Uniform Distribution Theory **9** (2014), no.2, 93–101

uniform distribution theory

NORMAL NUMBERS AND CANTOR EXPANSIONS

Ferdinánd Filip — Jan Šustek

ABSTRACT. A real number is called normal if every block of digits in its expansion occurs with the same frequency. A famous result of Borel is that almost every number is normal. We extend the definition of normal numbers to the case of Cantor series. The main result of this paper is that under some condition almost every number is normal in the new sense.

Communicated by Vladimír Baláž

1. Introduction

The concept of normal number was introduced by Borel. A number is called normal if in its base b expansion every block of digits occurs with the same frequency. In other words,

Definition 1. A real number $x \in (0,1)$ is called *simply normal to base* $b \geq 2$ if its base b expansion is $0.d_1(x)d_2(x)d_3(x)\ldots$ and

$$\lim_{N\to\infty} \frac{\#\{n \leq N \mid d_n(x) = A\}}{N} = \frac{1}{b} \quad \text{for every } A \in \{0,\ldots,b-1\}\,. \tag{1}$$

A number is called *normal to base* b if for every block of digits $A_1 \ldots A_L$, $L \geq 1$,

$$\lim_{N\to\infty} \frac{\#\{n \leq N-L \mid d_{n+1}(x) = A_1,\ldots,d_{n+L}(x) = A_L\}}{N} = \frac{1}{b^L}\,.$$

A number is called *absolutely normal* if it is normal to every base $b \geq 2$.

A famous result of Borel [1] is

Theorem 1. *Almost every real number is absolutely normal.*

2010 Mathematics Subject Classification: 11K16.
Keywords: Normal numbers, Cantor Expansion.

There are many ways to prove this theorem (see discussions in [3], [4] or [10]).

For the sake of brevity we will use Iverson bracket defined for a predicate P by

$$[P] = \begin{cases} 1 & \text{if } P \text{ is true,} \\ 0 & \text{if } P \text{ is false.} \end{cases}$$

Using this notation, one can rewrite condition (1) by

$$\lim_{N\to\infty} \frac{\sum\limits_{n=1}^{N} [d_n(x) = A]}{\sum\limits_{n=1}^{N} 1} = \frac{1}{b} \quad \text{for every } A \in \{0, \dots, b-1\}\,. \tag{2}$$

Notice that $\frac{1}{b}$ is the expected value of the function

$$\lim_{N\to\infty} \frac{\sum\limits_{n=1}^{N} [d_n = A]}{\sum\limits_{n=1}^{N} 1},$$

where $\{d_n\}_{n=1}^{\infty}$ is a sequence of independent uniformly distributed random variables $d_n \sim \mathcal{U}(\{0, \dots, b-1\})$.

In this paper, we first extend the definition of normal numbers to the case of Cantor series.

For a given sequence $\mathbf{b} = \{b_n\}_{n=1}^{\infty}$, $b_n \geq 2$, we will use the Cantor expansion of a real number x

$$x = \lfloor x \rfloor + \sum_{n=1}^{\infty} \frac{d_n(x)}{\prod\limits_{k=1}^{n} b_k}\,, \tag{3}$$

where $d_n(x) \in \{0, \dots, b_n - 1\}$. For almost every x, the digits $d_n(x)$ are uniquely determined by (3). Put $\mathbf{d}(x) := \{d_n(x)\}_{n=1}^{\infty}$.

Formula (2) does not seem to be good for Cantor expansions. For instance, if $b_n = 8$ and $b_p = b_q = b_r = 2$, then the events $d_n(x) = 0$ and $d_p(x) = d_q(x) = d_r(x) = 0$ occur with the same probability. Therefore it is good to add suitable weights depending on b_n. It seems that the most logical weights are $w_n = \log b_n$.

Denote by $h_{\mathbf{b}}(\mathbf{d}, A, N)$ the weighted "frequency" of a digit A in a finite sequence $\mathbf{d} = \{d_n\}_{n=1}^{N}$,

$$h_{\mathbf{b}}(\mathbf{d}, A, N) := \frac{\sum\limits_{n=1}^{N} [d_n = A] w_n}{\sum\limits_{n=1}^{N} w_n} = \frac{\log \prod\limits_{n=1}^{N} b_n^{[d_n = A]}}{\log \prod\limits_{n=1}^{N} b_n}\,.$$

Let $\mathbf{d}$ be a sequence of independent uniformly distributed random variables

$$d_n \sim \mathcal{U}(\{0,\ldots,b_n-1\})\,.$$

Then the expected value $R_{\mathbf{b}}(A,N)$ of the function $h_{\mathbf{b}}(\mathbf{d},A,N)$ is

$$\begin{aligned}
R_{\mathbf{b}}(A,N) &:= \mathbb{E}h_{\mathbf{b}}(\mathbf{d},A,N)\\
&= \frac{\mathbb{E}\sum\limits_{n=1}^{N}[d_n=A]\log b_n}{\sum\limits_{n=1}^{N}\log b_n} = \frac{\sum\limits_{n=1}^{N}\big(\mathbb{E}[d_n=A]\big)\log b_n}{\sum\limits_{n=1}^{N}\log b_n}\\
&= \frac{\sum\limits_{n=1}^{N}\frac{[A<b_n]}{b_n}\log b_n}{\sum\limits_{n=1}^{N}\log b_n} = \frac{\log\prod\limits_{n=1}^{N} b_n^{\frac{[A<b_n]}{b_n}}}{\log\prod\limits_{n=1}^{N} b_n}\,.
\end{aligned}$$

According to the previous text we introduce

Definition 2. A number (3) is called $\mathbf{b}$-*normal* if for every possible digit A

$$\liminf_{N\to\infty} h_{\mathbf{b}}(\mathbf{d}(x),A,N) = \liminf_{N\to\infty} R_{\mathbf{b}}(A,N)\,,$$

and

$$\limsup_{N\to\infty} h_{\mathbf{b}}(\mathbf{d}(x),A,N) = \limsup_{N\to\infty} R_{\mathbf{b}}(A,N)\,.$$

Our main result is

Theorem 2. *Let* $\mathbf{b}=\{b_n\}_{n=1}^{\infty}$ *be a sequence of integers with* $b_n\geq 2$ *for every* n *and with*

$$B := \limsup_{n\to\infty} b_n < \infty\,. \tag{4}$$

Then almost every real number (3) *is* $\mathbf{b}$-*normal.*

For results concerning unweighted frequencies of digits in Cantor expansions see, for instance, [2], [5], [6], [7], [8] or [9].

2. Preliminary results

The following lemma is Corollary 1 from [3].

Lemma 1. *Let* $\{b_n\}_{n=1}^{\infty}$ *be a sequence of integers with* $b_n \geq 2$ *and let for every* n *numbers* $\Theta_n(d) > 0$ *satisfy*

$$\sum_{d=0}^{b_n-1} \Theta_n(d) = b_n\,. \tag{5}$$

Then for almost every x *(with respect to Lebesgue measure) with Cantor expansion* (3) *the following product exists and is finite (possibly equal to zero)*

$$\prod_{n=1}^{\infty} \Theta_n(d_n(x)) < \infty\,. \tag{6}$$

Lemma 2. *The set of all* $\mathbf{b}$*-normal numbers is equal to*

$$\left\{ x \in \mathbb{R} \;\middle|\; \begin{array}{l} \liminf\limits_{N\to\infty} h_{\mathbf{b}}(\mathbf{d}(x), A, N) = \liminf\limits_{N\to\infty} R_{\mathbf{b}}(A, N) \\ \limsup\limits_{N\to\infty} h_{\mathbf{b}}(\mathbf{d}(x), A, N) = \limsup\limits_{N\to\infty} R_{\mathbf{b}}(A, N) \end{array} \textit{ for every } A < \limsup_{n\to\infty} b_n \right\}.$$

Denote the set mentioned in Lemma 2 by $\mathcal{Z}_{\mathbf{b}}$.

Proof. According to the definition it is sufficient to prove that

$$\lim_{N\to\infty} h_{\mathbf{b}}(\mathbf{d}(x), A, N) = \lim_{N\to\infty} R_{\mathbf{b}}(A, N) \quad \text{for all } A \geq \limsup_{n\to\infty} b_n.$$

For such A there are only finitely many n with $A < b_n$ and hence there are only finitely many n with $d_n(x) = A$. Thus, in the products in numerators in (7) below there are only finitely many terms different from 1. On the other hand, the denominators in (7) diverge. Hence,

$$\lim_{N\to\infty} \frac{\log \prod\limits_{n=1}^{N} b_n^{[d_n(x)=A]}}{\log \prod\limits_{n=1}^{N} b_n} = \lim_{N\to\infty} \frac{\log \prod\limits_{n=1}^{N} b_n^{\frac{[A<b_n]}{b_n}}}{\log \prod\limits_{n=1}^{N} b_n} = 0\,. \tag{7}$$

□

Lemma 3. *Consider the set*

$$\mathcal{J}_{\mathbf{b}} := \left\{ x \in \mathbb{R} \;\middle|\; \lim_{N\to\infty} \frac{\log \prod\limits_{n=1}^{N} b_n^{[d_n(x)=A]}}{\log \prod\limits_{n=1}^{N} b_n^{\frac{[A<b_n]}{b_n}}} = 1 \textit{ for all } A < \limsup_{n\to\infty} b_n \right\}.$$

Then $\mathcal{J}_{\mathbf{b}} \subseteq \mathcal{Z}_{\mathbf{b}}$.

P r o o f. For every $x \in \mathcal{J}_{\mathbf{b}}$ and every $A < \limsup_{n\to\infty} b_n$ we have

$$\lim_{N\to\infty} \frac{h_{\mathbf{b}}(\mathbf{d}(x), A, N)}{R_{\mathbf{b}}(A, N)} = \lim_{N\to\infty} \frac{\dfrac{\log \prod_{n=1}^{N} b_n^{[d_n(x)=A]}}{\log \prod_{n=1}^{N} b_n}}{\dfrac{\log \prod_{n=1}^{N} b_n^{\frac{[A<b_n]}{b_n}}}{\log \prod_{n=1}^{N} b_n}} = \lim_{N\to\infty} \frac{\log \prod_{n=1}^{N} b_n^{[d_n(x)=A]}}{\log \prod_{n=1}^{N} b_n^{\frac{[A<b_n]}{b_n}}} = 1\,.$$

Hence, $x \in \mathcal{Z}_{\mathbf{b}}$. $\square$

3. Proof of Theorem 2

Firstly, consider the function f defined for real numbers $b \geq 2$, $C \in \mathbb{R}$ and $V > 0$ by

$$f(b, C, V) := CV + \frac{b \log \frac{b}{b^C + b - 1}}{\log b}\,.$$

The function f is continuous (as a function of C) and has derivative

$$\frac{\partial f}{\partial C}(b, C, V) = V - \frac{b^{C+1}}{b^C + b - 1}\,.$$

It follows that

$$f(b, 0, V) = 0 \quad \text{and} \quad \frac{\partial f}{\partial C}(b, 0, V) = V - 1\,.$$

P r o o f. (of Theorem 2) Lemma 2 and Lemma 3 imply that if we put

$$P(A, V) := \left\{ x \in \mathbb{R} \;\middle|\; \liminf_{N\to\infty} \frac{\log \prod_{n=1}^{N} b_n^{[d_n(x)=A]}}{\log \prod_{n=1}^{N} b_n^{\frac{[A<b_n]}{b_n}}} < V \right\}, \tag{8}$$

$$Q(A, V) := \left\{ x \in \mathbb{R} \;\middle|\; \limsup_{N\to\infty} \frac{\log \prod_{n=1}^{N} b_n^{[d_n(x)=A]}}{\log \prod_{n=1}^{N} b_n^{\frac{[A<b_n]}{b_n}}} > V \right\}, \tag{9}$$

then the set of all non-normal numbers is a subset of

$$\mathbb{R}\setminus\mathcal{J}_{\mathbf{b}} = \bigcup_{A=0}^{B-1}\ \bigcup_{V\in(0,1)\cap\mathbb{Q}} P(A,V) \cup \bigcup_{A=0}^{B-1}\ \bigcup_{V\in(1,\infty)\cap\mathbb{Q}} Q(A,V)\,.$$

Thus, it is sufficient to prove that every set $P(A,V)$ (for $A<B$ and $V\in(0,1)$) and every set $Q(A,V)$ (for $A<B$ and $V\in(1,\infty)$) has zero Lebesgue measure.

1. We prove that for every $A<B$ and $V\in(0,1)$ the set $P(A,V)$ has zero Lebesgue measure. Since $f(b,0,V)=0$ and $\frac{\partial f}{\partial C}(b,0,V)<0$, we obtain that for every $b=2,\dots,B$ there exists a number $S(b,V)<0$ such that for every $C\in(S(b,V),0)$ we have $f(b,C,V)>0$. Set

$$E := \frac{1}{2}\max_{b=2,\dots,B} S(b,V) < 0$$

and for this number E put

$$H := \min_{b=2,\dots,B} f(b,E,V) > 0\,.$$

For every n and d we define numbers $\Theta_n(d)>0$ by

$$\Theta_n(d) := \begin{cases} 1 & \text{if } A\ge b_n\,,\\ \frac{b_n^{E+1}}{b_n^E+b_n-1} & \text{if } A<b_n \quad\text{and}\quad d=A\,,\\ \frac{b_n}{b_n^E+b_n-1} & \text{if } A<b_n \quad\text{and}\quad d\ne A\,. \end{cases} \tag{10}$$

One can easily check that condition (5) is satisfied. Thus, Lemma 1 implies (6) for almost every x.

Now consider two cases.

1a. First we assume that $A<\min\limits_{n\in\mathbb{N}} b_n$. Then $[A<b_n]=1$ for every n.

Let $x\in P(A,V)$. Equation (8) implies that there are infinitely many integers N such that

$$\frac{\log\prod\limits_{n=1}^{N} b_n^{[d_n(x)=A]}}{\log\prod\limits_{n=1}^{N} b_n^{\frac{[A<b_n]}{b_n}}} < V\,. \tag{11}$$

For every such number N we have

$$\begin{aligned}\prod_{n=1}^{N}\Theta_n(d_n(x)) &= \prod_{n=1}^{N}\left(\frac{b_n^{E+1}}{b_n^E+b_n-1}\right)^{[d_n(x)=A]}\left(\frac{b_n}{b_n^E+b_n-1}\right)^{1-[d_n(x)=A]}\\ &= \prod_{n=1}^{N} b_n^{E[d_n(x)=A]}\cdot\prod_{n=1}^{N}\frac{b_n}{b_n^E+b_n-1}\\ &> \prod_{n=1}^{N}\left(b_n^{\frac{1}{b_n}}\right)^{EV}\cdot\prod_{n=1}^{N}\frac{b_n}{b_n^E+b_n-1}\\ &= \prod_{n=1}^{N}\left(b_n^{\frac{1}{b_n}}\right)^{EV}\cdot\prod_{n=1}^{N}\left(b_n^{\frac{1}{b_n}}\right)^{\frac{b_n\log\frac{b_n}{b_n^E+b_n-1}}{\log b_n}}\\ &= \prod_{n=1}^{N}\left(b_n^{\frac{1}{b_n}}\right)^{f(b_n,E,V)} \geq \left(\prod_{n=1}^{N} b_n^{\frac{1}{b_n}}\right)^{H}.\end{aligned}$$

The above inequality is true for every $x\in P(A,V)$. However, for large N this inequality contradicts (6) since $H>0$ and $\prod_{n=1}^{\infty} b_n^{\frac{1}{b_n}}=\infty$. Hence, in case 1a the set $P(A,V)$ has zero Lebesgue measure.

1b. Now we assume that $A\geq\min_{n\in\mathbb{N}} b_n$. For every number N let M_N be the number of positive integers $n\leq N$ such that $A<b_n$. Note that the condition $A<B$ implies $\lim_{N\to\infty} M_N=\infty$. Let $\{\phi(m)\}_{m=1}^{\infty}$ be the increasing infinite sequence of all indices with $A<b_{\phi(m)}$. Then $[A<b_{\phi(m)}]=1$ for every m.

Let $x\in P(A,V)$. For every n with $A\geq b_n$ we have

$$\Theta_n(d_n(x)) = b_n^{[d_n(x)=A]} = b_n^{\frac{[A<b_n]}{b_n}} = 1\,.$$

Equation (8) implies that there are infinitely many integers N such that

$$\frac{\log\prod_{m=1}^{M_N} b_{\phi(m)}^{[d_{\phi(m)}(x)=A]}}{\log\prod_{m=1}^{M_N} b_{\phi(m)}^{\frac{[A<b_{\phi(m)}]}{b_{\phi(m)}}}} = \frac{\log\prod_{n=1}^{N} b_n^{[d_n(x)=A]}}{\log\prod_{n=1}^{N} b_n^{\frac{[A<b_n]}{b_n}}} < V\,. \tag{12}$$

Then (10) and (12) imply

$$\begin{aligned}
\prod_{n=1}^{N}\Theta_n(d_n(x)) = \prod_{m=1}^{M_N}\Theta_{\phi(m)}(d_{\phi(m)}(x))&\\
&= \prod_{m=1}^{M_N}\left(\frac{b_{\phi(m)}^{E+1}}{b_{\phi(m)}^{E}+b_{\phi(m)}-1}\right)^{[d_{\phi(m)}(x)=A]}\left(\frac{b_{\phi(m)}}{b_{\phi(m)}^{E}+b_{\phi(m)}-1}\right)^{1-[d_{\phi(m)}(x)=A]}\\
&= \prod_{m=1}^{M_N} b_{\phi(m)}^{E[d_{\phi(m)}(x)=A]}\cdot\prod_{m=1}^{M_N}\frac{b_{\phi(m)}}{b_{\phi(m)}^{E}+b_{\phi(m)}-1}\\
&> \prod_{m=1}^{M_N}\left(b_{\phi(m)}^{\frac{1}{b_{\phi(m)}}}\right)^{EV}\cdot\prod_{m=1}^{M_N}\frac{b_{\phi(m)}}{b_{\phi(m)}^{E}+b_{\phi(m)}-1}\\
&= \prod_{m=1}^{M_N}\left(b_{\phi(m)}^{\frac{1}{b_{\phi(m)}}}\right)^{EV}\cdot\prod_{m=1}^{M_N}\left(b_{\phi(m)}^{\frac{1}{b_{\phi(m)}}}\right)^{\frac{b_{\phi(m)}\log\frac{b_{\phi(m)}}{b_{\phi(m)}^{E}+b_{\phi(m)}-1}}{\log b_{\phi(m)}}}\\
&= \prod_{m=1}^{M_N}\left(b_{\phi(m)}^{\frac{1}{b_{\phi(m)}}}\right)^{f(b_{\phi(m)},E,V)} \geq \left(\prod_{m=1}^{M_N} b_{\phi(m)}^{\frac{1}{b_{\phi(m)}}}\right)^{H}.
\end{aligned}$$

The above inequality is true for every $x\in P(A,V)$. However, for large N this inequality contradicts (6) since $H>0$ and $\lim\limits_{N\to\infty}\prod_{m=1}^{M_N} b_{\phi(m)}^{\frac{1}{b_{\phi(m)}}}=\infty$ by (4) and $M_N\to\infty$. Hence, in case 1b the set $P(A,V)$ has zero Lebesgue measure.

2. Now we prove that for every $A<B$ and $V>1$ the set $Q(A,V)$ has zero Lebesgue measure. From the facts that $f(b,0,V)=0$ and $\frac{\partial f}{\partial C}(b,0,V)>0$ we obtain that for every b there exists a number $S(b,V)>0$ such that for every $C\in(0,S(b,V))$ we have $f(b,C,V)>0$. Set

$$E := \frac{1}{2}\min_{b=2,\dots,B} S(b,V) > 0$$

and for this number E put

$$H := \min_{b=2,\dots,B} f(b,E,V) > 0\,.$$

The rest of proof in case 2 reads exactly the same lines as in case 1 with the only exceptions that we write $Q(A,V)$ instead of $P(A,V)$ and that the inequalities in (11) and (12) are opposite. □

Acknowledgement. The authors were supported by grant no. P201/12/2351 of the Czech Science Foundation and by grant VEGA 1/1022/12.

REFERENCES

[1] BOREL, É. : *Les probabilités dénombrables et leurs applications arithmétiques*, Supplemento di rend. circ. Mat. Palermo **27** (1909), 247–271.

[2] ERDŐS, P.—RÉNYI, A.: *On Cantor's series with convergent* $\sum 1/q_n$, Ann. Univ. Sci. Budapest. Eötvös. Sect. Math. **2** (1959), 93–109.

[3] FILIP, F.—ŠUSTEK, J.: *An elementary proof that almost all real numbers are normal*, Acta Univ. Sapientiae, Mathematica **2** (2010), no. 1, 99–108.

[4] KHOSHNEVISAN, D.: *Normal numbers are normal*, Clay Mathematics Institute Annual Report (2006), p. 15 (continued on pp. 27–31).
`http://www2.maths.ox.ac.uk/cmi/library/annual_report/ar2006/06report_normalnumbers.pdf`

[5] MANCE, B.: *Construction of normal numbers with respect to the* Q*-Cantor series expansion for certain* Q, Acta Arith. **148** (2011), no. 2, 135–152.

[6] RÉNYI, A.: *On the distribution of the digits in Cantor's series*, Mat. Lapok **7** (1956), 77–100.

[7] ŠALÁT, T.: *Über die Cantorsche Reihen*, Czechoslovak Math. J. **18 (93)** (1968), 25–56.

[8] ŠALÁT, T.: *Einige metrische Ergebnisse in der Theorie der Cantorschen Reihen und Bairesche Kategorien von Mengen*, Studia Sci. Math. Hung. **6** (1971), 49–53.

[9] SCHWEIGER, F.: *Über den Satz von Borel-Rényi in der Theorie der Cantorschen Reihen*, Monatsh. Math. **74** (1970), 150–153.

[10] STRAUCH, O.—PORUBSKÝ, Š.: *Distribution of Sequences: a Sampler*. Peter Lang, 2005.

Received October 29, 2013
Accepted April 14, 2014

Ferdinánd Filip
Department of Mathematics and Informatics
J. Selye University
P.O.Box 54
945 01 Komárno
SLOVAKIA
E-mail: filip.ferdinand@ujs.sk

Jan Šustek
Department of Mathematics
Faculty of Science
University of Ostrava
30. dubna 22
701 03 Ostrava
CZECH REPUBLIC
E-mail: jan.sustek@osu.cz