Rendiconti di Matematica, Serie VII
Volume 19, Roma (1999), 615-634

On the expansions in non-integer bases

V. KOMORNIK – P. LORETI

RIASSUNTO: *In* [2] *Erdös, Horváth e Joó trovarono una proprietà di unicità per l'espansione del numero* 1 *in qualche base non intera* q*. Questi numeri* q *sono stati poi caretterizzati in* [3]*. Utilizzando questa caratterizzazione il più piccolo di tali numeri è stato determinato in* [8]*. Allouche e Cosnard* [1] *hanno trovato che questo numero è in relazione con la ben nota successione di Thue-Morse, e usando questa relazione hanno poi dimostrato la transcendenza di questo numero. Erdös e Joó* [4] *hanno costruito anche, per ogni intero positivo* N*, basi* q *per le quali* 1 *ha esattamente* N *differenti espansioni. La loro costruzione è stata generalizzata in* [5]*. La caratterizzazione di questi numeri rimane una questione aperta per* $N > 1$*. Lo scopo di questo lavoro è di fornire una condizione utile e sufficiente per* $N = 2$ *e di usare questo per la costruzione di numeri* q *piccoli per i quali il numero* 1 *ha esattamente* 2 *differenti espansioni.*

ABSTRACT: *In* [2] *Erdös, Horváth and Joó found a curious uniqueness property for the expansions of the number* 1 *in some noninteger bases* q*. These numbers* q *were then characterized in* [3]*. Using that characterization the smallest such number was determined in* [8]*. Allouche and Cosnard* [1] *discovered that this number is closely related to the classical Thue-Morse sequence, and using this relation they proved the transcendence of this number. Erdös and Joó* [4] *also constructed, for each positive integer* N*, bases* q *for which the number* 1 *has exactly* N *different expansions. Their construction was generalized in* [5]*. The characterization of these numbers remains an open question for* $N > 1$*. The purpose of this paper is to give a useful sufficient condition for* $N = 2$ *and to use this for the construction of small numbers* q *for which the number* 1 *has exactly* 2 *different expansions.*

KEY WORDS AND PHRASES: *Noninteger bases – Special sequences.*
A.M.S. CLASSIFICATION: 11A67 – 11B83

1 – A review of greedy expansions

In [11] Rényi introduced a generalization of the familiar dyadic expansions of real numbers and he studied the ergodic properties of the distribution of the digits. His results were extended by Parry [10].

Analogous problems, of probabiliste nature, appeared in some works of T. Varga on the expected length of the longest head run; they were studied more thoroughly by Erdös and Révész [6]. In this framework the earlier studies correspond to the case of unsymmetric coins where the probability of the head is some number $1/q$ instead of $1/2$.

A later work of Erdös, Horváth and Joó [2] revealed that all these questions have an essentialy combinatorial nature. This led to a series of papers by various authors; see, e.g., the review paper [7] of Joó and Schnitzer for a more detailed analysis. For the reader's convenience we recall in this and the next section some of these results, by giving direct and simpler proofs than the original ones.

The results of this section are very close (although not identical) to some theorems obtained earlier by Rényi and Parry in [10], [11]. We prove them in a different, direct way.

Fix a real number $1 < q \le 2$. By an *expansion* of a real number x we mean a *sequence* $c_1, c_2, \ldots$ of integers in $\{0, 1\}$ satisfying the equality

$$\sum_{i=1}^{\infty} \frac{c_i}{q^i} = x. \tag{1.1}$$

Proposition 1.1 (Cf. [11]). *A real number x has an expansion if and only if*

$$0 \le x \le \frac{1}{q-1}. \tag{1.2}$$

Proof. The necessity of the condition is obvious because

$$\sum_{i=1}^{\infty} \frac{1}{q^i} = \frac{1}{q-1}.$$

Conversely, if this condition is fulfilled, then an expansion can be obtained for example by the so-called *greedy* algorithm: we always choose the biggest possible value for c_i. More precisely, define recursively a sequence a_1, a_2, ... as follows. If for some positive integer n the numbers a_i are defined for all $i < n$ (no assumption if $n = 1$), then set $a_n = 1$ if

$$\sum_{i=1}^{n-1} \frac{a_i}{q^i} + \frac{1}{q^n} \le x \tag{1.3}$$

and set $a_n = 0$ otherwise. Since $x \ge 0$, we can start this construction and hence the definition is meaningful.

We claim that (a_i) is an expansion. First we note that we cannot have a *last* index n such that $a_n = 0$. Indeed, then we would have

$$\sum_{i=1}^{n} \frac{a_i}{q^i} > x - \frac{1}{q^n}$$

and

$$\sum_{i=1}^{n} \frac{a_i}{q^i} + \sum_{i=n+1}^{k} \frac{1}{q^i} \le x$$

for all $k > n$. These inequalities imply that

$$\sum_{i=n+1}^{\infty} \frac{1}{q^i} < \frac{1}{q^n}.$$

This is equivalent to $q > 2$, contradicting the choice of q. So we have either $a_n = 1$ for all n, or $a_n = 0$ for infinitely many indices n.

In the first case we obtain

$$\sum_{i=1}^{\infty} \frac{a_i}{q^i} = \sum_{i=1}^{\infty} \frac{1}{q^i} = \frac{1}{q-1} \ge x,$$

proving that the inequality in (1.3) is in fact an equality. In the second case we have

$$\sum_{i=1}^{n} \frac{a_i}{q^i} > x - \frac{1}{q^n}$$

for all n satisfying $a_n = 0$. Letting $n \to \infty$ we conclude again that the inverse inequality to (1.3) holds true. □

REMARK. The greedy algorithm also provides an expansion of $x = 0$ and of $x = 1$ if $q = 1$, given by $a_1 = x$ and $a_i = 0$ for all $i > 1$.

We are going to characterize the greedy expansions by using the *lexicographic order* between sequences: given two sequences (b_i) and (c_i), we write $(b_i) < (c_i)$ or $b_1 b_2 \cdots < c_1 c_2 \ldots$ if there exists a positive integer n such that $b_i = c_i$ for all $i < n$ but $b_n < c_n$. This is a complete ordering. Theorems 1.2 and 1.3 below are very close to former theorems of Parry.

DEFINITION. A sequence $a_1, a_2, \ldots$ of integers in $\{0, 1, \ldots, m\}$ is *distinguished* if

$$a_{n+1} a_{n+2} \cdots < a_1 a_2 \ldots \text{ whenever } a_n = 0. \tag{1.4}$$

THEOREM 1.2 (Cf. [10]). *Let us denote by (ε_i) the greedy expansion of 1 for $1 \le q \le 2$. Then the map $q \mapsto (\varepsilon_i)$ is a strictly increasing bijection of the closed interval $[1, 2]$ onto the set of distinguished sequences.*

THEOREM 1.3 (Cf. [10]). *Fix $1 < q \le 2$ arbitrarily and let us denote by (ε_i) the corresponding greedy expansion of 1. Furthermore, let us denote by (a_i) the greedy expansion of some x.*

(a) *Assume that the sequence (ε_i) is infinite, i.e., it contains infinitely many nonzero elements. Then the map $x \mapsto (a_i)$ is a strictly increasing bijection of the closed interval $[0, 1/(q-1)]$ onto the set of all sequences (a_i) satisfying*

$$a_{n+1} a_{n+2} \cdots < \varepsilon_1 \varepsilon_2 \ldots \quad \text{whenever} \quad a_n = 0. \tag{1.5}$$

(b) *Assume that the sequence (ε_i) is finite, i.e., all but finitely many elements vanish. Let ε_k be its last nonzero element. Then the map $x \mapsto (a_i)$ is a strictly increasing bijection of the closed interval $[0, 1/(q-1)]$ onto the set of all sequences (a_i) satisfying (1.5) and which are not eventually periodic with period $\varepsilon_1 \ldots \varepsilon_{k-1}(\varepsilon_k - 1)$.*

We need two lemmas. Fix $1 < q \le 2$ arbitrarily and denote by (ε_i) the corresponding greedy expansion of 1.

LEMMA 1.4. (a) *The greedy expansion* (a_i) *of every number* $x \in [0, 1/(q-1)]$ *satisfies the condition* (1.5).

(b) *If the sequence* (ε) *is finite with a last nonzero digit* ε_k*, then no greedy expansion is eventually periodic with the period* $\varepsilon_1, \dots, \varepsilon_{k-1}, \varepsilon_k - 1$.

PROOF. (a) If $(a_{n+i}) > (\varepsilon_i)$ for some n, then there exists an integer k such that

$$a_{n+i} = \varepsilon_i \quad \text{for} \quad i = 1, \dots, k-1$$

but

$$a_{n+k} > \varepsilon_k.$$

Then we have

$$\sum_{i=1}^{k} \frac{a_{n+i}}{q^i} > 1$$

by the definition of the sequence (ε_i). Hence

$$\sum_{i<n} \frac{a_i}{q^i} + \frac{a_n + 1}{q^n} < \sum_{i=1}^{\infty} \frac{a_i}{q^i} \le x.$$

This contradicts the definition of a_n unless $a_n = 1$.

If $(a_{n+i}) = (\varepsilon_i)$ for some n, then

$$\sum_{i<n} \frac{a_i}{q^i} + \frac{a_n + 1}{q^n} = \sum_{i=1}^{\infty} \frac{a_i}{q^i} = x$$

This contradicts the definition of a_n again unless $a_n = 1$.

(b) Assume on the contrary that the greedy expansion (a_n) of some x is eventually periodic with the period $\varepsilon_1, \dots, \varepsilon_{k-1}, \varepsilon_k - 1$, and let a_p be the last element of a period. Then changing $a_n = \varepsilon_k - 1$ to ε_k and a_n to 0 for all $n > p$, we obtain a lexicographically larger expansion of x, contradicting the definition of the greedy expansion. ∎

LEMMA 1.5. (a) *Let* (e_i) *be an infinite expansion of* 1*. Let* (a_i) *be an expansion of some real number* x*, satisfying*

$$a_{n+1}a_{n+2}\cdots < e_1e_2\ldots \quad \textit{whenever} \quad a_n = 0. \tag{1.6}$$

Then (a_i) *is the greedy expansion of* x*.*

(b) *Let* (e_i) *be a finite expansion of* 1 *and denote* e_k *its last nonzero element. Let* (a_i) *be an expansion of some real number* x*, satisfying* (1.6)*, and assume that* (a_i) *is not eventually periodic with period* $e_1\ldots e_{k-1}(e_k-1)$*. Then* (a_i) *is the greedy expansion of* x*.*

PROOF. There is nothing to prove if $a_n = 1$ for all n. Otherwise fix n such that $a_n = 0$. We have to show that

$$\sum_{i=1}^{\infty}\frac{a_{n+i}}{q^i} < 1 \quad \text{whenever} \quad a_n = 0. \tag{1.7}$$

Using the hypothesis of the lemma we can construct a sequence of integers

$$n = k_0 < k_1 < \ldots$$

satisfying for each $j = 1, 2, \ldots$ the conditions

$$a_{k_{j-1}+i} = e_i \quad \text{for all} \quad 1 \le i < k_j - k_{j-1}$$

and

$$a_{k_j} < e_{k_j-k_{j-1}}.$$

(a) If the sequence (e_i) is infinite, then

$$\sum_{i=1}^{\infty}\frac{a_{n+i}}{q^i} \le \sum_{j=1}^{\infty}\Big(\sum_{i=1}^{k_j-k_{j-1}}\frac{e_i}{q^{k_{j-1}+i}} - \frac{1}{q^{k_j}}\Big) < \sum_{j=1}^{\infty}\Big(\frac{1}{q^{k_{j-1}}} - \frac{1}{q^{k_j}}\Big) = 1,$$

proving (1.7).

(b) If the sequence (e_i) is finite, then the above proof leads to (1.7) with $\le$ instead of $<$. A closer inspection of the proof reveals that we obtain $=$ exactly if $k_j - k_{j-1} = k$ for every j and if the sequence (a_{n+i}) is periodic with period $e_1\ldots e_{k-1}(e_k-1)$. However, this case was excluded. □

Proof of Theorem 1.2. It follows at once from the definition of the greedy expansion that the map $q \mapsto (\varepsilon_i)$ is strictly increasing.

Lemma 1.4 shows that for any $1 < q \leq 2$ the greedy expansion (ε_i) of 1 is a distinguished sequence. This is also for true $q = 1$ because the sequence $1000\dots$ is obviously distinguished.

Conversely, let (e_i) be a distinguished sequence satisfying

$$\sum_{i=1}^{\infty} \frac{e_i}{q^i} = 1$$

for some $1 \leq q \leq 2$. We claim that then (e_i) is the greedy expansion of 1 for this q. Indeed, if $e_n = 0$ for some n, then we deduce from Lemma 1.5 that

$$\sum_{i<n} \frac{e_i}{q^i} + \frac{e_n + 1}{q^n} > \sum_{i=1}^{\infty} \frac{e_i}{q^i} = 1.$$

This proves that (e_i) is the greedy expansion of 1. ∎

Proof of Theorem 1.3. Again, the strict increasingness of the map $x \mapsto (a_i)$ follows from the definition of greedy expansions. Furthermore, Lemma 1.4 shows that (a_i) satisfies the condition (1.5) for every x.

Conversely, let (a_i) be an expansion of some number x and assume that the condition (1.5) is satisfied. If $a_n = 0$ for some n, then applying Lemma 1.5 we obtain that

$$\sum_{i<n} \frac{a_i}{q^i} + \frac{a_n + 1}{q^n} > \sum_{i=1}^{\infty} \frac{a_i}{q^i} = x.$$

This proves that (a_i) is the greedy expansion of x. ∎

2 – Lazy and unique expansions

In this section we recall some results from [3] and [5].

Fixing $1 < q \leq 2$ again, we may define for every $0 \leq y \leq 1/(q-1)$ another expansion (b_i) of y by the so-called *lazy* algorithm: we choose always the smallest possible nonnegative integer b_i. More precisely, define recursively a sequence $b_1, b_2, \dots$ as follows. If for some positive integer

n the numbers b_i are defined for all $i < n$ (no assumption if $n = 1$), then set $b_n = 0$

$$\sum_{i=1}^{n} \frac{b_i}{q^i} + \sum_{i>n} \frac{1}{q^i} \geq y, \tag{2.1}$$

and set $b_n = 1$ otherwise. It follows at once from this definition that (b_i) is the lazy expansion of y if and only if the sequence $(a_i) := (1 - b_i)$ is the greedy expansion of $x := (q-1)^{-1} - y$. Using this "duality" relation we deduce from Proposition 1.1 that every $0 \leq y \leq 1/(q-1)$ has a lazy expansion. Furthermore, it follows at once from the definitions that if (a_i) and (b_i) are the greedy and lazy expansions of some x and if there exists another expansion (c_i) of x, then

$$(b_i) < (c_i) < (a_i).$$

In other words, the greedy expansion is the greatest expansion and the lazy expansion is the smallest expansion of a given number x with respect to the lexicographic order.

Using this duality, we deduce from Lemma 1.5 the

PROPOSITION 2.1(Cf. [5]). *Let (e_i) be an infinite expansion of* 1. (*This means that $e_i > 0$ for infinitely many indices i.*) *If an expansion (b_i) of some number y satisfies the condition*

$$(1 - b_{n+i}) < (e_i) \quad \textit{whenever} \quad b_n > 0, \tag{2.2}$$

then (b_i) is the lazy expansion of y.

REMARK. This condition is not necessary. For example, the expansion of 0 is unique and hencre lazy, but the corresponding sequence $b_i \equiv 0$ for all i does not satisfy (2.2). It would be interesting to find a necessary and sufficient condition.

Now we can determine those numbers q for which the greedy and lazy expansions of 1 coincide. Equivalently, these are the numbers q for which the expansion of 1 is unique. Let us introduce a subset of the set of distinguished sequences:

Definition. A sequence $a_1, a_2, \ldots$ of integers in {0,1} is 1-*admissible* if

$$(a_{n+i}) < (a_i) \text{ whenever } a_n = 0 \tag{2.3}$$

and

$$(1 - a_{n+i}) < (a_i) \text{ whenever } a_n = 1. \tag{2.4}$$

In the sequel we shall often write $\overline{\varepsilon_i}$ instead of $1 - \varepsilon_i$ and $\overline{s}$ instead of $\overline{\varepsilon_1} \ldots \overline{\varepsilon_n}$ if $s = \varepsilon_1 \ldots \varepsilon_n$ for brevity. Thus the condition (2.4) may be rewritten in the form

$$(\overline{a_{n+i}}) < (a_i) \text{ whenever } a_n = 1.$$

Theorem 2.2 (Cf. [3]). *The number* 1 *has a unique expansion for a given* q *if and only if the greedy expansion* (ε_i) *of* 1 *is a* 1*-admissible sequence. In other words, the bijection* $q \mapsto (\varepsilon_i)$ *of Theorem* 1.2 *establishes also a bijection between the set of numbers* q *having the uniqueness property and the set of* 1*-admissible sequences.*

It follows from this theorem that there are continuum many q's having this curious uniqueness property. See [8] for the determination of the smallest such q, and [1] for the proof of its transcendence.

3 – Numbers with exactly two expansions

Let $\varepsilon = (\varepsilon_i)$ be a sequence of zeroes and ones, satisfying the condition

$$\varepsilon_{n+1}\varepsilon_{n+2} \cdots < \varepsilon_1\varepsilon_2 \ldots \quad \text{whenever} \quad \varepsilon_n = 0. \tag{3.1}$$

Then $\varepsilon_1 = 1$. Indeed, otherwise applying (3.1) we would obtain by induction that $\varepsilon_i \equiv 0$, contradicting (3.1). Hence there exists a unique $q \in [1, 2]$ satisfying

$$\sum_{i=1}^{\infty} \varepsilon_i q^{-i} = 1. \tag{3.2}$$

By Theorem 1.2 (ε_i) is the corresponding greedy expansion of 1.

Assume that there exists a positive integer m satisfying

$$\varepsilon_m = 1 \tag{3.3}$$

and

$$\varepsilon_i + \varepsilon_{i+m} \in \{0, 1\} \quad \text{for all} \quad i \geq 1. \tag{3.4}$$

Then we can define another expansion

$$\sum_{i=1}^{\infty} \delta_i q^{-i} = 1 \tag{3.5}$$

of 1 by setting

$$\delta_i = \begin{cases} \varepsilon_i & \text{if } i < m, \\ 0 & \text{if } i = m, \\ \varepsilon_i + \varepsilon_{i-m} & \text{if } i > m. \end{cases} \tag{3.6}$$

Indeed, thanks to (3.4), $\delta = (\delta_i)$ is a sequence of zeroes and ones. Furthermore, using (3.2) and (3.3) we have

$$\sum_{i=1}^{\infty} \delta_i q^{-i} = \sum_{i<m} \varepsilon_i q^{-i} + \sum_{i>m} (\varepsilon_i + \varepsilon_{i-m}) q^{-i} =$$
$$= \sum_{i \neq m} \varepsilon_i q^{-i} + q^{-m} \sum_{i=1}^{\infty} \varepsilon_i q^{-i} = \sum_{i=1}^{\infty} \varepsilon_i q^{-i} = 1.$$

Now we give a *sufficient* condition for q in order to have exactly two expansions of 1. In the sequel we write for brevity $\overline{\varepsilon_i}$ instead of $1 - \varepsilon_i$ and also $\overline{s}$ instead of $\overline{\varepsilon_1 \varepsilon_2} \ldots$ if $s = \varepsilon_1 \varepsilon_2 \ldots$ is a finite or infinite sequence of zeroes and ones.

THEOREM 3.1. *Assume* (3.1), (3.3), (3.4) *and define* q *by* (3.2). *Furthermore, define the sequence* (δ_i) *by* (3.6), *and assume that*

$$\overline{\delta_{n+1} \delta_{n+2} \ldots} < \varepsilon_1 \varepsilon_2 \ldots \quad \textit{whenever} \quad \delta_n = 1, \tag{3.7}$$

$$\overline{\varepsilon_{n+1} \varepsilon_{n+2} \ldots} < \varepsilon_1 \varepsilon_2 \ldots \quad \textit{whenever} \quad \varepsilon_n = 1 \quad \textit{and} \quad n > m, \tag{3.8}$$

$$\delta_{n+1} \delta_{n+2} \cdots < \varepsilon_1 \varepsilon_2 \ldots \quad \textit{whenever} \quad \delta_n = 0 \quad \textit{and} \quad n > m. \tag{3.9}$$

Then for this q *there are exactly two different expansions: those given by* (3.2) *and* (3.5).

Proof. The sequence (ε_i) is infinite: otherwise (3.7) is not satisfied for the last $\delta_n = 1$. Hence $q > 1$. Applying Theorem 1.2 and Proposition 2.1 we deduce from (3.1) and (3.7) that (3.2) and (3.5) are the greedy and lazy expansions of 1.

It remains to verify that if a sequence (ρ_i) of zeroes and ones satisfies the strict inequalities

$$(\delta_i) < (\rho_i) < (\varepsilon_i),$$

then

$$\sum_{i=1}^{\infty} \rho_i q^{-i} \neq 1. \tag{3.10}$$

Fix such a sequence (ρ_i), then $\rho_i = \delta_i = \varepsilon_i = 1$ for all $i < m$. Since $\delta_m = 0$ and $\varepsilon_m = 1$, we have either $\rho_m = \delta_m$ or $\rho_m = \varepsilon_m$. We distinguish two cases.

First case: $\rho_m = 0$. Then there is an integer $n > m$ such that $\rho_i = \delta_i$ for all $i < n$ and $\delta_n = 0 < 1 = \rho_n$. Using (3.9) and applying Lemma 1.5 with $(e_i) = (\varepsilon_i)$ and $(a_i) = (\delta_{n+i})$, we obtain that

$$\sum_{i=1}^{\infty} \delta_{n+i} q^{-i} < 1.$$

Therefore

$$\begin{aligned}\Big(\sum_{i=1}^{\infty} \rho_i q^{-i}\Big) - 1 &= \Big(\sum_{i=1}^{\infty} \rho_i q^{-i}\Big) - \Big(\sum_{i=1}^{\infty} \delta_i q^{-i}\Big) = \\ &= q^{-n} + \sum_{i=n+1}^{\infty} (\rho_i - \delta_i) q^{-i} \geq q^{-n} - \sum_{i=n+1}^{\infty} \delta_i q^{-i} = \\ &= q^{-n}\Big(1 - \sum_{i=1}^{\infty} \delta_{n+i} q^{-i}\Big) > 0,\end{aligned}$$

proving (3.10).

Second case: $\rho_m = 1$. Then there is an integer $n > m$ such that $\rho_i = \varepsilon_i$ for all $i < n$ and $\rho_n = 0 < 1 = \varepsilon_n$. Using (3.8) and applying Lemma 1.5 with $(e_i) = (\varepsilon_i)$ and $(a_i) = (\overline{\varepsilon_{n+i}})$, we obtain that

$$\sum_{i=1}^{\infty} \overline{\varepsilon_{n+i}} q^{-i} < 1.$$

Hence

$$\Big(\sum_{i=1}^{\infty}\rho_i q^{-i}\Big)-1=\Big(\sum_{i=1}^{\infty}\rho_i q^{-i}\Big)-\Big(\sum_{i=1}^{\infty}\varepsilon_i q^{-i}\Big)=$$

$$=-q^{-n}+\sum_{i=n+1}^{\infty}(\rho_i-\varepsilon_i)q^{-i}\le -q^{-n}+\sum_{i=n+1}^{\infty}\overline{\varepsilon_i}q^{-i}=$$

$$=-q^{-n}\Big(1-\sum_{i=1}^{\infty}\overline{\varepsilon_{n+i}}q^{-i}\Big)<0,$$

implying (3.10) again. ∎

4 – Small numbers with exactly two expansions

A sequence (ε_i) satisfying the conditions of Theorem 3.1 is called *admissible*, and the corresponding unique positive solution of the equation

$$\sum_{n=1}^{\infty}\varepsilon_n q^{-n}=1$$

is also called *admissible*. By Theorem 3.1 for admissible q's there are exactly two different expansions. In this section we are looking for the smallest admissible q's.

Let us consider the sequence (ε'_n) given by

$$111\ 001\ 001\ 001\ \cdots = 111\ \underline{\underline{001}};$$

the symbol $\underline{\underline{s}}$ will be used here and in the sequel to denote the period s of a periodic sequence. (Note that this sequence is not admissible but 1-admissible; hence it has the uniqueness property.) Denoting the unique positive solution of the equation

$$\sum_{n=1}^{\infty}\varepsilon_n q^{-n}=1$$

by q', we have $q'\approx 1.871349313$.

We shall prove the

THEOREM 4.1. *All admissible numbers are greater than q'. On the other hand, q' is an accumulation point of the set of admissible numbers.*

First we note that the sequence (ε'_n) satisfies (3.1) (with ε'_i instead of ε_i), so that it corresponds to the greedy expansion of 1 for q', but (3.3) and (3.4) cannot be satisfied simultaneously for any m, so that q' is not admissible.

Now choose an arbitrarily large positive integer k and insert between the kth and $(k+1)$th block 001 a block $100\dots0$ formed by one followed by $3k+4$ zeroes. Then we obtain an admissible sequence ε with $m = 3k+4$. For example, for $k = 1$ we have $m = 7$ and

$$\begin{aligned}\varepsilon &= 111001\ 10000000\ 001\ 001\ 001\ 001\ 001\ 001\ \dots\\ \delta &= 111001\ 01110011\ 001\ 001\ 001\ 101\ 101\ 101\ \dots;\end{aligned}$$

for $k = 2$ we obtain $m = 10$ and

$$\begin{aligned}\varepsilon &= 111001001\ 10000000000\ 001\ 001\ 001\ 001\ 001\ 001\ 001\ \dots\\ \delta &= 111001001\ 01110010011\ 001\ 001\ 001\ 001\ 101\ 101\ 101\ \dots.\end{aligned}$$

One can readily verify that the corresponding numbers q tend to q' as $k \to \infty$.

In order to complete the proof of the theorem, it remains to establish the inequality $q \geq q'$ for every admissible number q. Equivalently, it suffices to prove the inequality

$$(\varepsilon_n) \geq (\varepsilon'_n)$$

(in the lexicographic order) for every admissible sequence (ε_n). We may assume without loss of generality that

$$\varepsilon_1 \dots \varepsilon_5 \leq 11100. \tag{4.1}$$

in the lexicographic order. (The above examples show that such sequences exist.) We need a series of lemmas.

Lemma 4.2. *Every admissible sequence ε contains infinitely many one digits and infinitely many zero digits.*

Proof. If ε had a last one digit $\varepsilon_j = 1$, then δ would also have a last one digit $\delta_{j+m} = 1$, and then (3.7) would not hold for $n = j + m$. The other property follows at once from (3.4). ∎

It follows in particular from this lemma that there exists a first one digit $\varepsilon_k = 1$ in the subsequence $\varepsilon_{m+1}\varepsilon_{m+2}\dots$.

LEMMA 4.3. *For every admissible sequence ε the sequence δ contains infinitely many zero digits.*

PROOF. Assume on the contrary that δ has a last zero digit $\delta_n = 0$. Since $\delta_m = 0$, we have necessarily $n \geq m$. If $n > m$, then (3.9) is not satisfied. Hence we have necessarily $n = m$, so that $\delta_i = 1$ for all $i > m$.

Now we deduce from (3.6) that

$$\varepsilon_{m+1}\dots\varepsilon_{2m} = \overline{\varepsilon_1\dots\varepsilon_m},$$

and then by induction that ε is $2m$-periodic. However, this contradicts (3.1). ∎

It follows from this lemma that there exists a first zero digit $\delta_l = 0$ in the subsequence $\delta_{m+1}\delta_{m+2}\dots$.

LEMMA 4.4. *Every admissible sequence ε begins with* 111.

PROOF. Let ε be an admissible sequence. Then $\varepsilon_1 = 1$ by (3.1).

Next we claim that $\varepsilon_2 = 1$. Indeed, assume on the contrary that (ε_i) begins with 10. Then ε cannot contain 11 by (3.1). Therefore $\varepsilon \leq 1010\dots$. Next we deduce from (3.8) that the subsequence $\varepsilon_{k+1}\varepsilon_{k+2}\dots$ does not contain 00 either. Hence $\varepsilon_n\varepsilon_{n+1}\dots = 0101\dots$ for some $n > m$ and then $\varepsilon_{n+1}\varepsilon_{n+2}\dots \geq \varepsilon_1\varepsilon_2\dots$, contradicting (3.1).

Now we prove that $\varepsilon_3 = 1$. Assume on the contrary that ε begins with 110. Then ε cannot contain 111 by (3.1) and hence $\varepsilon \leq 110110\dots$. Next we deduce from (3.8) that the subsequence $\varepsilon_{k+1}\varepsilon_{k+2}\dots$ does not contain 000 either. Using (3.6) it follows that out of three consecutive elements of the subsequence $\delta_{k+m+1}\varepsilon_{k+m+2}\dots$ at most one can be zero. Using also the preceding lemma, there exists $n > m+k+1$ with $\delta_n = 0$ and $\delta_{n+1}\delta_{n+2} \geq 110110\dots$. But this contradicts (3.9). ∎

Until now we did not use (4.1). From now on we shall need it. Thanks to Lemma 4.4 all these sequences begin with 11100. Hence $m \geq 3$, and it follows from (3.3), (3.4) and (3.6) that

$$\begin{aligned} &\varepsilon_m = 1, \quad \varepsilon_{m+1} = \varepsilon_{m+2} = \varepsilon_{m+3} = \varepsilon_{2m} = 0, \\ &\delta_m = 0, \quad \delta_{m+1} = \delta_{m+2} = \delta_{m+3} = \delta_{2m} = 1. \end{aligned}$$

Hence $k, l \geq m + 4$.

Lemma 4.5. *There is no* 0000 *in the subsequence* $\varepsilon_{k+1}\varepsilon_{k+2}\ldots$.

Proof. Since $\varepsilon_k = 1$, otherwise ε would contain 10000 beginning with $\varepsilon_n = 1$ for some $n \geq k > m$, contradicting (3.9). ∎

Lemma 4.6. *There is no* 1111 *in the subsequence* $\delta_{l+1}\delta_{l+2}\ldots$.

Proof. Since $\delta_l = 0$, otherwise δ would contain 01111 beginning with $\delta_n = 1$ for some $n \geq l > m$, contradicting (3.9). ∎

Lemma 4.7. *There is no* 11 *in the subsequence*

$$\varepsilon_K\varepsilon_{K+1}\ldots, \qquad K = \max\{k - m + 2, l - m + 2\}.$$

Proof. Assume on the contrary that $\varepsilon_n = \varepsilon_{n+1} = 1$ for some $n \geq K$. Then $\delta_{m+n} = \delta_{m+n+1} = 1$ by (3.6) and $\varepsilon_{m+n} = \varepsilon_{m+n+1} = 0$ by (3.4). We distinguish four cases according to the values of ε_{m+n-1} and ε_{m+n+2}:

First case: $\varepsilon_{m+n-1} = \varepsilon_{m+n+2} = 0$. This is impossible by Lemma 4.5 because $m + n - 1 \geq k + 1$. (See the figure below.)

$$\begin{array}{ccccc} \varepsilon: & 0 & 0 & 0 & 0 \\ & & 1 & 1 & \\ \delta: & & 1 & 1 & \end{array}$$

Second case: $\varepsilon_{m+n-1} = \varepsilon_{m+n+2} = 1$. Then $\delta_{m+n-1} = \delta_{m+n} = \delta_{m+n+1} = \delta_{m+n+2} = 1$ by (3.6). But this is impossible by Lemma 4.6 because $m + n - 1 \geq l + 1$.

$$\begin{array}{ccccc} \varepsilon: & 1 & 0 & 0 & 1 \\ & & 1 & 1 & \\ \delta: & 1 & 1 & 1 & 1 \end{array}$$

Third case: $\varepsilon_{m+n-1}=0$ and $\varepsilon_{m+n+2}=1$. Then $\varepsilon_{m+n+3}=1$ because (3.8) implies that $\varepsilon_{m+n-1}=\varepsilon_{m+n}=\varepsilon_{m+n+1}=0$ must be followed by $\varepsilon_{m+n+2}=\varepsilon_{m+n+3}=1$. (We use again the relation $m+n-1\geq k+1$.) But then $\delta_{m+n}=\delta_{m+n+1}=\delta_{m+n+2}=\delta_{m+n+3}=1$, contradicting Lemma 4.6 again (because $m+n\geq l+1$).

$$\begin{array}{cccccc} \varepsilon: & 0 & 0 & 0 & 1 & 1\\ & & 1 & 1 & & \\ \delta: & & 1 & 1 & 1 & 1 \end{array}$$

Fourth case: $\varepsilon_{m+n-1}=1$ and $\varepsilon_{m+n+2}=0$. Then $\delta_{m+n-1}=\delta_{m+n}=\delta_{m+n+1}=1$, and applying (3.9) we conclude that $\delta_{m+n+2}=\delta_{m+n+3}=0$. (To apply (3.9) we need again the relation $m+n-1\geq l+1$.) This implies by (3.6) that $\varepsilon_{m+n}=\varepsilon_{m+n+1}=\varepsilon_{m+n+2}=\varepsilon_{m+n+3}=0$, contradicting Lemma 4.5 because $m+n\geq k+1$.

$$\begin{array}{cccccc} \varepsilon: & 1 & 0 & 0 & 0 & 0\\ & & 1 & 1 & & \\ \delta: & 1 & 1 & 1 & 0 & 0 \end{array}$$

□

Lemma 4.8. *There is no* 000 *in the subsequence*

$$\varepsilon_L\varepsilon_{L+1}\ldots,\qquad L=\max\{k+1,l-m-1\}.$$

Proof. Assume on the contrary that $\varepsilon_n=\varepsilon_{n+1}=\varepsilon_{n+2}=0$ for some $n\geq L$. Then $\varepsilon_{n+3}=\varepsilon_{n+4}=1$ by (3.8) (because $n\geq k+1$), contradicting the preceding lemma (because $n+3\geq K$). □

Lemma 4.9. *There is no* 101 *in the subsequence*

$$\varepsilon_M\varepsilon_{M+1}\ldots,\qquad M=\max\{m+k+1,l+1\}.$$

Proof. (See the figure below.) Assume on the contrary that $\varepsilon_n\varepsilon_{n+1}\varepsilon_{n+2} = 101$ for some $n \geq M$. Then we have $\delta_n = \delta_{n+2} = 1$ and $\varepsilon_{n-m} = \varepsilon_{n+2-m} = 0$ by (3.6). Then $\varepsilon_{n+1-m} = 1$ by Lemma 4.8 (because $n - m \geq L$), and then $\delta_{n+1} = 1$ by (3.6). Since $\delta_n = \delta_{n+1} = \delta_{n+2} = 1$ and $n \geq l + 1$, (3.9) implies that $\delta_{n+3} = \delta_{n+4} = 0$, and therefore $\varepsilon_{n+3-m} = \varepsilon_{n+4-m} = 0$ by (3.6). But then $\varepsilon_{n+2-m} = \varepsilon_{n+3-m} = \varepsilon_{n+4-m} = 0$, contradicting the preceding lemma (because $n + 2 - m \geq L$).

$$\begin{array}{lcccccccccc} \varepsilon: & 0 & 1 & 0 & 0 & 0 & \dots & 1 & 0 & 1 & & \\ \delta: & & & & & & \dots & 1 & 1 & 1 & 0 & 0 \end{array}$$

∎

Now we can prove the crucial

Proposition 4.10. *Every admissible sequence ε satisfying* (4.1) *begins in fact with* 11100 *and is eventually periodic with period* 100. *Moreover, denoting by ε_k the first one digit in the subsequence $\varepsilon_{m+1}\varepsilon_{m+2}\dots$ and by δ_l the first zero digit in the subsequence $\delta_{m+1}\delta_{m+2}\dots$, the subsequence*

$$\varepsilon_N\varepsilon_{N+1}\dots, \quad N = \max\{m + k + 1, l + 1\}$$

is already periodic with one of the three periods 100, 010 *or* 001.

Proof. By Lemma 4.4 every admissible sequence satisfying (4.1) begins in fact with 11100.

Next we note that $N = \max\{K, L, M\}$. Therefore the three preceding lemmas imply that the subsequence $\varepsilon_M\varepsilon_{M+1}\dots$ does not contain three consecutive zeroes and every one digit is followed by at least two zeroes. Hence this subsequence is periodic with one of the periods 100, 010 or 001. ∎

Before proceeding to the proof of Theorem 4.1, let us note two important corollaries of the above proposition.

Remark.
- There are only countably many admissible sequences satisfying (4.1).
- If ε is an admissible sequence beginning with 11100, then m cannot be a multiple of 3. Indeed, otherwise (3.6) would lead to an eventually periodic sequence δ with period 200.

PROOF OF THEOREM 4.1. It remains to prove that every admissible sequence (ε_n) satisfies the inequality

$$(\varepsilon_n) \geq 111\ 001\ 001\ 001\dots$$

We already know that (ε_n) begins with 111. Assume on the contrary that for some $k \geq 1$ the sequence (ε_n) begins with

$$\underbrace{111}_{1}\ \underbrace{001}_{2}\dots\ \underbrace{001}_{k}\ 000. \tag{4.2}$$

By (3.4) we cannot have $m = 1$ or $m = 2$, and by the last remark above m cannot be a multiple of 3. Since $\varepsilon_m = 1$ by (3.3), we must have $m \geq 3k + 4$.

Then the sequence (δ_n) also begins with (4.2). Applying (3.7) we conclude that the following three digits of (δ_n) are 111. Therefore $m \geq 3k+7$ (because $\delta_m = 0$), so that the following three digits of (ε_n) are also 111, and then by (3.1) it must be followed by 000. In particular, we must have $m \geq 3k + 10$.

Repeating these arguments we obtain that m must be bigger than any integer, which is impossible. □

5 – Algebraic properties of q-expansions

We have the

THEOREM 5.1. *If ε is an eventually periodic sequence, then $1 < q < 2$ given by* (3.2) *is irrational.*

PROOF. Assume first that ε is periodic from the beginning with a period of length m. Then setting $x = 1/q$ we have

$$\begin{aligned} 1 &= (\varepsilon_1 x^1 + \dots + \varepsilon_m x^m)(1 + x^m + x^{2m} + \dots) = \\ &= (\varepsilon_1 x^1 + \dots + \varepsilon_m x^m)/(1 - x^m), \end{aligned}$$

whence

$$1 - x^m = \varepsilon_1 x^1 + \dots + \varepsilon_m x^m$$

or

$$(1+\varepsilon_m)x^m+\varepsilon_{m-1}x^{m-1}+\dots+\varepsilon_1 x-1=0.$$

Assuming on the contrary that q is rational, we may write $x=a/b$ with two relative prime integers a and b. Since $0.5<x<1$, we have $a\ge 2$. Rewriting the equation in the form

$$(1+\varepsilon_m)a^m+\varepsilon_{m-1}a^{m-1}b+\dots+\varepsilon_1 ab^{m-1}-b^m=0,$$

we see that all terms but the last are divisible by a. (We need here the fact that $a\ge 2$.) Hence the equality cannot hold.

Now assume that ε is periodic with a period of length m, but not from the beginning. Then there is a positive integer n such that the subsequence $\varepsilon_{n+1}\varepsilon_{n+2}\dots$ is periodic with a period of length m, but $\varepsilon_n\ne\varepsilon_{n+m}$. Setting again $x=1/q$, we have

$$1=\varepsilon_1x+\dots+\varepsilon_nx^n+(\varepsilon_{n+1}x^{n+1}+\dots+\varepsilon_{n+m}x^{n+m})(1+x^m+x^{2m}+\dots)=$$
$$=\varepsilon_1x+\dots+\varepsilon_nx^n+(\varepsilon_{n+1}x^{n+1}+\dots+\varepsilon_{n+m}x^{n+m})/(1-x^m)$$

whence

$$1-x^m=(\varepsilon_1x+\dots+\varepsilon_nx^n)(1-x^m)+(\varepsilon_{n+1}x^{n+1}+\dots+\varepsilon_{n+m}x^{n+m}).$$

Since $\varepsilon_n\ne\varepsilon_{n+m}$, this is an equation of order $n+m$ with integer coefficients and with principal coefficient $\varepsilon_n-\varepsilon_{n+m}=\pm 1$. Now we recall the elementary result from algebra that if a real number x satisfies an equation $x^n+c_1x^{n-1}+\dots+c_n$ with integral coefficients, then x is either an integer or an irrational number. (The proof is analogous to the usual proof of the irrationality of $\sqrt{2}$, see, e.g., [9], p. 15.)

Since in our case $x=1/q$ satisfies $0<x<1$, it cannot be an integer; applying this theorem we conclude that x is irrational. □

REFERENCES

[1] J.-P. Allouche – M. Cosnard: *Non-integer bases, iteration of continuous functions, and an arithmetical self-similar set*, to appear.

[2] P. Erdös – M. Horváth – I. Joó: *On the uniqueness of the expansions* $1=\sum q^{-n_i}$, Acta Math. Hungar., **58** (1991), 333-342.

[3] P. Erdös – I. Joó – V. Komornik: *Characterization of the unique expansions* $1 = \sum q^{-n_i}$ *and related problems*, Bull. Soc. Math. France, **118** (1990), 377-390.

[4] P. Erdös – I. Joó: *On the number of expansions* $1 = \sum q^{-n_i}$, Ann. Univ. Sci. Budapest. Eötvös Sect. Math., (35) 1992, 129-132.

[5] P. Erdös – I. Joó – V. Komornik: *On the number of q-expansions*, Ann. Univ. Sci. Budapest. Eötvös Sect. Math., (37) 1994, 109-118.

[6] P. Erdös – P. Révész: *On the length of the longest head-run*, Colloq. Math. Soc. J. Bolyai, **16** (1977), 219-228.

[7] I. Joó – F. J. Schnitzer: *On some problems concerning expansions by noninteger bases*, Anz. Österreich. Akad. Wiss. Math.-Natur. Kl., **133** (1996), 3-10 (1997).

[8] V. Komornik – P. Loreti: *Unique developments in non-integer bases*, Amer. Math. Monthly, **105** (1998), 636-639.

[9] I. Niven: *Irrational Numbers*, The Mathematical Associaton of America, 1956.

[10] W. Parry: *On the β-expansion of real numbers*, Acta Math. Hungar., **11** (1960), 401-416.

[11] A. Rényi: *Representations for real numbers and their ergodic properties*, Acta Math. Hungar., **8** (1957), 477-493.

Lavoro pervenuto alla redazione il 9 luglio 1999
ed accettato per la pubblicazione il 23 febbraio 2000.
Bozze licenziate il 19 aprile 2000

Indirizzo degli Autori:

V. Komornik – Institut de Recherche Mathématique Avancée – Université Louis Pasteur et CNRS – 7, rue René Descartes – 67084 Strasbourg Cédex, France
E-mail: komornik@math.u-strasbg.fr

P. Loreti – Dipartimento Me. Mo. Mat. – Università di Roma “La Sapienza” – Via A. Scarpa, 16 – 00161 Roma, Italy
Email: loreti@dmmm.uniroma1.it