ACTA ARITHMETICA
108.2 (2003)

Extremely non-normal continued fractions

by

L. Olsen (St. Andrews)

1. Introduction and statement of results. Let $\mathbb{P}$ denote the irrational numbers in the closed unit interval, i.e.

$$\mathbb{P} := [0,1] \setminus \mathbb{Q}.$$

For $x \in \mathbb{P}$, let

$$x = \cfrac{1}{a_1(x) + \cfrac{1}{a_2(x) + \cfrac{1}{a_3(x) + \dots}}} \tag{1.1}$$

where $a_n(x) \in \mathbb{N}$ for all n, denote the simple (infinite) continued fraction expansion of x. For a positive integer n and a digit $i \in \mathbb{N}$, we write

$$\Pi(x,i;n) = \frac{|\{1 \le j \le n \mid a_j(x) = i\}|}{n}$$

for the frequency of the digit i among the first n digits in the continued fraction expansion of x. A classical result due to Lévy [Lé] says that for Lebesgue almost all $x \in \mathbb{P}$ we have

$$\Pi(x,i;n) \to \frac{1}{\log 2} \log \frac{(i+1)^2}{i(i+2)} \tag{1.2}$$

for all $i \in \mathbb{N}$; the reader is referred to the textbook [Bi, p. 45] for a contemporary proof of this based on the ergodic theorem. In analogy with normal numbers (cf. [KN]), we will say that a number $x \in \mathbb{P}$ is *continued fraction normal* (c-f-normal) if it satisfies (1.2). Hence, using this terminology, Lévy's result says that Lebesgue almost all $x \in \mathbb{P}$ are c-f-normal.

2000 *Mathematics Subject Classification*: Primary 11K50; Secondary 11K16, 30B70, 28A80.

Key words and phrases: continued fractions, divergence points, Baire category, Hausdorff dimension, packing dimension.

In this paper we will prove that from a topological viewpoint, most numbers fail to be c-f-normal in a very spectacular way. We will show that (in the Baire sense) most numbers are as far away from being c-f-normal as possible. Similar results for sets of numbers whose N-adic expansion/Lüroth expansion deviates significantly from the N-adic expansion/Lüroth expansion of Lebesgue almost all numbers have been obtained by Olsen [Ol2] and Šalát [Ša1].

We first introduce some notation. For a positive integer n and a finite string $\mathbf{i} = i_1 \dots i_k \in \mathbb{N}^k$ of length k with entries $i_j \in \mathbb{N}$, we write

$$\Pi(x, \mathbf{i}; n) = \frac{|\{1 \le j \le n \mid a_j(x) = i_1, \dots, a_{j+k-1}(x) = i_k\}|}{n}$$

for the frequency of the string $\mathbf{i}$ among the first n digits in the simple continued fraction expansion of x, and let

$$\Pi_k(x; n) = (\Pi(x, \mathbf{i}; n))_{\mathbf{i} \in \mathbb{N}^k}$$

denote the vector of frequencies $\Pi(x, \mathbf{i}; n)$ of all strings $\mathbf{i} \in \mathbb{N}^k$ of length k. We define the subset Δ_k of ℓ^1 by

$$\Delta_k = \Big\{ (p_{\mathbf{i}})_{\mathbf{i} \in \mathbb{N}^k} \;\Big|\; p_{\mathbf{i}} \ge 0,\ \sum_{\mathbf{i}} p_{\mathbf{i}} = 1 \Big\},$$

i.e. Δ_k denotes the simplex of probability vectors indexed by strings $\mathbf{i} = i_1 \dots i_k$ of length k with entries $i_j \in \mathbb{N}$. We will always equip Δ_k with the 1-norm $\|\cdot\|_1$. The vector $\Pi_k(x; n)$ of frequencies of strings of length k among the first n digits in the simple continued fraction expansion of x clearly belongs to Δ_k. We will quantify the non-normality of x by considering the extent to which the sequence $(\Pi_k(x; n))_n$ fills up the simplex Δ_k. Of course, in general, it is not true that the sequence $(\Pi_k(x; n))_n$ fills up a substantial part of Δ_k for any x. For example, consider strings of length 3. By considering all possible ways a string of length 2, such as $37 \in \mathbb{N}^2$ (i.e. 37 represents the string of length 2 whose first digit equals 3 and whose second digit equals 7), can arise it is easily seen that

$$\Big| \sum_{i \in \mathbb{N}} \Pi(x, i37; n) - \sum_{i \in \mathbb{N}} \Pi(x, 37i; n) \Big| \le \frac{1}{n}$$

for *all* x. This implies that for *each* x, all but finitely many points in the sequence $(\Pi_3(x; n))_n$ will be very close to the subsimplex

$$\Delta_3 \cap \Big\{ (x_{\mathbf{i}})_{\mathbf{i} \in \mathbb{N}^3} \in \ell^1 \;\Big|\; \sum_{i \in \mathbb{N}} x_{i37} = \sum_{i \in \mathbb{N}} x_{37i} \Big\}. \tag{1.3}$$

Hence, in general the sequence $(\Pi_k(x; n))_n$ will not fill up a significant part of the simplex Δ_k, and the full simplex Δ_k is not the "correct" object to consider. Rather we need to consider the subsimplex defined by slicing Δ_k

by various planes corresponding to the subsimplex in (1.3). Motivated by this, we define the subsimplex S_k of shift invariant probability vectors in $\mathbb{R}^{\mathbb{N}^k}$ by

$$\mathsf{S}_k = \Big\{(p_{\mathbf{i}})_{\mathbf{i}\in\mathbb{N}^k} \,\Big|\, p_{\mathbf{i}} \geq 0, \sum_{\mathbf{i}} p_{\mathbf{i}} = 1, \sum_{i} p_{i\mathbf{i}} = \sum_{i} p_{\mathbf{i}i} \text{ for all } \mathbf{i} \in \mathbb{N}^{k-1}\Big\}. \tag{1.4}$$

Observe that $\Delta_1 = \mathsf{S}_1$. We will now prove that the subsimplex S_k is the "correct" object. Let $\mathsf{A}_k(x)$ denote the set of accumulation points of the sequence $(\Pi_k(x;n))_n$ with respect to $\|\cdot\|_1$, i.e.

$$\mathsf{A}_k(x) = \{\mathbf{p} \in \Delta_k \mid \text{there exists a subsequence } (\Pi_k(x;n_m))_m \text{ such that } \|\Pi_k(x;n_m) - \mathbf{p}\|_1 \to 0\}. \tag{1.5}$$

The next result says that the subsimplex S_k is, indeed, the "correct" simplex to consider: all accumulation points of $(\Pi_k(x;n))_n$ belong to S_k.

THEOREM 0. *Let $x \in [0,1]$. Then*

$$\mathsf{A}_k(x) \subseteq \mathsf{S}_k.$$

Proof. Let $\mathbf{p} = (p_{\mathbf{i}})_{\mathbf{i}\in\mathbb{N}^k}$ be an accumulation point of the sequence $(\Pi_k(x;n))_n$ with respect to $\|\cdot\|_1$. We can thus find a strictly increasing sequence $(n_m)_m$ of positive integers such that

$$\|\Pi_k(x;n_m) - \mathbf{p}\|_1 \to 0. \tag{1.6}$$

By considering all possible ways a string $\mathbf{i} \in \mathbb{N}^{k-1}$ of length $k-1$ can arise it follows that

$$\Big|\sum_{i\in\mathbb{N}} \Pi(x, i\mathbf{i}; n) - \sum_{i\in\mathbb{N}} \Pi(x, \mathbf{i}i; n)\Big| \leq 1/n. \tag{1.7}$$

It follows from (1.6) and (1.7) that if $\mathbf{i} \in \mathbb{N}^{k-1}$, then

$$\begin{aligned}\Big|\sum_{i\in\mathbb{N}} p_{i\mathbf{i}} - \sum_{i\in\mathbb{N}} p_{\mathbf{i}i}\Big| &\leq \Big|\sum_{i\in\mathbb{N}} p_{i\mathbf{i}} - \sum_{i\in\mathbb{N}} \Pi(x, i\mathbf{i}; n_m)\Big| \\ &\quad + \Big|\sum_{i\in\mathbb{N}} \Pi(x, i\mathbf{i}; n_m) - \sum_{i\in\mathbb{N}} \Pi(x, \mathbf{i}i; n_m)\Big| \\ &\quad + \Big|\sum_{i\in\mathbb{N}} \Pi(x, \mathbf{i}i; n_m) - \sum_{i\in\mathbb{N}} p_{\mathbf{i}i}\Big| \\ &\leq \|\Pi_k(x;n_m) - \mathbf{p}\|_1 + 1/n_m + \|\Pi_k(x;n_m) - \mathbf{p}\|_1 \to 0.\end{aligned}$$

This implies that $\sum_{i\in\mathbb{N}} p_{i\mathbf{i}} = \sum_{i\in\mathbb{N}} p_{\mathbf{i}i}$ for all $\mathbf{i} \in \mathbb{N}^{k-1}$. ∎

We will say that the number x is *extremely non-k-continued fraction normal* (extremely non-k-c-f-normal) if the set of accumulation points of the sequence $(\mathit{\Pi}_k(x;n))_n$ (with respect to $\|\cdot\|_1$) equals S_k, and we will denote the set of extremely non-k-c-f-normal numbers by $\mathbb{E}_k$, i.e.

$$\mathbb{E}_k = \{x \in \mathbb{P} \mid \mathsf{A}_k(x) = \mathsf{S}_k\}.$$

We will say that a number is *extremely non-continued fraction normal* (extremely non-c-f-normal) if it is extremely non-k-c-f-normal for all k. We let $\mathbb{E}$ denote the set of extremely non-c-f-normal numbers, i.e.

$$\mathbb{E} = \bigcap_k \mathbb{E}_k.$$

Hence, the numbers in $\mathbb{E}$ are as far away from being c-f-normal as possible. Our main result (Theorem 1 below) states, somewhat surprisingly, that the set $\mathbb{E}$ is extremely big from a topological viewpoint.

THEOREM 1. (1) *The set* $\mathbb{E}$ *is comeager in* $\mathbb{P}$, *i.e.* $\mathbb{P} \setminus \mathbb{E}$ *is of the first category* (*in* $\mathbb{P}$). *In particular*, $\mathbb{E}$ *is of the second category* (*in* $\mathbb{P}$).

(2) *The set* $\mathbb{E}$ *is comeager in* $[0,1]$, *i.e.* $[0,1] \setminus \mathbb{E}$ *is of the first category* (*in* $[0,1]$). *In particular*, $\mathbb{E}$ *is of the second category* (*in* $[0,1]$).

Theorem 1 shows that from a topological point of view, a typical number in $[0,1]$ is as far away from being c-f-normal as possible. The proof of Theorem 1 is given in Section 2.

Define the set $\mathbb{S}$ by

$$\mathbb{S} = \{x \in \mathbb{P} \mid \text{the sequence } (\mathit{\Pi}(x,i;n))_n \text{ is dense in } [0,1] \text{ for all } i \in \mathbb{N}\}.$$

Šalát [Ša2] proved that $\mathbb{S}$ is comeager. Since clearly $\mathbb{E} \subseteq \mathbb{E}_1 \subseteq \mathbb{S}$ it follows immediately from Theorem 1 that $\mathbb{S}$ is comeager.

As an immediate corollary to Theorem 1, we obtain the packing dimension $\operatorname{Dim}\mathbb{E}$ of $\mathbb{E}$; the reader is referred to [Fa] for the definition of Dim.

COROLLARY 2. $\operatorname{Dim}\mathbb{E} = 1$.

Proof. It follows from [Ed, Exercise (1.8.4)] that if E is a subset of $\mathbb{R}$ with $\operatorname{Dim} E < 1$, then E is of the first category. This and Theorem 1 imply that $\operatorname{Dim}\mathbb{E} = 1$. ∎

Let dim denote the Hausdorff dimension; the reader is referred to [Fa] for the definition. We have not considered the problem of computing the Hausdorff dimension of the set $\mathbb{E}$. However, it follows from [Ol1, Ol2] that the Hausdorff dimension of the set of numbers whose N-adic expansion satisfies a similar condition of extreme non-normality equals 0, and we therefore make the following conjecture.

CONJECTURE 3. $\dim\mathbb{E} = 0$. *In fact*, *for each positive integer* k, *we have* $\dim\mathbb{E}_k = 0$.

2. Proof of Theorem 1. We start by introducing some notation. Let $\mathbb{N}^* = \bigcup_n \mathbb{N}^n$, and define $\pi : \mathbb{N}^{\mathbb{N}} \to \mathbb{P}$ by

$$\pi(\omega) = \cfrac{1}{\omega_1 + \cfrac{1}{\omega_2 + \cfrac{1}{\omega_3 + \dots}}}$$

for $\omega = \omega_1\omega_2\ldots \in \mathbb{N}^{\mathbb{N}}$. If $\omega = \omega_1\omega_2\ldots \in \mathbb{N}^{\mathbb{N}}$ and m is a positive integer or if $\omega = \omega_1\ldots\omega_n \in \mathbb{N}^n$ and m is a positive integer with $m \le n$, then we will write $\omega|m = \omega_1\ldots\omega_m$. For $\omega \in \mathbb{N}^n$, we write $|\omega| = n$ and we define the cylinder $[\omega]$ generated by ω by

$$[\omega] = \{\sigma \in \mathbb{N}^{\mathbb{N}} \mid \sigma|n = \omega\}.$$

For $\mathbf{i} = i_1\ldots i_k \in \mathbb{N}^k$ and $\omega = \omega_1\ldots\omega_{n+k-1} \in \mathbb{N}^{n+k-1}$ write

$$\mathsf{P}(\omega,\mathbf{i}) = \frac{|\{1 \le i \le n \mid \omega_i = i_1,\ldots,\omega_{i+k-1} = i_k\}|}{n}$$

for the frequency of the string $\mathbf{i}$ among the digits of the string ω. Let

$$\mathsf{P}_k(\omega) = (\mathsf{P}(\omega,\mathbf{i}))_{\mathbf{i}\in\mathbb{N}^k}$$

denote the vector of all frequencies of strings $\mathbf{i}$ of length k among the digits of ω.

We now turn towards the proof of Theorem 1. Let

$$\mathsf{S}_k^* = \bigcup_N \Big\{ (p_{\mathbf{i}})_{\mathbf{i}\in\mathbb{N}^k} \;\Big|\; p_{\mathbf{i}} \ge 0, \sum_{\mathbf{i}} p_{\mathbf{i}} = 1, \sum_i p_{i\mathbf{i}} = \sum_i p_{\mathbf{i}i} \text{ for all } \mathbf{i} \in \mathbb{N}^{k-1},$$
$$p_{\mathbf{i}} = 0 \text{ for } \mathbf{i} \in \mathbb{N}^k \setminus \{1,\ldots,N\}^k \Big\}.$$

The set S_k^* is clearly a dense and separable subset of $(\mathsf{S}_k, \|\cdot\|_1)$. We can therefore find a sequence $(\mathbf{q}_{k,m})_m$ in S_k^* that is dense in $(\mathsf{S}_k, \|\cdot\|_1)$. For positive integers k and m write

$$\mathbb{E}_{k,m} = \{x \in \mathbb{P} \mid \mathbf{q}_{k,m} \text{ is an accumulation point of } (\Pi_k(x;n))_n\}.$$

We clearly have

$$\mathbb{E} = \bigcap_{k,m} \mathbb{E}_{k,m},$$

and it therefore suffices to prove that $\mathbb{E}_{k,m}$ is comeager for all k and m. Therefore, fix positive integers k and m. Since the set $\mathbb{P}$ of irrationals is a Baire space, in order to prove that $\mathbb{E}_{k,m}$ is a comeager subset of $\mathbb{P}$, it suffices to construct a set $E \subseteq \mathbb{P}$ satisfying the following three conditions:

(1) $E \subseteq \mathbb{E}_{k,m}$;
(2) E is dense in $\mathbb{P}$;
(3) E is a $\mathcal{G}_\delta$ set.

We will now proceed to construct a set E with the desired properties. Write

$$\mathbf{q}_{k,m} = \mathbf{q} = (q_{\mathbf{i}})_{\mathbf{i}\in\mathbb{N}^k}.$$

Since $\mathbf{q}$ belongs to S_k^*, there exists a positive integer N such that $q_{\mathbf{i}} = 0$ for $\mathbf{i} \in \mathbb{N}^k \setminus \{1,\dots,N\}^k$. Put

$$Z_n = \left\{ \omega \in \bigcup_{l\geq knN^k} \{1,\dots,N\}^l \;\middle|\; \|\mathsf{P}_k(\omega) - \mathbf{q}\|_1 \leq \frac{1}{n} \right\},$$
$$\widehat{Z}_n = \left\{ \omega \in \bigcup_{l\geq knN^k} \{1,\dots,N\}^l \;\middle|\; \|\mathsf{P}_k(\omega) - \mathbf{q}\|_1 \leq \frac{5}{n} \right\}.$$

Furthermore, for subsets $W, W_1, \dots, W_n$ of $\mathbb{N}^*$ and $\omega \in \mathbb{N}^*$ we will write

$$W_1 \dots W_n = \{\omega_1 \dots \omega_n \mid \omega_i \in W_i\},$$
$$\omega W = \{\omega\sigma \mid \sigma \in W\}, \qquad [W] = \{[\sigma] \mid \sigma \in W\}.$$

LEMMA 2.1. *Let n be a positive integer and $\omega \in \mathbb{N}^*$. Then there exists an integer $Q \geq n$ such that*

$$\omega \underbrace{Z_n \dots Z_n}_{Q\ times} \subseteq \widehat{Z}_n.$$

Proof. Let

$$\sigma = \omega\sigma_1 \dots \sigma_Q \in \omega \underbrace{Z_n \dots Z_n}_{Q \text{ times}} \qquad \text{with } \sigma_i \in Z_n.$$

Write $\omega = \omega_1 \dots \omega_s$ and $M = \max_i \omega_i$. For each $\mathbf{i} \in \mathbb{N}^k$ we clearly have

$$\frac{\sum_i |\sigma_i| \mathsf{P}(\sigma_i, \mathbf{i})}{|\sigma|} \leq \mathsf{P}(\sigma, \mathbf{i}) \leq \frac{|\omega| + \sum_i |\sigma_i| \mathsf{P}(\sigma_i, \mathbf{i}) + Qk}{|\sigma|}.$$

Since no $\mathbf{i} \in \mathbb{N}^k \setminus \{1,\dots,N\}^k$ is a substring of any σ_i (because $\sigma_i \in Z_n$), this implies that

$$\begin{aligned}
(2.1)\quad \|\mathsf{P}_k(\sigma) - \mathbf{q}\|_1 &\leq \left\| \mathsf{P}_k(\sigma) - \sum_i \frac{|\sigma_i|}{|\sigma|} \mathsf{P}_k(\sigma_i) \right\|_1 + \left\| \sum_i \frac{|\sigma_i|}{|\sigma|} \mathsf{P}_k(\sigma_i) - \mathbf{q} \right\|_1 \\
&= \sum_{\mathbf{i}\in\{1,\dots,N\}^k} \left| \mathsf{P}(\sigma, \mathbf{i}) - \sum_i \frac{|\sigma_i|}{|\sigma|} \mathsf{P}_k(\sigma_i) \right| \\
&\quad + \sum_{\mathbf{i}\in\mathbb{N}^k\setminus\{1,\dots,N\}^k} \left| \mathsf{P}(\sigma, \mathbf{i}) - \sum_i \frac{|\sigma_i|}{|\sigma|} \mathsf{P}_k(\sigma_i) \right| \\
&\quad + \left\| \sum_i \frac{|\sigma_i|}{|\sigma|} \mathsf{P}_k(\sigma_i) - \mathbf{q} \right\|_1
\end{aligned}$$

$$\le \sum_{\mathbf{i}\in\{1,\dots,N\}^k} \frac{|\omega|+Qk}{|\sigma|} + \sum_{\substack{\mathbf{i}\in\mathbb{N}^k\setminus\{1,\dots,N\}^k \\ \mathsf{P}(\omega,\mathbf{i})\neq 0}} \frac{|\omega|}{|\sigma|}$$
$$+ \Big\| \sum_i \frac{|\sigma_i|}{|\sigma|}\,\mathsf{P}_k(\sigma_i) - \mathbf{q}\Big\|_1$$
$$\le N^k \frac{|\omega|+Qk}{QnkN^k} + M^k \frac{|\omega|}{QnkN^k} + \Big\| \sum_i \frac{|\sigma_i|}{|\sigma|}\,\mathsf{P}_k(\sigma_i) - \mathbf{q}\Big\|_1$$
$$= \frac{1}{n} + c\,\frac{1}{Q} + \Big\| \sum_i \frac{|\sigma_i|}{|\sigma|}\,\mathsf{P}_k(\sigma_i) - \mathbf{q}\Big\|_1$$

where we have used the fact that $|\sigma| \ge \sum_{i=1}^{Q} |\sigma_i| \ge QnkN^k$ and written

$$c = \frac{|\omega|}{nk}\bigg(1 + \frac{M^k}{N^k}\bigg).$$

Also

$$\Big\| \sum_i \frac{|\sigma_i|}{|\sigma|}\,\mathsf{P}_k(\sigma_i) - \mathbf{q}_{k,m}\Big\|_1 \le \Big\| \sum_i \frac{|\sigma_i|}{|\sigma|}\,\mathsf{P}_k(\sigma_i) - \sum_i \frac{|\sigma_i|}{\sum_j |\sigma_j|}\,\mathsf{P}_k(\sigma_i)\Big\|_1 \tag{2.2}$$
$$+ \Big\| \sum_i \frac{|\sigma_i|}{\sum_j |\sigma_j|}\,\mathsf{P}_k(\sigma_i) - \mathbf{q}_{k,m}\Big\|_1$$
$$\le \sum_i |\sigma_i| \bigg| \frac{1}{|\sigma|} - \frac{1}{\sum_j |\sigma_j|} \bigg|$$
$$+ \sum_i \frac{|\sigma_i|}{\sum_j |\sigma_j|} \|\mathsf{P}_k(\sigma_i) - \mathbf{q}_{k,m}\|_1$$
$$\le \sum_i \frac{|\sigma_i|}{\sum_j |\sigma_j|} \frac{|\omega|}{|\sigma|} + \sum_i \frac{|\sigma_i|}{\sum_j |\sigma_j|} \frac{1}{n}$$
$$= \frac{|\omega|}{|\sigma|} + \frac{1}{n} \le \frac{|\omega|}{QnkN^k} + \frac{1}{n}.$$

It follows from (2.1) and (2.2) that

$$\|\mathsf{P}_k(\omega) - \mathbf{q}\|_1 \le \frac{1}{n} + c\,\frac{1}{Q} + \frac{|\omega|}{QnkN^k} + \frac{1}{n}.$$

Hence, by choosing Q large enough we can ensure that $Q \ge n$ and $\|\mathsf{P}_k(\omega) - \mathbf{q}\|_1 \le 5/n$. ■

LEMMA 2.2. *There exist functions* $u_n : \mathbb{N}^* \to \mathbb{N}^*$, $Q_n : \mathbb{N}^* \to \mathbb{N}$, *with the following properties*: *for all* $\omega \in \mathbb{N}^*$ *we have*

$$\pi([u_n(\omega)])^- \subseteq \pi([\omega])^\circ, \tag{2.3}$$

$$u_n(\omega)\ \underbrace{Z_n \dots Z_n}_{Q_n(\omega)\ times} \subseteq \widehat{Z}_n, \tag{2.4}$$

$$Q_n(\omega) \geq n. \tag{2.5}$$

In (2.3), *the closure and interior are with respect to the space* $\mathbb{P}$.

Proof. Let $\omega \in \mathbb{N}^*$. Now choose $\sigma \in \pi^{-1}(\pi([\omega 22]))$. Hence $\pi(\sigma) \in \pi([\omega 22])$, and we can thus choose a positive integer m such that $\pi([\sigma|m]) \subseteq \pi([\omega 2])^- \subseteq \pi([\omega])^\circ$. We now define $u_n(\omega)$ by $u_n(\omega) = \sigma|m$. Also, by Lemma 2.1 we can find an integer $Q_n(\omega)$ such that (2.4) and (2.5) are satisfied. ∎

Let $u_n : \mathbb{N}^* \to \mathbb{N}^*$ and $Q_n : \mathbb{N}^* \to \mathbb{N}$ be as in Lemma 2.2. Now we define $\Gamma_n \subseteq \mathbb{N}^*$ by

$$\Gamma_0 = \mathbb{N}^*,$$

$$\Gamma_1 = \bigcup_{\omega \in \Gamma_0} u_1(\omega)\ \underbrace{Z_1 \dots Z_1}_{Q_1(\omega)\ \text{times}},$$

$$\Gamma_2 = \bigcup_{\omega \in \Gamma_1} u_2(\omega)\ \underbrace{Z_2 \dots Z_2}_{Q_2(\omega)\ \text{times}}, \qquad \dots$$

and

$$E_n = \bigcup_{\omega \in \Gamma_n} \pi([\omega]).$$

Finally, let

$$E = \bigcap_n E_n.$$

We will now prove that E has the properties (1)–(3) listed before Lemma 2.1.

We first prove that Z_n is non-empty for all n. In order to prove this we will need the following result from [Ol1]. For $x \in [0,1]$, let

$$x = \sum_{n=1}^{\infty} \frac{\varepsilon_{N,n}(x)}{N^n},$$

where $\varepsilon_{N,n}(x) \in \{0,1,\dots,N-1\}$ for all n, denote the unique non-terminating N-adic expansion of x. For a positive integer n and a finite string $\mathbf{i} = i_1 \dots i_k \in \{0,1,\dots,N-1\}^k$ we write

$$\Lambda_N(x,\mathbf{i};n) = \frac{|\{1 \leq i \leq n \mid \varepsilon_{N,i}(x) = i_1, \dots, \varepsilon_{N,i+k-1}(x) = i_k\}|}{n}$$

for the frequency of the string $\mathbf{i}$ among the first n digits in the N-adic expansion of x, and let

$$\Lambda_N^k(x;n) = (\Lambda_N(x,\mathbf{i};n))_{\mathbf{i} \in \{0,1,\dots,N-1\}^k}$$

denote the vector of frequencies $\Lambda_N(x,\mathbf{i};n)$ of all strings $\mathbf{i}\in\{0,1,\dots,N-1\}^k$ of length k. Let

$$\Delta_N^k=\Big\{(p_{\mathbf{i}})_{\mathbf{i}\in\{0,1,\dots,N-1\}^k}\;\Big|\;p_{\mathbf{i}}\ge 0,\ \sum_{\mathbf{i}}p_{\mathbf{i}}=1\Big\},$$

and

$$\mathsf{S}_N^k=\Big\{(p_{\mathbf{i}})_{\mathbf{i}\in\{0,1,\dots,N-1\}^k}\;\Big|\;p_{\mathbf{i}}\ge 0,\ \sum_{\mathbf{i}}p_{\mathbf{i}}=1,\ \sum_i p_{i\mathbf{i}}=\sum_i p_{\mathbf{i}i}\ \text{for all }\mathbf{i}\in\{0,1,\dots,N-1\}^{k-1}\Big\},$$

i.e. Δ_N^k (S_N^k) denotes the simplex of (shift invariant) probability vectors indexed by strings $\mathbf{i}=i_1\dots i_k$ of length k with entries $i_j\in\{0,1,\dots,N-1\}$. Define $H_N^k:\Delta_N^k\to\mathbb{R}$ by

$$H_N^k(\mathbf{p})=-\frac{1}{\log N}\sum_{\mathbf{i}\in\{0,1,\dots,N-1\}^{k-1}}\sum_i p_{\mathbf{i}i}\log\frac{p_{\mathbf{i}i}}{\sum_j p_{\mathbf{i}j}}$$

for $\mathbf{p}=(p_{\mathbf{i}})_{\mathbf{i}\in\{0,1,\dots,N-1\}^k}$ (as usual, we put $0\log 0=0$). The following result is proved in [Ol1, Theorem 1].

THEOREM 2.3. *Let* $\mathbf{p}\in\Delta_N^k$.

(1) *If* $\mathbf{p}\notin\mathsf{S}_N^k$, *then*

$$\{x\in[0,1]\mid\lim_n\|\Lambda_N^k(x;n)-\mathbf{p}\|_1=0\}=\emptyset.$$

(2) *If* $\mathbf{p}\in\mathsf{S}_N^k$, *then*

$$\dim\{x\in[0,1]\mid\lim_n\|\Lambda_N^k(x;n)-\mathbf{p}\|_1=0\}=H_N^k(\mathbf{p}).$$

We can now prove that Z_n is non-empty.

LEMMA 2.4. $Z_n\ne\emptyset$ *for all* n.

Proof. Recall that $\mathbf{q}=(q_{\mathbf{i}})_{\mathbf{i}\in\mathbb{N}^k}$ where $q_{\mathbf{i}}=0$ for $\mathbf{i}\in\mathbb{N}^k\setminus\{1,\dots,N\}^k$. For $\mathbf{i}=i_1\dots i_k\in\{0,1,\dots,N-1\}^k$, write $\widetilde{\mathbf{i}}=\widetilde{i}_1\dots\widetilde{i}_k\in\{1,\dots,N\}^k$ where $\widetilde{i}_j=i_j+1$. Define $\widetilde{\mathbf{q}}$ by

$$\widetilde{\mathbf{q}}=(q_{\widetilde{\mathbf{i}}})_{\mathbf{i}\in\{0,1,\dots,N-1\}^k}.$$

Also, define $\widetilde{\mathsf{P}}(\omega)$ for $\omega\in\bigcup_n\mathbb{N}^{n+k-1}$ by

$$\widetilde{\mathsf{P}}(\omega)=(\mathsf{P}(\omega,\widetilde{\mathbf{i}}))_{\mathbf{i}\in\{0,1,\dots,N-1\}^k}.$$

Finally, let

$$X=\Big\{\omega\in\bigcup_{l\ge knN^k}\{1,\dots,N\}^l\;\Big|\;\|\widetilde{\mathsf{P}}(\omega)-\widetilde{\mathbf{q}}\|_1\le\frac{1}{n}\Big\}.$$

It is clear that $X \subseteq Z_n$. We now claim that $X \neq \emptyset$. Since $\mathbf{q} = (q_{\mathbf{i}})_{\mathbf{i} \in \mathbb{N}^k} \in \mathsf{S}_k$ and $q_{\mathbf{i}} = 0$ for $\mathbf{i} \in \mathbb{N}^k \setminus \{1, \ldots, N\}^k$, we conclude that $\widetilde{\mathbf{q}} \in \mathsf{S}_N^k$, and it therefore follows immediately from Theorem 2.3 that

$$\dim\{x \in [0,1] \mid \lim_n \|\Lambda_N^k(x;n) - \widetilde{\mathbf{q}}\|_1 = 0\} = H_N^k(\widetilde{\mathbf{q}}) > 0,$$

whence $\{x \in [0,1] \mid \lim_n \|\Lambda_N^k(x;n) - \widetilde{\mathbf{q}}\|_1 = 0\} \neq \emptyset$. We can thus choose $x \in [0,1]$ such that $\lim_n \|\Lambda_N^k(x;n) - \widetilde{\mathbf{q}}\|_1 = 0$. Put $\omega = \omega_1 \omega_2 \ldots \in \{1, \ldots, N\}^{\mathbb{N}}$ where $\omega_i = \varepsilon_{N,i}(x) + 1$. Then $\omega|m$ lies in X for m large enough. ■

Proposition 2.5. $E \subseteq \mathbb{E}_{k,m}$.

Proof. Let $x \in E$. We must now find a sequence $(n_l)_l$ of integers with $\lim_l n_l = \infty$ such that $\|\Pi_k(x;n_l) - \mathbf{q}\|_1 \to 0$. Since $x \in E = \bigcap_n E_n$, we conclude that for each positive integer n, we can find $\gamma_n \in \Gamma_n$ such that $x \in \pi([\gamma_n])$. We now define the sequence $(n_l)_l$ by $n_l = |\gamma_l| - (k-1)$ and claim that $n_l \to \infty$ and $\|\Pi_k(x;n_l) - \mathbf{q}\|_1 \to 0$. We first prove that $n_l \to \infty$. However, this is obvious since $\gamma_l \in \Gamma_l$. Next, we prove that $\|\Pi_k(x;n_l) - \mathbf{q}\|_1 \to 0$. It follows from (2.4) and the definition of Γ_n that $\Gamma_n \subseteq \widehat{Z}_n$ for all n. In particular, we conclude that $\gamma_l \in \Gamma_l \subseteq \widehat{Z}_l$. Using this and the fact that $\Pi_k(x;n_l) = \mathsf{P}_k(\gamma_l)$, we conclude that

$$\|\Pi_k(x;n_l) - \mathbf{q}\|_1 = \|\mathsf{P}_k(\gamma_l) - \mathbf{q}\|_1 \leq 5/l \to 0. \ ■$$

Proposition 2.6. *E is dense in $\mathbb{P}$.*

Proof. Let $x \in \mathbb{P}$ and $r > 0$. We must now find $t \in E \cap B(x,r)$. We first observe that there exists $\omega \in \mathbb{N}^*$ such that

$$\pi([\omega]) \subseteq B(x,r). \tag{2.6}$$

Next, since $Z_n \neq \emptyset$ for all n (cf. Lemma 2.4), we can choose strings $\omega_n \in \mathbb{N}^*$ inductively as follows. Let

$$\begin{aligned} &\omega_0 = \omega \in \Gamma_0, \\ &\omega_1 \in u_1(\omega_0)\ \underbrace{Z_1 \ldots Z_1}_{Q_1(\omega_0)\text{ times}} \subseteq \Gamma_1, \\ &\omega_2 \in u_2(\omega_1)\ \underbrace{Z_2 \ldots Z_2}_{Q_2(\omega_1)\text{ times}} \subseteq \Gamma_2, \qquad \ldots \end{aligned}$$

For each n we have

$$\pi([\omega_{n+1}])^- \subseteq \pi([u_n(\omega_n)])^- \subseteq \pi([\omega_n])^\circ \subseteq \pi([\omega_n]) \subseteq \pi([\omega_n])^-. \tag{2.7}$$

In (2.7), the closure and interior are with respect to the space $\mathbb{P}$. It follows from (2.7) that

$$\bigcap_n \pi([\omega_n])^- = \bigcap_n \pi([\omega_n]). \tag{2.8}$$

It also follows from (2.7) that $(\pi([\omega_n])^-)_n$ is a decreasing sequence of non-empty compact subsets of $[0,1]$, and the intersection $\bigcap_n \pi([\omega_n])^-$ is therefore non-empty. Now pick any $t \in \bigcap_n \pi([\omega_n])^-$. We claim that $t \in E \cap B(x,r)$. We first prove that $t \in E$. Using (2.8) we see that $t \in \bigcap_n \pi([\omega_n])^- = \bigcap_n \pi([\omega_n]) \subseteq \bigcap_n E_n = E$. Next we prove that $t \in B(x,r)$. We clearly have (using (2.6)) $t \in \pi([\omega_0]) = \pi([\omega]) \subseteq B(x,r)$. ∎

PROPOSITION 2.7. *E is a $\mathcal{G}_\delta$ set.*

Proof. For a positive integer n we define the set G_n by

$$G_n = \bigcup_{\omega \in \Gamma_n} \pi([\omega])^\circ$$

where there interior is with respect to the space $\mathbb{P}$. The set G_n is clearly open (in $\mathbb{P}$). We now have

$$(2.9)\qquad E_{n+1} = \bigcup_{\omega \in \Gamma_{n+1}} \pi([\omega]) \subseteq \bigcup_{\sigma \in \Gamma_n} \pi([u_{n+1}(\sigma)\, \underbrace{Z_{n+1} \ldots Z_{n+1}}_{Q_{n+1}(\sigma)\text{ times}}]) \subseteq \bigcup_{\sigma \in \Gamma_n} \pi([u_{n+1}(\sigma)]) \subseteq \bigcup_{\sigma \in \Gamma_n} \pi([\sigma])^\circ = G_n,$$

and

$$(2.10)\qquad G_n = \bigcup_{\omega \in \Gamma_n} \pi([\omega])^\circ \subseteq \bigcup_{\omega \in \Gamma_n} \pi([\omega]) = E_n.$$

It follows immediately from (2.9) and (2.10) that $E = \bigcap_n E_n = \bigcap_n G_n$. Since each G_n is open, this shows that E is $\mathcal{G}_\delta$. ∎

Proof of Theorem 1. (1) Since $\mathbb{P}$ is a Baire space, it follows immediately from Propositions 2.5–2.7 that the set $\mathbb{E}$ is comeager.

(2) This statement easily follows from Theorem 1(1). ∎

References

[Bi] P. Billingsley, *Ergodic Theory and Information*, Wiley, 1965.

[Ed] G. Edgar, *Integral, Probability, and Fractal Measures*, Springer, New York, 1998.

[Fa] K. J. Falconer, *Fractal Geometry*, Wiley, 1990.

[KN] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley-Interscience, New York, 1974.

[Lé] P. Lévy, *Théorie de l'addition des variables aléatoires*, Gauthier-Villars, Paris, 1937.

[Ol1] L. Olsen, *Applications of multifractal divergence points to sets of numbers defined by their N-adic expansion*, Math. Proc. Cambridge Philos. Soc., to appear.

[Ol2] —, *Extremely non-normal numbers*, preprint, 2002.

[Ša1] T. Šalát, *Zur metrischen Theorie der Lürothschen Entwicklungen der reellen Zahlen*, Czechoslovak Math. J. 93 (1968), 489–522.

[Ša2] —, *Bemerkung zu einem Satz von P. Lévy in der metrischen Theorie der Kettenbrüche*, Math. Nachr. 41 (1969), 91–94.

Department of Mathematics
University of St. Andrews
St. Andrews, Fife KY16 9SS, Scotland
E-mail: lo@st-and.ac.uk

Received on 29.7.2002 (4335)