# On the expansions of a real number to several integer bases

Yann Bugeaud

**Abstract.** *Only very little is known on the expansions of a real number to several integer bases. We establish various results showing that the expansions of a real number in two multiplicatively independent bases cannot both be simple, in a suitable sense. We also construct explicitly a real number $\xi$ which is rich to all integer bases, that is, with the property that, for every integer $b \geq 2$, every finite block of letters on the alphabet $\{0, 1, \ldots, b-1\}$ occurs in the b-ary expansion of $\xi$.*

## 1. Introduction

Throughout this paper, $\lfloor x \rfloor$ denotes the greatest integer less than or equal to $x$ and $\lceil x \rceil$ denotes the smallest integer greater than or equal to $x$. Let $b \geq 2$ be an integer. For a non-zero real number $\xi$, write

$$\xi = \lfloor \xi \rfloor + \sum_{k \geq 1} \frac{a_k}{b^k} = \lfloor \xi \rfloor + 0.a_1 a_2 \ldots,$$

where each digit $a_k$ is an integer from $\{0, 1, \ldots, b-1\}$ and an infinity of the $a_k$ are not equal to $b-1$. The sequence $(a_k)_{k \geq 1}$ is uniquely determined by the fractional part of $\xi$. With a slight abuse of notation, we call it the b-ary expansion of $\xi$.

We begin by introducing several notions which are commonly used to measure the complexity of the b-ary expansion of a real number. For a positive integer $n$, let $p(n, \xi, b)$ denote the total number of distinct blocks of $n$ digits in the b-ary expansion of $\xi$, that is,

$$p(n, \xi, b) = \text{Card}\{(a_{j+1}, \ldots, a_{j+n}) : j \geq 0\}.$$

Obviously, we have $1 \leq p(n, \xi, b) \leq b^n$, and both inequalities are sharp. If $\xi$ is rational, then its b-ary expansion is ultimately periodic, thus, the numbers $p(n, \xi, b)$, $n \geq 1$, are uniformly bounded by a constant depending only on

1

$\xi$ and $b$. If $\xi$ is irrational, then, by a classical result of Morse and Hedlund [18, 19], we know that $p(n, \xi, b) \geq n + 1$ for every positive integer $n$, and this inequality are sharp. Furthermore, since

$$p(n + n', \xi, b) \leq p(n, \xi, b) \cdot p(n', \xi, b) \qquad (1.1)$$

holds for every positive integers $n, n'$, the sequence $(\log p(n, \xi, b))_{n \geq 1}$ is sub-additive, thus the sequence $((\log p(n, \xi, b))/n)_{n \geq 1}$ converges.

**Definition.** *Let $b \geq 2$ be an integer. A real number $\xi$ is rich to base $b$ if $p(n, \xi, b) = b^n$ for every $n \geq 1$. The entropy of $\xi$ to base $b$, denoted by $E(\xi, b)$, is defined by*

$$E(\xi, b) = \lim_{n \to +\infty} \frac{\log p(n, \xi, b)}{n}.$$

The notion of richness (also called disjunctiveness) was considered for instance by Compton [11] and Hertling [15]. Note that the terminology 'rich' is used since 2008 in combinatorics on words with a different meaning; see [12].

It follows from (1.1) that $E(\xi, b) = \log b$ if, and only if, $p(n, \xi, b) = b^n$ for every $n \geq 1$, that is, if, and only if, $\xi$ is rich to base $b$. Observe that the entropy of a real number to base $b$ belongs to $[0, \log b]$. Furthermore, the set of real numbers $\xi$ such that $E(\xi, b) = 0$ for some $b \geq 2$ has zero Hausdorff dimension; see e.g. [17].

The first result of the present paper asserts that there are irrational real numbers that are not rich to every integer base.

A real number $\xi$ is simply normal to base $b$ if every digit from $\{0, 1, \ldots, b-1\}$ occurs in its $b$-ary expansion with the same frequency $1/b$. It is normal to base $b$ if it is simply normal to each base $b^m$ with $m \geq 1$. A well-known theorem of Borel [6] asserts that almost every real number is normal to every integer base. Furthermore, any real number normal to base $b$ is clearly rich to base $b$.

Two positive integers $x$ and $y$ are called *multiplicatively independent* if the only pair of integers $(m, n)$ such that $x^m y^n = 1$ is the pair $(0, 0)$. Answering a question of Steinhaus, Cassels [10] (for $b = 3$) and Schmidt [21] (for every $b$), independently, established the following result.

**Theorem CS.** *Let $b$ and $b'$ be two integers $\geq 2$. Normality to base $b$ implies normality to base $b'$ if, and only if, $b$ and $b'$ are multiplicatively dependent.*

The conclusion of Theorem CS holds if normality is replaced by richness, as shown by El-Zanati and Transue [14] (actually, the 'only if' sense of the next result was previously established in [21]).

2

**Theorem EZT.** *Let $b$ and $b'$ be two integers $\geq 2$. Richness to base $b$ implies richness to base $b'$ if, and only if, $b$ and $b'$ are multiplicatively dependent.*

Starting with the Champernowne number

$$0.1234567891011121314\ldots,$$

several explicit examples of real numbers normal to a given base have been constructed; see [9] for references. However, it remains an open question to construct explicitly a real number which is normal to two multiplicatively independent bases; see, however, [25, 22] and [4, 5] for computable constructions of real numbers normal to all integer bases. The situation is much better understood when normality is replaced by the weaker property of richness. Namely, Hertling [15] has given explicit examples of real numbers rich to infinitely many pairwise multiplicatively independent bases, and not rich to some given base. Our second result, inspired by his work, gives an explicit construction of a real number rich to every integer base.

Apart from the results stated above, very few is known on the expansion of a given irrational number in two multiplicatively independent bases. We address the following question.

**Problem 1.** *Are there irrational real numbers having a 'simple' expansion in two multiplicatively independent bases?*

We have to explain what we mean by 'simple', since there are several ways to measure the complexity of $\xi$. If we use the notion of entropy, then Problem 1 can be rephrased as follows.

**Problem 2.** *Are there irrational real numbers having zero entropy in two multiplicatively independent bases?*

We are unable to answer Problem 2. However, Theorem 1 below gives some (very) partial information in this direction.

A different point of view for 'simplicity' was taken in [3]. For an integer $b \geq 2$, a real number $\xi$, and a positive integer $n$, we count the number of non-zero digits of $\xi$ among its $n$ first $b$-ary digits by setting

$$\mathcal{NZ}(n, \xi, b) = \mathrm{Card}\{1 \leq k \leq n : a_k \neq 0\}.$$

An irrational real number $\xi$ could be considered 'simple' to base $b$ if $\mathcal{NZ}(n, \xi, b)$ is small for every large integer $n$ (which means that $\xi$ has only few non-zero digits).

A third point of view, addressed in [8], aims to estimate the asymptotic behaviour of the number of digit changes in the $b$-ary expansion of $\xi$. Following [8], we define the function $\mathcal{DC}$, 'number of digit changes', by

$$\mathcal{DC}(n, \xi, b) = \mathrm{Card}\{1 \leq k \leq n : a_k \neq a_{k+1}\},$$

for any positive integer $n$. An irrational real number could be considered 'simple' to base $b$ if it has only few digit changes in its $b$-ary expansion.

Since
$$\mathcal{DC}(n,\xi,b) \leq 2\mathcal{NZ}(n,\xi,b) + 1, \qquad (1.2)$$
a lower bound for $\mathcal{DC}(n,\xi,b)$ implies a lower bound for $\mathcal{NZ}(n,\xi,b)$. However, the converse does not hold.

Our third result asserts that, if $\xi$ is non-zero and the bases $b$ and $b'$ are coprime, then, for $n$ sufficiently large, $\mathcal{DC}(n,\xi,b)$ and $\mathcal{DC}(n,\xi,b')$ cannot be simultaneously very small.

The present paper is organized as follows. Our results are stated in Section 2 and proved in Sections 3 to 6.

## 2. Main results

We begin by establishing the existence of real numbers whose entropy to base $b$ is strictly less than $\log b$, for every $b \geq 2$. This is a straightforward consequence of a result of Akhunzhanov [1, 2]. Combined with arguments from [7], we obtain a slightly stronger statement.

**Theorem 1.** *Let $\varepsilon$ be a positive real number and $b_0 \geq 2$ be an integer. There exist a positive real number $c$, depending only on $\varepsilon$ and $b_0$, and uncountably many real numbers $\xi$ such that*

$$E(\xi, b_0) < \varepsilon$$

*and*

$$E(\xi, b) \leq \log b - \frac{1}{b^{cb(\log b)}}, \qquad \text{for every } b \geq 2.$$

The main ingredient in the proof of Theorem 1 is the notion of $(\alpha, \beta)$-game, introduced by Schmidt [23].

Theorem 1 implies the existence of real numbers which are not rich to every integer base. In the opposite direction, we give below an explicit construction of a real number which is rich to every integer base.

Let $b$ be an odd prime number. For an integer $j \geq 2$, let $W_j$ be an integer whose representation in base $j$ comprises each digit from $\{0, 1, \ldots, j-1\}$ exactly once and whose representation in base $b$ finishes by the digit 1. Clearly, we have
$$W_j \leq j^j.$$
Let $(a_j)_{j \geq 2}$ be a sequence of integers $\geq 2$ such that $a_j = 2$ for every even index $j$ and $a_{2j+1} = bj$ for $j \geq 1$.

For $j \geq 2$ not a power of $b$, let $(p_{k,bj}/q_{k,bj})_{k \geq 1}$ denote the sequence of convergents to the irrational number $(\log bj)/(\log j)$.

Put $c_2 = 1$ and $n_2 = 2$. Let $j$ be a positive integer such that $c_2, \ldots, c_{2j}$ and $n_2, \ldots, n_{2j}$ have been constructed. To shorten the notation, set $\ell_j := \lceil bj \log(bj) \rceil$ and $m_j = \lceil (\log bj)/(\log j) \rceil$. Note that the $b$-ary expansion of $W_{bj}$ has less than $\ell_j$ digits. If $j$ is not a power of $b$, then set

$$n_{2j+1} = p_{6(\ell_j + m_j \log n_{2j})+1, bj}, \qquad c_{2j+1} = W_{bj}.$$

If $j$ is a power of $b$, then set

$$n_{2j+1} = n_{2j}^2 + \ell_j + 2, \qquad c_{2j+1} = W_{bj}.$$

In every case, set

$$n_{2j+2} = \lceil b^{n_{2j+1}} \log j \rceil, \qquad c_{2j+2} = 1.$$

We are ready to state our second result.

**Theorem 2.** *Let $b$ be an odd prime number. With the sequences $(c_j)_{j \geq 2}$ and $(n_j)_{j \geq 2}$ defined above, the real number*

$$\xi = \sum_{j \geq 2} \frac{c_j}{b^{n_j}}$$

*is rich to every integer base.*

Now, we turn our attention to Problem 1 and we show that a real number cannot have too few digit changes simultaneously in two coprime bases.

**Theorem 3.** *Let $b$ and $b'$ be multiplicatively independent positive integers. Let $\xi$ be an irrational real number. If $b$ and $b'$ are coprime, then there exist an integer $n_0$ and a positive real number $c$ such that*

$$\mathcal{DC}(n, \xi, b) + \mathcal{DC}(n, \xi, b') \geq c \log n, \quad \text{for } n \geq n_0, \tag{2.1}$$

*and*

$$\mathcal{NZ}(n, \xi, b) + \mathcal{NZ}(n, \xi, b') \geq c \log n, \quad \text{for } n \geq n_0. \tag{2.2}$$

*If $b$ and $b'$ are not coprime, then there are a positive real number $c$ and arbitrarily large integers $n$ such that*

$$\mathcal{DC}(n, \xi, b) + \mathcal{DC}(n, \xi, b') \geq c \log n \tag{2.3}$$

*and*

$$\mathcal{NZ}(n, \xi, b) + \mathcal{NZ}(n, \xi, b') \geq c \log n. \tag{2.4}$$

The assumption that the bases $b$ and $b'$ are multiplicatively independent is necessary. Indeed, for a very rapidly increasing sequence $(n_k)_{k \geq 1}$ of positive integers, and $b \geq 2$, $m$ and $\ell$ positive integers, the number

$$\sum_{k \geq 1} \frac{1}{b^{m\ell n_k}} = \sum_{k \geq 1} \frac{1}{(b^\ell)^{mn_k}} = \sum_{k \geq 1} \frac{1}{(b^m)^{\ell n_k}}$$

has only few non-zero digits in base $b^\ell$ and in base $b^m$.

It should be pointed out that, if $b$ and $b'$ are not coprime, then there are irrational real numbers $\xi$ for which (2.3) and (2.4) do not hold for every large integer $n$. Indeed, take for example $b = 6$, $b' = 10$ and

$$\xi = \sum_{k \geq 1} \frac{1}{2^{n_k}},$$

for a very rapidly increasing sequence $(n_k)_{k \geq 1}$ of positive integers. Then, there exist integers $a_{k,h}$ in $\{0, 1, \ldots, 5\}$ and $b_{k,h}$ in $\{0, 1, \ldots, 9\}$ such that

$$\xi = \sum_{k \geq 1} \frac{3^{n_k}}{6^{n_k}} = \sum_{k \geq 1} \sum_{h=0}^{\lceil n_k (\log 3)/(\log 6) \rceil} \frac{a_{k,h}}{6^{n_k - h}}$$

$$= \sum_{k \geq 1} \frac{5^{n_k}}{10^{n_k}} = \sum_{k \geq 1} \sum_{h=0}^{\lceil n_k (\log 5)/(\log 10) \rceil} \frac{b_{k,h}}{10^{n_k - h}}.$$

Consequently, if $n_k > 4n_{k-1}$, then among the first $n_k/4$ digits of $\xi$ written in base 6 or in base 10, there are at most $n_{k-1}$ non-zero digits, and

$$\mathcal{NZ}(n_k/4, \xi, 6) + \mathcal{NZ}(n_k/4, \xi, 10) \leq n_{k-1},$$

$$\mathcal{DC}(n_k/4, \xi, 6) + \mathcal{DC}(n_k/4, \xi, 10) \leq n_{k-1}.$$

Since $n_k$ can be chosen arbitrarily large, this proves our claim.

In particular, Theorem 3 asserts that an irrational real number cannot have very few non-zero digits simultaneously in base 2 and in base 3. This gives a very partial answer to Problem 1.

A related question has been investigated in 1973 by Senge and Straus [24], who proved that the number of integers, the sum of whose digits in each of the bases $b$ and $b'$ lies below a fixed bound, is finite if, and only if, $b$ and $b'$ are multiplicatively independent. Their proof rests on the Thue–Siegel–Roth theorem and, hence, is ineffective. Using Baker's theory of linear forms in logarithms, Stewart [26] succeeded in establishing an effective version of Senge and Straus' theorem. He showed that if $b$ and $b'$ are multiplicatively

independent, then, for every $c \geq 1$, each integer $m$ whose sum of digits in base $b$ as well as in base $b'$ is bounded by $c$ satisfies

$$m < \exp\exp\{\kappa c \log(3c)\}, \tag{2.5}$$

where $\kappa$ is a positive constant which is effectively computable in terms of $b$ and $b'$ only. A similar result holds for the number of digit changes. Note that, when $b$ and $b'$ are not coprime, then an elementary argument based on the consideration of $p$-adic valuations, for a prime number $p$ dividing $b$ and $b'$, gives a slight strengthening of (2.5).

We highlight a much weaker question than Problem 2.

**Problem 3.** *Let $b$ and $b'$ be multiplicatively independent positive integers. Let $\xi$ be an irrational real number. Is it true that*

$$p(n, \xi, b) + p(n, \xi, b') \geq 2n + 3, \quad (n \geq 1)?$$

A positive answer to Problem 3 would mean that no real number can have a Sturmian expansion in two multiplicatively independent bases.

Our last result is a metric statement concerning simple normality to distinct bases. Apparently, Hertling [16] was the first to establish the correct analogue of Theorem CS when normality is replaced by simple normality.

**Theorem H.** *Simple normality to base $b$ implies simple normality to base $b'$ if and only if $b$ is a power of $b'$.*

More precisely, Hertling proved that, for integers $b$ and $b'$ both $\geq 2$ such that $b$ is not a power of $b'$, the set of real numbers which are simply normal to base $b$, but not simply normal to base $b'$, is uncountable. Our last theorem is a refinement of his result.

**Theorem 4.** *Let $a$ and $b$ be integers $\geq 2$ such that $a$ is not an integer power of $b$. Then, the set of real numbers which are simply normal to base $a$ but not simply normal to base $b$ has full Hausdorff dimension.*

When $a$ and $b$ are multiplicatively independent integers, Theorem 4 has been proved by Nagasaka [20]. Thus, we are only concerned with multiplicatively dependent integers $a$ and $b$. Our proof differs from that of the weaker result established in [16]. We make use of a classical theorem of Eggleston [13], see Theorem E in Section 6.

## 3. Proof of Theorem 1

Before giving an outline of the proof of Theorem 1, we claim that, for a real number $\xi$, an integer $b \geq 2$, and positive integers $n$ and $t$, we have

$$p(tn, \xi, b) \leq tb^t p(n, \xi, b^t). \tag{3.1}$$

To see this, observe that any block of $tn$ consecutive digits of $\xi$ to base $b$ is composed of $t_1$ digits of $\xi$ to base $b$ followed by $n - 1$ digits of $\xi$ to base $b^t$ and by $t - t_1$ digits of $\xi$ to base $b$, for some integer $t_1$ with $0 \le t_1 \le t - 1$.

The key ingredient of the proof of Theorem 1 is the notion of Schmidt's game, introduced in [23]. By means of a suitable modification of the original procedure, Akhunzhanov [1, 2] proved that there are real numbers $\xi$ such that

$$||b^n \xi|| \ge \exp\{-5000b(\log b)^2\}, \text{ for every } n \ge 1 \text{ and every } b \ge 2, \quad (3.2)$$

where $|| \cdot ||$ denotes the distance to the nearest integer. Let $\xi$ be with this property and $b \ge 2$ be an integer. Set

$$t = \lceil 5000b(\log b) \rceil.$$

Property (3.2) implies that $\xi$ has (at least) one missing digit in its expansion to base $b^t$, yielding that

$$p(n, \xi, b^t) \le (b^t - 1)^n, \quad \text{for } n \ge 1.$$

We deduce from (3.1) that

$$p(tn, \xi, b) \le t b^t (b^t - 1)^n,$$

and, by taking the logarithm, dividing by $tn$ and letting $n$ tend to infinity, it follows that

$$E(\xi, b) \le \frac{\log(b^t - 1)}{t} \le \log b - \frac{1}{tb^t}. \quad (3.3)$$

We can do slightly better by combining [1, 2] with ideas from [7], where the Schmidt game is played on a Cantor set. Let $\varepsilon$ be a positive real number. Let $b_0 \ge 2$ be an integer. Let $t$ be a positive integer, and consider the Cantor set $\mathcal{C}_t$ composed of the numbers $\xi$ in $[0, 1]$ having only the digits $0$ and $b_0^t - 1$ in base $b_0^t$. Then, clearly,

$$p(n, \xi, b_0^t) \le 2^n.$$

We deduce from (3.1) that

$$p(tn, \xi, b_0) \le t b_0^t 2^n,$$

and

$$E(\xi, b_0) \le \frac{\log 2}{t} < \varepsilon, \quad (3.4)$$

if $t$ is chosen large enough.

Instead of playing the Schmidt game on a real interval of length 2 as in [1, 2], we play it on the set $\mathcal{C}_t$. The analogue of Lemma 4 from [2] shows

that a countable intersection of winning sets on $\mathcal{C}_t$ is also winning on $\mathcal{C}_t$. There is no additional difficulty, therefore we omit the details of the proof. Proceeding in this way, we establish that there are uncountably many $\xi$ satisfying (3.4) and (3.3) for every $b \geq 2$. This proves the theorem. $\qquad \square$

## 4. Proof of Theorem 2

The proof of Theorem 2 depends on a series of lemmas, some of which having been already established by Hertling [15].

**Lemma 1.** *Let $b$ be an odd prime number. Let $a \geq 2$ be an integer coprime with $b$. Let $\mu_b$ be the order of $a$ modulo $b$ and $\ell_b$ be the greatest integer $\ell$ such that $b^\ell$ divides $a^{\mu_b} - 1$. For any positive integer $n$ exceeding $\ell_b$ and any integer $c$ prime to $b$, every block on $\{0, 1, \ldots, a-1\}$ of length at most $(n - \ell_b) \log_a b$ occurs in the $a$-ary expansion of $c/b^n$.*

Although this lemma can be found in [15], we give a proof.
*Proof.* The length of the periodic part of the $a$-ary expansion of $c/b^n$ is at most equal to the order of $a$ modulo $b^n$, thus it is less than $b^n$. Let $0.a_1 a_2 \ldots$ be the $a$-ary expansion of the fractional part of $c/b^n$. Let $w$ be a word of length $L := \lfloor (n - \ell_b) \log_a b \rfloor$ on $\{0, 1, \ldots, a-1\}$ and set

$$I_w := \{\xi \in [0, 1) : \text{the } a\text{-ary expansion of } \xi \text{ begins with } w\}.$$

For every $m \geq 0$, we have $a_{m+1} \ldots a_{m+L} = w$ if, and only if, $ca^m$ modulo $b^n$ lies in the interval $b^n \cdot I_w \subset [0, b^n)$. Since the length of $b^n \cdot I_w \subset [0, b^n)$ is at least equal to

$$b^n \cdot a^{-L} \geq b^n \cdot b^{-n+\ell_b} = b^{\ell_b},$$

there exists an integer in $\{0, 1, \ldots, b^n - 1\}$ which is congruent to $c$ modulo $b^{\ell_b}$ and belongs to $b^n \cdot I_w$. On the other hand, the subgroup of $(\mathbf{Z}/b^n\mathbf{Z})^*$ generated by $a^{\mu_b}$ is equal to the set of classes $x$ in $(\mathbf{Z}/b^n\mathbf{Z})^*$ which are congruent to 1 modulo $b^{\ell_b}$. Since $b$ and $c$ are coprime, we have

$$\{c \cdot a^{\mu_b j} \bmod b^n : j \geq 0\} = \{x \bmod b^n : x \equiv c \bmod b^{\ell_b}\},$$

and there exists an integer $j$ such that $c \cdot a^{\mu_b j}$ modulo $b^n$ lies in $b^n \cdot I_w$. This shows that the word $w$ occurs in the $a$-ary expansion of $c/b^n$. $\qquad \square$

The next lemma is a version of Lemma 5.1 of Hertling [15].

**Lemma 2.** *Let $a \geq 2$ and $b \geq 2$ be coprime integers. Let $(m_i)_{i \geq 1}$ be an increasing sequence of positive integers. For $i \geq 1$, let $d_i$ be an integer*

coprime with $b$ and satisfying $0 < d_i \leq b^{m_i/2}$. If for some positive integers $M$ and $j$ we have

$$m_{j+1} \geq 2(M \log_b a + m_j + 2 + \log_b a),$$

then the prefixes of length $M$ of the $a$-ary expansions of $\sum_{i=1}^{j} d_i b^{-m_i}$ and $\sum_{i \geq 1} d_i b^{-m_i}$ coincide.

*Proof.* As pointed out in the proof of Lemma 5.1 of [15], the $a$-ary expansion of any rational number $d/b^{m_j}$, where the integer $d$ is coprime with $b$, cannot have $\lfloor (\log_a b) m_j \rfloor + 1$ consecutive digits $(a-1)$. Denoting by $d'_1, \ldots, d'_M$ the first $M$ digits of the $a$-ary expansion of $\sum_{i=1}^{j} d_i b^{-m_i}$, this shows that

$$\sum_{i=1}^{j} \frac{d_i}{b^{m_i}} < \sum_{i=1}^{M} \frac{d'_i}{a^i} + \frac{1}{a^M} - \frac{1}{a^{M+1+\lfloor (\log_a b) m_j \rfloor}}$$

$$\leq \sum_{i=1}^{M} \frac{d'_i}{a^i} + \frac{1}{a^M} - \frac{1}{a^{M+1} b^{m_j}}.$$

Combined with the inequalities

$$\sum_{i \geq j+1} d_i b^{-m_i} \leq \sum_{i \geq j+1} b^{-m_i/2} \leq b^{-m_{j+1}/2}(1 + b^{-1/2} + b^{-1} + \ldots)$$

$$\leq 4 b^{-m_{j+1}/2} \leq a^{-M-1} b^{-m_j},$$

this proves the lemma. □

**Lemma 3.** *Let $b$ and $j$ be multiplicatively independent positive integers. Let $(p_k/q_k)_{k \geq 1}$ denote the sequence of convergents to $(\log bj)/(\log j)$. For any $k \geq 3bj$ and any positive integer $d$ with $d < (bj)^{(k \log 2)/2}$, the representation in base $bj$ of $dj^{p_{2k+1}}$ begins with the representation of $d$ in base $bj$.*

*Proof.* Since $j$ and $b$ are multiplicatively independent, the real number $(\log bj)/(\log j)$ is irrational and we have

$$-\frac{1}{q_{2k+1}^2} < \frac{\log bj}{\log j} - \frac{p_{2k+1}}{q_{2k+1}} < 0,$$

for every $k \geq 1$. This implies that, for $k \geq 3bj$, we have

$$j^{p_{2k+1}} > (bj)^{q_{2k+1}}$$

and

$$j^{p_{2k+1}} < (bj)^{q_{2k+1}} j^{1/q_{2k+1}} < (bj)^{q_{2k+1}}(1 + 2(\log j)/q_{2k+1})$$

$$< (bj)^{q_{2k+1}} + (bj)^{q_{2k+1}-2-(\log q_{2k+1})/(2 \log bj)}.$$

Since $q_{2k+1} \geq 2^k$, this shows that the representation in base $bj$ of $j^{p_{2k+1}}$ begins with a digit 1 followed by at least $\lceil (k \log 2)/(2 \log bj) \rceil$ zeros. Consequently, for any positive integer $d$ with $d < (bj)^{(k \log 2)/2}$, the representation in base $bj$ of $dj^{p_{2k+1}}$ begins with the representation of $d$ in base $bj$. $\qquad \square$

The next lemma, whose proof is omitted, is an easy observation.

**Lemma 4.** *Let $b \geq 2$ and $j \geq 2$ be integers. Let $d$ and $n$ be positive integers with $d < b^n$. Then the $bj$-ary expansion of $d/b^n$ terminates and, setting*

$$\frac{d}{b^n} = \frac{dj^n}{(bj)^n} := \sum_{k \geq 1} \frac{d_k}{(bj)^k},$$

*we have $d_k = 0$ for $k \geq n+1$ and for $k \leq n - \lfloor (n \log j + \log d)/\log(bj) \rfloor - 1$.*

*Completion of the proof of Theorem 2.*

We keep the notation of Section 2. Let $a \geq 2$ be an integer. Let $w$ be a finite word on $\{0, 1, \ldots, a-1\}$ of length $L$. We have to distinguish three cases.

Assume first that $a = b$. Since $b \geq 3$, the $b$-ary expansion of $W_{bj}$ has no more than $\ell_j$ digits for $j \geq 2$, and it follows from Lemma 4 that the $b$-ary expansion of $\xi$ is the concatenation of blocks of digit 0 and of the $b$-ary expansions of $c_j$ for $j \geq 2$. From the definition of $W_{b^L}$, it follows that $w$ occurs in the $b$-ary expansion of $\xi$. Consequently, $\xi$ is rich to any base which is an integral power of $b$.

Assume now that $b$ does not divide $a$, that is, that $a$ and $b$ are coprime. Let $J$ be a large integer. Note that $b$ and $c_{2J+1}$ are coprime since the last digit in the representation of $c_{2J+1}$ in base $b$ is the digit 1. Consequently, the denominator of the lowest form of the rational number $\sum_{j=1}^{2J+1} c_j/b^{n_j}$ is $b^{n_{2J+1}}$ and the $a$-ary expansion of this rational is purely periodic. By Lemma 1, if $J$ is large enough, then $w$ occurs in the $a$-ary expansion of $\sum_{j=1}^{2J+1} c_j/b^{n_j}$, which is of length at most $b^{n_{2J+1}}$. Again, for $J$ large enough, our choices of $n_{2J+2}$ and $c_{2J+2}$ imply that the assumption of Lemma 2 is satisfied with $j = 2J+1$ and $M = b^{n_{2J+1}}$. This implies that $w$ occurs in the $a$-ary expansion of $\xi$.

Finally, assume that $b$ divides $a$, but that $a$ is not a power of $b$. Set $J = a^L/b$. By Lemma 4, we may assume that $L$ is large enough to guarantee that the $a$-ary expansions of $c_j/b^{n_j}$, $j \geq 2J$, do not overlap. Note that

$$\frac{c_{2J+1}}{b^{n_{2J+1}}} = \frac{c_{2J+1}}{b^{p_{6(\ell_J + m_J \log n_{2J})+1, bJ}}} = c_{2J+1} \cdot \frac{(a^L/b)^{p_{6(\ell_J + m_J \log n_{2J})+1, bJ}}}{a^{p_{6(\ell_J + m_J \log n_{2J})+1, bJ}}}.$$

It then follows from Lemma 3 that the representation in base $a^L$ of the rational integer $c_{2J+1}(a^L/b)^{p_{6(\ell_J + m_J \log n_{2j})+1, bJ}}$ begins with the representation

11

in base $a^L$ of $c_{2J+1}$. Since $w$ can be viewed as a letter to base $a^L$, our choice of $c_{2J+1}$ and the definition of $W_{bJ}$ imply that the $a$-ary expansion of $c_{2J+1}/b^{n_{2J+1}}$ contains $w$.

This completes the proof of Theorem 2.

## 5. Proof of Theorem 3

We use repeatedly the elementary fact that, if the $b$-ary expansion of a rational number $\zeta$ reads

$$0.a_1 a_2 \ldots a_n aaa \ldots,$$

with $a, a_1, \ldots a_n$ in $\{0, 1, \ldots, b-1\}$ and $a \neq a_n$, then

$$\{b^n \zeta\} = 0.aa \ldots = \frac{a}{b-1}$$

and there exists an integer $r$ such that $\zeta = r/(b^n(b-1))$. A difficulty arises since the latter rational number may not be written under its lowest form. To see this, just observe that

$$
\begin{aligned}
\zeta &= \frac{a_n + a_{n-1}b + \ldots + a_1 b^{n-1}}{b^n} + \frac{a}{b^n(b-1)} \\
&= \frac{(a_n - a) + b(a_n - a_{n-1}) + \ldots + b^{n-1}(a_2 - a_1) + a_1 b^n}{b^n(b-1)},
\end{aligned}
$$

and $a_n - a$ has no reason to be coprime with $b$. Note however that, since $a$ is not equal to $a_n$, there exists a prime number $p$ and a positive integer $v$ such that $p^v$ divides $b$, but $p^v$ does not divide $a_n - a$. This shows that if $\zeta = A/B$ in its reduced form, then $p^n$ divides $B$.

Set $\lambda = (\log b)/(\log b')$ and $c = 2 + \lceil 2/\lambda \rceil$.

Assume that $b$ and $b'$ are coprime. Let $N$ be a large positive integer such that $\mathcal{DC}(N, \xi, b) = \mathcal{DC}(2N + c, \xi, b)$. This implies that the $(N+1)$-th, $(N+2)$-th, ..., until the $(2N+c+1)$-th digit in the $b$-ary expansion of $\xi$ are all the same. Let $n$ be the smallest positive integer such that the $(n+1)$-th, $(n+2)$-th, ..., until the $(2N+c+1)$-th digit in the $b$-ary expansion of $\xi$ are all the same. We have $n \leq N$ and there exists an integer $r$ such that

$$\left| \xi - \frac{r}{b^n(b-1)} \right| \leq \frac{1}{b^{2N+c+1}}.$$

Let $h$ be the integer defined by the inequalities

$$h + 1 \leq \lambda(2N - n + c - 1) < h + 2. \tag{5.1}$$

If the rational integer $r'$ satisfies

$$r(b')^h(b'-1) = r'b^n(b-1), \qquad (5.2)$$

then our choice of $n$ and the discussion at the beginning of this section imply that there exists a prime number $p$ dividing $b$ such that $p^n$ divides $(b')^h(b'-1)$. Since $b$ and $b'$ are coprime, $p$ does not divide $b'$, and we get that $n$ is bounded by a constant depending only on $b$ and $b'$. This shows that if $N$ is sufficiently large, in terms of $b, b'$ and $\xi$, then (5.2) holds for no rational integer $r'$. Observe that

$$2b^n(b')^h(b-1)(b'-1) \le 2b^{n+1}b^{2N-n+c-1} \le b^{2N+c+1}.$$

Consequently, if $N$ is large enough, then for any rational integer $r'$, the triangle inequality gives

$$\left| \xi - \frac{r'}{(b')^h(b'-1)} \right| \ge \left| \frac{r}{b^n(b-1)} - \frac{r'}{(b')^h(b'-1)} \right| - \left| \xi - \frac{r}{b^n(b-1)} \right|$$

$$\ge \frac{1}{b^n(b')^h(b-1)(b'-1)} - \frac{1}{b^{2N+c+1}}$$

$$\ge \frac{1}{2b^n(b')^h(b-1)(b'-1)}$$

$$\ge \frac{1}{(b')^{\lambda(n+1)+h+2}}.$$

This implies that, in the $b'$-ary expansion of $\xi$, the $(h+1)$-th, $(h+2)$-th, ..., until the $(\lambda(n+1)+h+2)$-th digit cannot be all the same, whence

$$\mathcal{DC}(\lambda(n+1)+h+1, \xi, b') \ge \mathcal{DC}(h, \xi, b') + 1.$$

It then follows from (5.1) that

$$\mathcal{DC}(2\lambda N + \lambda c, \xi, b') \ge \mathcal{DC}(\lambda N + (c-1)\lambda - 2, \xi, b') + 1$$
$$\ge \mathcal{DC}(\lambda N, \xi, b') + 1,$$

since $n \le N$ and $(c-1)\lambda - 2 > 0$, by our choice of $c$.

Set $u_1 = 1$ and $u_{n+1} = 2u_n + c$ for $n \ge 1$. A rapid calculation shows that $u_n \le (c+1)2^n$ for $n \ge 1$. We thus have proved that

$$\mathcal{DC}(u_{n+1}, \xi, b) - \mathcal{DC}(u_n, \xi, b) + \mathcal{DC}(\lambda u_{n+1}, \xi, b') - \mathcal{DC}(\lambda u_n, \xi, b') \ge 1,$$

for every integer $n$ large enough. Consequently, setting $\lambda' = \max\{1, \lambda\}$, we get

$$\mathcal{DC}(\lambda'(c+1)2^n, \xi, b) + \mathcal{DC}(\lambda'(c+1)2^n, \xi, b') \ge n,$$

for every integer $n$ large enough. This implies (2.1). Using (1.2), we get immediately (2.2).

Assume now that the bases $b$ and $b'$ are not coprime. We keep the above notation. If, for every large integer $N$, equality (5.2) holds for no integer $r'$, then we can proceed exactly as above to get the same result.

Consequently, we assume that there are integers $n, h, r, r'$ with

$$h + 1 \leq \lambda(2N - n + c - 1) < h + 2$$

and

$$r(b')^h(b' - 1) = r'b^n(b - 1). \tag{5.3}$$

We make the further assumption that there exists a prime number $p$ that divides $b$ but does not divide $b'$. Set $u = v_p(b)$, where $v_p$ is the $p$-adic valuation. It follows that

$$v_p(r) = un + \delta, \tag{5.4}$$

for some integer $\delta$ exceeding $-(\log b')/(\log p)$.

Observe also that, by our choice of $n$, there are $d$ in $\{0, 1, \ldots, b-1\}$ and an integer $R < b^n$ such that

$$\frac{r}{b^n(b-1)} = \frac{R}{b^n} + \frac{d}{b^n(b-1)}, \tag{5.5}$$

thus

$$r = (b-1)R + d. \tag{5.6}$$

We study the representation of $R$ in base $b$. Write

$$R = \sum_{j=0}^{n-1} d_j b^j, \tag{5.7}$$

where $d_j$ is in $\{0, 1, \ldots, b-1\}$.

Let $t$ be a real number strictly greater than $(\log b)/(u \log p)$. Let $J$ be a positive integer such that $d_J = d_{J+1} = \ldots = d_H$, where $H = \lfloor t(J+2) \rfloor$. We observe that

$$d + (b-1)\sum_{j=0}^{J-1} d_j b^j - d_J b^J \neq 0,$$

since, by the choice of $n$, the digits $d$ and $d_0$ are different. Furthermore, we check that

$$v_p\left(d + (b-1)\sum_{j=0}^{J-1} d_j b^j - d_J b^J\right) \leq (J+2)\frac{\log b}{\log p},$$

14

$$(b-1)\sum_{j=J}^{H} d_j b^j - d_H b^{H+1} + d_J b^J = 0,$$

and

$$v_p\left((b-1)\sum_{j=H+1}^{n-1} d_j b^j + d_H b^{H+1}\right) \geq u(H+1) \geq ut(J+2).$$

Recall that we have

$$v_p\left(d + (b-1)\sum_{j=0}^{n-1} d_j b^j\right) = un + \delta.$$

By (5.5), (5.6), (5.7), and our choice of $t$, we get a contradiction if $(J+2)(\log b) < (un+\delta)(\log p)$. Consequently, for every positive integer $J$ satisfying the latter inequality, there exists $j$ with $J \leq j \leq t(J+2)-1$ such that $d_j$ and $d_{j+1}$ are different. This shows that there are at least some constant times $\log n$ digit changes in the $b$-ary expansion of $R$. This proves (2.3) and (2.4) when there is a prime dividing $b$ but not dividing $b'$.

It remains for us to explain why the same argument works if $b$ and $b'$ have the same set of prime divisors. Assume that $b = p_1^{e_1} \cdots p_\ell^{e_\ell}$ and $b = p_1^{f_1} \cdots p_\ell^{f_\ell}$, where the $p_i$ are distinct primes and the $e_i$ and $f_i$ are positive integers. Since $b$ and $b'$ are multiplicatively independent, we may assume that the quotients $e_1/f_1$ and $e_2/f_2$ are different. Looking at (5.3), we see that for at least one of the primes $p_1$ or $p_2$ we have an equality of type (5.4). Namely, there is $i \in \{1,2\}$ and a positive integer $c$ such that $v_{p_i}(r) \geq cn$. We then continue exactly as above and get the same conclusion. $\qquad\square$

## 6. Proof of Theorem 4

We recall a result of Eggleston [13].

**Theorem E.** *Let $b \geq 2$ be an integer. Let $p_0, p_1, \ldots, p_{b-1}$ be non-negative real numbers whose sum is equal to 1. Then, the set of real numbers $\xi$ in $(0,1)$ whose $b$-ary expansion $0.a_1 a_2 a_3 \ldots$ satisfies*

$$\lim_{N \to +\infty} \frac{\text{Card}\{1 \leq n \leq N : a_n = j\}}{N} = p_j \quad \text{for } j = 0, \ldots, b-1$$

*has Hausdorff dimension*

$$-\frac{1}{\log b} \sum_{j=0}^{b-1} p_j \log p_j.$$

Taking into account the results of Nagasaka [20], Theorem 4 is an immediate consequence of Theorem 5 below.

**Theorem 5.** *Let $b \geq 2$ be an integer. Let $m, n$ be coprime positive integers with $n \geq 2$. The set of real numbers which are simply normal to base $b^m$ but not simply normal to base $b^n$ has Hausdorff dimension 1.*

*Proof.* Recalling that simple normality to base $b^{\ell m}$ implies simple normality to base $b^m$, we assume that $m$ is greater than $n$. We work in base $B = b^{2mnb^m}$. Let $w$ be a word on $\{0, 1, \ldots, b-1\}$ obtained as the concatenation of $2n-2$ copies of $0^m$, of $2n-1$ copies of $1^m$, and of $2n$ copies of each other word of length $m$ on $\{0, 1, \ldots, b-1\}$. Clearly, the length of $w$ is then

$$m(4n - 3) + 2nm(b^m - 2) = 2nmb^m - 3m.$$

Put

$$w_0 = w0^m1^m0^m \qquad \text{and} \qquad w_1 = w0^{2m}1^m.$$

The words $w_0$ and $w_1$ have length $2nmb^m$, thus they can be written to base $b^n$ as words of length $2mb^m$. Writing $2m + h = jn$ and $m - h' = j'n$ with $0 \leq h, h' < n$ and $j, j'$ integers, the last $jn$ digits of $w_0$ to base $b$ are $0^h1^{n-h}, 1^n, \ldots, 1^n, 1^{n-h'}0^{h'}, 0^n, \ldots, 0^n$, while the last $jn$ digits of $w_1$ to base $b$ read $0^n, \ldots, 0^n, 0^{n-h'}1^{h'}, 1^n, \ldots, 1^n$. Since $n$ does not divide $m$, we have $h' \neq 0$, thus, at least one word among $w_0$ and $w_1$ cannot be written as a concatenation, in some order, of $2m$ copies of every letter to base $b^n$. Without any loss of generality, we may assume that $w_0$ shares this property.

Let $\mathcal{B}$ be the set of real numbers $\xi$ whose expansion to base $B$ has the property that every letter, except $w_0$, occurs with the same frequency $1/(B-1)$. Then, by Theorem E, this set has Hausdorff dimension $\log(B-1)/\log B$.

Our choice of $w_0$ implies that each element of $\mathcal{B}$ is simply normal to base $b^m$, but not simply normal to base $b^n$. Recalling that simple normality to base $b^{\ell m}$ implies simple normality to base $b^m$, we get our result by replacing $m$ by an arbitrarily large multiple of $m$ coprime with $n$. $\qquad\square$

# References

[1] R. K. Akhunzhanov, *On nonnormal numbers*, Mat. Zametki 72 (2002), 150–152 (in Russian). English translation in Math. Notes 72 (2002), 135–137.

[2] R. K. Akhunzhanov, *On the distribution modulo 1 of exponential sequences*, Mat. Zametki 76 (2004), 163–171 (in Russian). English translation in Math. Notes 76 (2004), 153–160.

[3] D. H. Bailey, J. M. Borwein, R. E. Crandall and C. Pomerance, *On the binary expansions of algebraic numbers*, J. Théor. Nombres Bordeaux 16 (2004), 487–518.

[4] V. Becher and S. Figueira, *An example of a computable absolutely normal number*, Theoret. Comput. Sci. 270 (2002), 947–958.

[5] V. Becher, S. Figueira, and R. Picchi, *Turing's unpublished algorithm for normal numbers*, Theoret. Comput. Sci. 377 (2007), 126–138.

[6] É. Borel, *Les probabilités dénombrables et leurs applications arithmétiques*, Rend. Circ. Math. Palermo 27 (1909), 247–271.

[7] R. Broderick, Y. Bugeaud, L. Fishman, D. Kleinbock and B. Weiss, *Schmidt's game, fractals, and numbers normal to no base*, Math. Res. Lett. 17 (2010), 307–321.

[8] Y. Bugeaud, *On the b-ary expansion of an algebraic number*, Rend. Sem. Mat. Univ. Padova 118 (2007), 217–233.

[9] Y. Bugeaud, Distribution modulo one and Diophantine approximation. Cambridge Univ. Press. To appear.

[10] J. W. S. Cassels, *On a problem of Steinhaus about normal numbers*, Colloq. Math. 7 (1959), 95–101.

[11] K. J. Compton, *On rich words*. In: Combinatorics on words (Waterloo, Ont., 1982), 39–61, Academic Press, Toronto, ON, 1983.

[12] A. de Luca, A. Glen, and L. Q. Zamboni, *Rich, Sturmian, and trapezoidal words*, Theoret. Comput. Sci. 407 (2008), 569–573.

[13] H. G. Eggleston, *The fractional dimension of a set defined by decimal properties*, Quart. J. Math. Oxford Ser. 20 (1949), 31–36.

[14] S. I. El-Zanati and W. R. R. Transue, *On dynamics of certain Cantor sets*, J. Number Theory 36 (1990), 246–253.

[15] P. Hertling, *Disjunctive ω-words and real numbers*, J. UCS 2 (1996), 549–568.

[16] P. Hertling, *Simply normal numbers to different bases*, J. UCS 8 (2002), 235–242.

[17] Ch. Mauduit and C. G. Moreira, *Generalized Hausdorff dimensions of sets of real numbers with zero entropy expansion*, Ergodic Theory Dynam. Systems. To appear.

[18] M. Morse and G. A. Hedlund, *Symbolic dynamics*, Amer. J. Math. 60 (1938), 815–866.

[19] M. Morse and G. A. Hedlund, *Symbolic dynamics II*, Amer. J. Math. 62 (1940), 1–42.

[20] K. Nagasaka, *Nonnormal numbers to different bases and their Hausdorff dimension*, Tsukuba J. Math. 10 (1986), 89–99.

[21] W. M. Schmidt, *On normal numbers*, Pacific J. Math. 10 (1960), 661–672.

[22] W. M. Schmidt, *Über die Normalität von Zahlen zu verschiedenen Basen*, Acta Arith. 7 (1961/1962), 299–309.

[23] W. M. Schmidt, *On badly approximable numbers and certain games*, Trans. Amer. Math. Soc. 123 (1966), 178–199.

[24] H. G. Senge and E. G. Straus, PV-*numbers and sets of multiplicity*, Period. Math. Hungar. 3 (1973), 93–100.

[25] W. Sierpiński, *Démonstration élémentaire du théorème de M. Borel sur les nombres absolument normaux et détermination effective d'un tel nombre*, Bull. Soc. Math. France 45 (1917), 127–132.

[26] C. L. Stewart, *On the representation of an integer in two different bases*, J. reine angew. Math. 319 (1980), 63–72.

Yann Bugeaud
Université de Strasbourg
Mathématiques
7, rue René Descartes
67084 STRASBOURG (FRANCE)

`bugeaud@math.unistra.fr`