-1)\}$$

such that for each integer i, the word $\sigma(\rho_J(w_i))$ is equal to the concatenation $\prod_{i_1=0}^{k}\prod_{i_2=0}^{k} u_{i,i_1,i_2}$ of the $(k+1)^2$ blocks of x, where each block u_{i,i_1,i_2} starts at position $(p_1-1)(p_2-1)p_1^{i_1}p_2^{i_2}(i-1)+1$ and has length $(p_1-1)(p_2-1)p_1^{i_1}p_2^{i_2}$.

Note for each $0 \le i_1, i_2 \le k$, the length of u_{i,i_1,i_2} is the same for all $i \ge 1$, and note that

$$x = u_{1,i_1,i_2}u_{2,i_1,i_2}u_{3,i_1,i_2}\cdots$$

is the decomposition of x in blocks of length $(p_1-1)(p_2-1)p_1^{i_1}p_2^{i_2}$. What is important is that both ρ_J and σ are fixed and do not depend on w_i. The permutation σ is not made explicit by the statement of the lemma, because it is not used in the sequel but it can be easily recovered from the proof of the lemma.

Proof. The positions in $\tau_P(x)$ of word w_i are the integers from $(p_1p_2)^{k+1}(i-1)+1$ to $(p_1p_2)^{k+1}i$, so they are of the form $(p_1p_2)^{k+1}(i-1)+j$ where $1 \le j \le (p_1p_2)^{k+1}$. By definition, the function ρ_J removes the symbols at a position of the form $(p_1p_2)^{k+1}(i-1)+j$, where $j \in J$. It follows that the word $\rho_J(w_i)$ contains the symbols at a position $(p_1p_2)^{k+1}(i-1)+j$, where j can be written $j = p_1^{i_1}p_2^{i_2}m$ for $0 \le i_1, i_2 \le k$, and m divisible by neither p_1 nor p_2.

Fix i_1 and i_2 such that $0 \le i_1, i_2 \le k$, and consider all symbols of w_i at positions of the form $(p_1p_2)^{k+1}(i-1)+j$, where $j = p_1^{i_1}p_2^{i_2}m$ for $0 \le i_1, i_2 \le k$ and m is divisible by neither p_1 or p_2. Note that in p_1p_2 consecutive integers, exactly p_1+p_2-1 of them are divisible by either p_1 or p_2. Since $1 \le m \le p_1^{k+1-i_1}p_2^{k+1-i_2}$, there are exactly $(p_1-1)(p_2-1)p_1^{k-i_1}p_2^{k-i_2}$ possible values for m. By definition of the Toeplitz transform τ_P, these symbols are at consecutive positions in x. More precisely, they are the symbols from position $(p_1-1)(p_2-1)p_1^{k-i_1}p_2^{k-i_2}(i-1)+1$ to position $(p_1-1)(p_2-1)p_1^{k-i_1}p_2^{k-i_2}i$ in x. Calling this word $u_{i,k-i_1,k-i_2}$ for each $0 \le i_1, i_2 \le k$ provides the decomposition of $\rho_J(w_i)$. □

We continue with two lemmas that show that condition (II) is quite robust. First we prove that if condition (II) of Theorem 4 holds for words of some given lengths, then it also holds for all shorter words.

Lemma 10. *Fix an integer $k \ge 0$, and let $(\ell_{i_1,i_2})_{0\le i_1,i_2\le k}$ be a family of nonnegative integers. If*

$$\lim_{N\to\infty}\frac{1}{N}\#\left\{n : \begin{array}{l} 1 \le n \le N, \text{ and for all } 0 \le i_1, i_2 \le k, \\ u_{i_1,i_2} \text{ occurs in } x \text{ at position } (p_1-1)(p_2-1)p_1^{i_1}p_2^{i_2}n \end{array}\right\}$$

is equal to $b^{-\sum_{0\le i_1,i_2\le k}|u_{i_1,i_2}|}$ for all families $(u_{i_1,i_2})_{0\le i_1,i_2\le k}$ of finite words such that $|u_{i_1,i_2}| = \ell_{i_1,i_2}$, then it also holds for all families $(v_{i_1,i_2})_{0\le i_1,i_2\le k}$ of finite words such that $|v_{i_1,i_2}| \le \ell_{i_1,i_2}$.

Proof. The result follows from the equality

$$\#\left\{n: \begin{array}{l} 1 \le n \le N, \text{ and for all } 0 \le i_1, i_2 \le k, \\ v_{i_1,i_2} \text{ occurs in } x \text{ at position } (p_1-1)(p_2-1)p_1^{i_1}p_2^{i_2}n \end{array}\right\}$$
$$= \sum_{w_{i_1,i_2}} \#\left\{n: \begin{array}{l} 1 \le n \le N, \text{ and for all } 0 \le i_1, i_2 \le k, \\ v_{i_1,i_2}w_{i_1,i_2} \text{ occurs in } x \text{ at position } (p_1-1)(p_2-1)p_1^{i_1}p_2^{i_2}n \end{array}\right\}$$

where the summation ranges over all families $(w_{i_1,i_2})_{0\le i_1,i_2\le k}$ of finite words such that $|w_{i_1,i_2}| = \ell_{i_1,i_2} - |v_{i_1,i_2}|$ for each $0 \le i_1, i_2 \le k$. □

The following lemma shows that offsets can be freely added in condition (II) of Theorem 4.

Lemma 11. *Fix a nonnegative integer k, and let $(\delta_{i_1,i_2})_{0\le i_1,i_2\le k}$ be a family of integers. Then, for every family $(u_{i_1,i_2})_{0\le i_1,i_2\le k}$ of finite words*

$$\lim_{N\to\infty} \frac{1}{N}\#\left\{n: \begin{array}{l} 1 \le n \le N, \text{ and for all } 0 \le i_1, i_2 \le k, \\ u_{i_1,i_2} \text{ occurs in } x \text{ at position } (p_1-1)(p_2-1)p_1^{i_1}p_2^{i_2}n \end{array}\right\}$$

is equal to $b^{-\sum_{0\le i_1,i_2\le k}|u_{i_1,i_2}|}$ if and only if for every family $(v_{i_1,i_2})_{0\le i_1,i_2\le k}$ of finite words

$$\lim_{N\to\infty} \frac{1}{N}\#\left\{n: \begin{array}{l} 1 \le n \le N, \text{ and for all } 0 \le i_1, i_2 \le k, \\ v_{i_1,i_2} \text{ occurs in } x \text{ at position } (p_1-1)(p_2-1)p_1^{i_1}p_2^{i_2}n + \delta_{i_1,i_2} \end{array}\right\}$$

is equal to $b^{-\sum_{0\le i_1,i_2\le k}|v_{i_1,i_2}|}$.

Note that if some δ_{i_1,i_2} is negative, then the position $(p_1-1)(p_2-1)p_1^{i_1}p_2^{i_2}n+\delta_{i_1,i_2}$ may not exist for small values of n because it is negative. However, this does not really matter because we are considering a limit when N goes to infinity.

Proof. In order to replace δ_{i_1,i_2} by $\delta_{i_1,i_2}+1$ for a single pair (i_1, i_2), it suffices to sum up for all possible symbols a in A, all equalities of condition (II) for av_{i_1,i_2} to get the equality of condition (II) for v_{i_1,i_2} with $\delta_{i_1,i_2}+1$.

In order to replace δ_{i_1,i_2} by $\delta_{i_1,i_2} - (p_1-1)(p_2-1)p_1^{i_1}p_2^{i_2}$ for all pairs (i_1, i_2), it suffices to replace n by $n-1$. These two replacements allow us to get any possible change of offsets. □

We now come to the proof of Theorem 4.

Proof of Theorem 4. We first prove that condition I implies condition II. We suppose that $\tau_P(x)$ is normal. By Lemma 10, it suffices to prove condition II when the length of each word u_{i_1,i_2} is $(p_1-1)(p_2-1)p_1^{i_1+\ell}p_2^{i_2+\ell}$ for some fixed but arbitrary integer ℓ. Consider the decomposition $\tau_P(x) = w_1w_2w_3\cdots$, where each word w_i has length $(p_1p_2)^{k+1}$, and let y be the sequence $y = w'_1w'_2w'_3\cdots$, where each word w'_i is given by $w'_i = \rho_J(w_i)$. Since w'_i is obtained from w_i by removing symbols at fixed positions, the sequence y is also normal. Here we use the result that selecting digits along a periodic sequence of positions preserves normality [19]. By Lemma 9, the symbols of each word w'_i can be rearranged by a permutation σ to the concatenation

$$\prod_{i_1=0}^{k}\prod_{i_2=0}^{k} z_{i,i_1,i_2},$$

where z_{i,i_1,i_2} is the block in x of length $(p_1-1)(p_2-1)p_1^{i_1}p_2^{i_2}$ that starts at position $(p_1-1)(p_2-1)p_1^{i_1}p_2^{i_2}(i-1)+1$. It follows that the concatenation $w_iw_{i+1}\cdots w_{i+(p_1p_2)^\ell-1}$ of $(p_1p_2)^\ell$ such words can be also rearranged by another permutation to the concatenation

$$\prod_{i_1=0}^{k}\prod_{i_2=0}^{k}\tilde{z}_{i,i_1,i_2},$$

where $\tilde{z}_{i,i_1,i_2}$ is the block in x of length $(p_1-1)(p_2-1)p_1^{i_1+\ell}p_2^{i_2+\ell}$ that starts at position $(p_1-1)(p_2-1)p_1^{i_1+\ell}p_2^{i_2+\ell}(i-1)+1$.

Since y is normal, all words of length $(p_1^{k+1}-1)(p_2^{k+1}-1)(p_1p_2)^\ell$ occur with the same frequency in y as a concatenation $w_iw_{i+1}\cdots w_{i+(p_1p_2)^\ell-1}$. It follows that for every family $(u_{i_1,i_2})_{0\le i_1,i_2\le k}$ of finite words such that u_{i_1,i_2} has length

$$(p_1-1)(p_2-1)p_1^{i_1+\ell}p_2^{i_2+\ell}$$

for $0\le i_1,i_2\le k$, the limit

$$\lim_{N\to\infty}\frac{1}{N}\#\left\{n: \begin{array}{l} 1\le n\le N, \text{ and for all } 0\le i_1,i_2\le k, \\ u_{i_1,i_2} \text{ occurs in } x \text{ at position } (p_1-1)(p_2-1)p_1^{i_1}p_2^{i_2}n+1 \end{array}\right\}$$

has the same value. By Lemma 11, the offset $+1$ in the position

$$(p_1-1)(p_2-1)p_1^{i_1}p_2^{i_2}n+1$$

can be removed, and the result is established.

We now prove the converse. By a result of Long [13], a number x is normal to base b if and only if there exists positive integers $m_1<m_2<\cdots$ such that x is simply normal to the bases b^{m_i}, $i\ge 1$. Therefore, it is sufficient to prove that, for infinitely many length k, all words of length k have the expected frequency in $\tau_P(x)$. We claim that for each word w of length $(p_1p_2)^\ell$ for any integer ℓ, the frequency of w at positions multiple of $(p_1p_2)^\ell$ is the expected one, namely $|A|^{-(p_1p_2)^\ell}$. Now suppose that ℓ is fixed, and let k be an integer to be fixed later. Factorize $\tau_P(x)=w_1w_2w_3\cdots$, where each word w_i has length $(p_1p_2)^{k+1}$, and let y be the sequence $y=w'_1w'_2w'_3\cdots$, where $w'_i=\rho_J(w_i)$. By Lemma 9, and by the hypothesis of condition II, the sequence y is normal. For each integer i, the word w_i is obtained from w'_i by inserting $p_1^{k+1}+p_2^{k+1}-1$ symbols. Each word w_i contains $(p_1p_2)^{k+1-\ell}$ blocks of size $(p_1p_2)^\ell$. These $p_1^{k+1}+p_2^{k+1}-1$ inserted symbols can spoil at most $p_1^{k+1}+p_2^{k+1}-1$ blocks of size $(p_1p_2)^\ell$ but this number of possible spoiled blocks becomes negligible with respect to the total number of such blocks when k goes to infinity. Hence normality follows by taking k great enough. This concludes the proof that $\tau_P(x)$ is normal. □

Acknowledgments

This paper was initiated in November 2016 during the workshop "Normal Numbers: Arithmetic, Computational and Probabilistic Aspects" at the Erwin Schrödinger International Institute for Mathematics and Physics (ESI) in Vienna. We thank the ESI for bringing the three of us together, and for providing a stimulating atmosphere for mathematical discussions.

The authors are very grateful to the anonymous referee for reading the first version of this paper with exceptional accurateness, and for making many suggestions for possible improvements. The comments of the referee helped significantly to improve the presentation of the results and proofs in this paper.

References

[1] Christoph Aistleitner, *Metric number theory, lacunary series and systems of dilated functions*, Uniform distribution and quasi-Monte Carlo methods, Radon Ser. Comput. Appl. Math., vol. 15, De Gruyter, Berlin, (2014), pp. 1–16. MR3287357

[2] Robert B. Ash, *Information theory*, Dover Publications, Inc., New York, (1990). Corrected reprint of the 1965 original. MR1088248

[3] Alan Baker, *Transcendental number theory*, Cambridge University Press, London-New York, (1975). MR0422171

[4] Verónica Becher and Olivier Carton, *Normal numbers and computer science*, Sequences, groups, and number theory, Trends Math., Birkhäuser/Springer, Cham, (2018), pp. 233–269. MR3799929

[5] Verónica Becher, Olivier Carton, and Pablo Ariel Heiber, *Finite-state independence*, Theory Comput. Syst. **62** (2018), no. 7, 1555–1572, DOI 10.1007/s00224-017-9821-6. MR3832108

[6] Bernard Bercu, Bernard Delyon, and Emmanuel Rio, *Concentration inequalities for sums and martingales*, SpringerBriefs in Mathematics, Springer, Cham, (2015). MR3363542

[7] E. Borel. Les probabilités dénombrables et leurs applications arithmétiques. *Rendiconti del Circolo Matematico di Palermo*, **27** (1909) 247–271.

[8] Yann Bugeaud, *Distribution modulo one and Diophantine approximation*, Cambridge Tracts in Mathematics, vol. 193, Cambridge University Press, Cambridge, (2012). MR2953186

[9] J. W. S. Cassels, *On a problem of Steinhaus about normal numbers*, Colloq. Math. **7** (1959), 95–101, DOI 10.4064/cm-7-1-95-101. MR0113863

[10] Thomas M. Cover and Joy A. Thomas, *Elements of information theory*, 2nd ed., Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, (2006). MR2239987

[11] Harry Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation*, Math. Systems Theory **1** (1967), 1–49, DOI 10.1007/BF01692494. MR0213508

[12] Konrad Jacobs and Michael Keane, $0-1$-*sequences of Toeplitz type*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **13** (1969), 123–131, DOI 10.1007/BF00537017. MR0255766

[13] Calvin T. Long, *Note on normal numbers*, Pacific J. Math. **7** (1957), 1163–1165. MR0091959

[14] Walter Philipp, *Empirical distribution functions and strong approximation theorems for dependent random variables. A problem of Baker in probabilistic number theory*, Trans. Amer. Math. Soc. **345** (1994), no. 2, 705–727, DOI 10.2307/2154995. MR1249469

[15] S. S. Pillai, *On normal numbers*, Proc. Indian Acad. Sci., Sect. A. **12** (1940), 179–184. MR0002324

[16] Wolfgang M. Schmidt, *Über die Normalität von Zahlen zu verschiedenen Basen* (German), Acta Arith. **7** (1961/1962), 299–309, DOI 10.4064/aa-7-3-299-309. MR0140482

[17] Eric V. Slud, *Distribution inequalities for the binomial law*, Ann. Probability **5** (1977), no. 3, 404–412. MR0438420

[18] R. Tijdeman, *On integers with many small prime factors*, Compositio Math. **26** (1973), 319–330. MR0325549

[19] Donald D. Wall, *Normal Numbers*, ProQuest LLC, Ann Arbor, MI, 1950. Thesis (Ph.D.)–University of California, Berkeley. MR2937990

Institute of Analysis and Number Theory, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria

Email address: aistleitner@math.tugraz.at

Departamento de Computación and ICC CONICET, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Pabellón I, Ciudad Universitaria, C1428EGA Buenos Aires, Argentina

Email address: vbecher@dc.uba.ar

Institut de Recherche en Informatique Fondamentale, Université Paris Diderot, Case 7014, 75205 Paris Cedex 13, France

Email address: olivier.carton@irif.fr