Univoque numbers and an avatar of Thue-Morse

Jean-Paul Allouche
CNRS, LRI, UMR 8623
Université Paris-Sud, Bâtiment 490
F-91405 Orsay Cedex (France)
allouche@lri.fr

Christiane Frougny
LIAFA, CNRS UMR 7089
Case 7014
75205 Paris Cedex 13 (France)
and Université Paris 8
Christiane.Frougny@liafa.jussieu.fr

Abstract

Univoque numbers are real numbers $\lambda > 1$ such that the number 1 admits a unique expansion in base λ, i.e., a unique expansion $1 = \sum_{j \geq 0} a_j \lambda^{-(j+1)}$, with $a_j \in \{0, 1, \ldots, \lceil \lambda \rceil - 1\}$ for every $j \geq 0$. A variation of this definition was studied in 2002 by Komornik and Loreti, together with sequences called *admissible sequences*. We show how a 1983 study of the first author gives both a result of Komornik and Loreti on the smallest admissible sequence on the set $\{0, 1, \ldots, b\}$, and a result of de Vries and Komornik (2007) on the smallest univoque number belonging to the interval $(b, b+1)$, where b is any positive integer. We also prove that this last number is transcendental. An avatar of the Thue-Morse sequence, namely the fixed point beginning in 3 of the morphism $3 \to 31$, $2 \to 30$, $1 \to 03$, $0 \to 02$, occurs in a "universal" manner.

2000 Math. Subject Classification. 11A63, 11B83, 11B85, 68R15, 11J81.
Keywords. Beta-expansion, univoque numbers, iteration of continuous functions, Thue-Morse sequence, uniform morphism, automatic sequence, transcendence.

1 Introduction

Komornik and Loreti determined in [18] the smallest *univoque* real number in the interval $(1, 2)$, i.e., the smallest number $\lambda \in (1, 2)$ such that 1 has a unique expansion $1 = \sum_{j \geq 0} a_j/\lambda^{j+1}$ with $a_j \in \{0, 1\}$ for every $j \geq 0$. The word "univoque" in this context seems to have been introduced (with a slightly different meaning) by Daróczy and Kátai in [12, 13], while characterizing unique expansions of the real number 1 was done by Erdős, Joó, and Komornik in [15]. The first author and Cosnard showed in [4] how the result of [18] parallels (and can be deduced from) their study of a certain set of binary sequences arising in the iteration of unimodal continous functions of the unit interval that was done in [11, 2, 1]. The relevant sets of binary sequences occurring in references [2, 1], resp. in reference [18], can be defined by

$$\Gamma := \{A \in \{0,1\}^{\mathbb{N}},\ \forall k \geq 0,\ \overline{A} \leq \sigma^k A \leq A\}$$

$$\Gamma_{strict} := \{A \in \{0,1\}^{\mathbb{N}},\ \forall k \geq 1,\ \overline{A} < \sigma^k A < A\}$$

where σ is the *shift* on sequences and the *bar* operation replaces 0's by 1's and 1's by 0's, i.e., if $A = (A_n)_{n \geq 0}$, then $\sigma A := (a_{n+1})_{n \geq 0}$, and $\overline{A} := (1 - a_n)_{n \geq 0}$; furthermore $\leq$ denotes the lexicographical order on sequences induced by $0 < 1$, the notation $A < B$ meaning as usual that $A \leq B$ and $A \neq B$. The smallest univoque number in $(1, 2)$ and the smallest nonperiodic sequence of the set Γ both involve the Thue-Morse sequence (see for example [6] for more on this sequence).

It is tempting to generalize these sets to alphabets with more than 2 letters.

Definition 1 For b a positive integer, we will say that the real number $\lambda > 1$ is $\{0, 1, \ldots, b\}$-*univoque* if the number 1 has a unique expansion as $1 = \sum_{j \geq 0} a_j \lambda^{-(j+1)}$, where a_j belongs to $\{0, \ldots, b\}$ for all $j \geq 0$. Furthermore, if $\lambda > 1$ is $\{0, 1, \ldots, \lceil \lambda \rceil - 1\}$-univoque, we will simply say that λ is *univoque.*

Remark 1 If $\lambda > 1$ is $\{0, 1, \ldots, b\}$-univoque for some positive integer b, then $\lambda \leq b + 1$. Also note that any integer $q \geq 2$ is univoque, since there is exactly one expansion of 1 as $1 = \sum_{j \geq 0} a_j q^{-(j+1)}$, with $a_j \in \{0, 1, \ldots, q-1\}$, namely $1 = \sum_{j \geq 0} (q-1) q^{-(j+1)}$.

Komornik and Loreti studied in [19] the reals λ belonging to the interval $(1, b+1]$ that are $\{0, 1, \ldots, b\}$-univoque. For their study, they introduced *admissible sequences* on the alphabet $\{0, 1, \ldots, b\}$. Denote, as above, by σ the *shift* on sequences, and by *bar* the operation that replaces every $t \in \{0, 1, \ldots, b\}$ by $b - t$, i.e., if $A = (a_n)_{n \geq 0}$, then $\overline{A} := (b - a_n)_{n \geq 0}$. Also denote by $\leq$ the lexicographical order on sequences induced by the natural order on $\{0, 1, \ldots, b\}$. Then, a sequence $A = (a_n)_{n \geq 0}$ on $\{0, 1, \ldots, b\}$ is *admissible* if

$$\begin{array}{lll} \forall k \geq 0 \text{ such that } a_k < b, & \sigma^{k+1}A & < & A, \\ \forall k \geq 0 \text{ such that } a_k > 0, & \sigma^{k+1}A & > & \overline{A}. \end{array}$$

(Note that our notation is not exactly the notation of [19], since our sequences are indexed by $\mathbb{N}$ and not $\mathbb{N} \setminus \{0\}$.) These sequences have the following property: the map that associates with a real $\lambda \in (1, b+1]$ the sequence of coefficients $(a_j)_{j \geq 0} \in \{0, 1, \ldots, b\}$ of the greedy (i.e., the lexicographically largest) expansion of 1, $1 = \sum_{j \geq 0} a_j \lambda^{-(j+1)}$, is a bijection from the set of $\{0, 1, \ldots, b\}$-univoque λ's to the set of admissible sequences on $\{0, 1, \ldots, b\}$ (see [19]).

Now there are two possible generalizations of the result of [18] about the smallest univoque number in $(1, 2)$, i.e., the smallest admissible binary sequence. One is to look at the smallest (if any) admissible sequence on the alphabet $\{0, 1, \ldots, b\}$, as did Komornik and Loreti in [19], the other is to look at the smallest (if any) univoque number in $(b, b+1)$, as did de Vries and Komornik in [14].

It happens that the first author already studied a generalization of the set Γ to the case of more than 2 letters (see [1, Part 3]). Interestingly enough this study was not related to the iteration of continuous functions as was the study of Γ, but only introduced as a tempting formal arithmetico-combinatorial generalization of the study of the set of binary sequences Γ to a similar set of sequences with more than two values.

The purpose of the present paper is threefold:

1) *to show how the 1983 study [1, Part 3, p. 63–90] gives both the result of Komornik and Loreti in [19] on the smallest admissible sequence on $\{0, 1, \ldots, b\}$, and the result of de Vries and Komornik in [14] on the smallest number univoque number λ belonging to $(b, b+1)$ where b is any positive integer;*

2) *to bring to light a* universal *morphism that governs the smallest elements in 1) above, and to show that the infinite sequence generated by this morphism is an avatar of the Thue-Morse sequence;*

3) *to prove that the smallest univoque number belonging to $(b, b+1)$ (where b is any positive integer) is transcendental.*

The paper consists of five sections. In Section 2 below we recall some results of [1, Part 3, p. 63–90] on the generalization of the set Γ to a $(b+1)$-letter alphabet. Then we give some properties of the lexicographically least nonperiodic sequence of this set, completing results of [1, Part 3, p. 63–90]. In Section 3 we give two corollaries of the properties of this least sequence: one gives the result in [19], the other gives the result in [14]. The transcendence results are proven in the last section.

2 The generalized Γ and Γ_{strict} sets

Definition 2 Let b be a positive integer, and $\mathcal{A}$ be a finite ordered set with $b+1$ elements. Let $\alpha_0 < \alpha_1 < \ldots < \alpha_b$ be the elements of $\mathcal{A}$. The *bar* operation is defined on $\mathcal{A}$ by $\overline{\alpha_j} = \alpha_{b-j}$. We extend this operation to $\mathcal{A}^{\mathbb{N}}$ by $\overline{(a_n)_{n\geq 0}} := (\overline{a_n})_{n\geq 0}$. Let σ be the *shift* on $\mathcal{A}^{\mathbb{N}}$, defined by $\sigma((a_n)_{n\geq 0}) := (a_{n+1})_{n\geq 0}$.

We define the sets $\Gamma(\mathcal{A})$ and $\Gamma_{strict}(\mathcal{A})$ by:

$$\Gamma(\mathcal{A}) := \{A = (a_n)_{n\geq 0} \in \mathcal{A}^{\mathbb{N}},\ a_0 = \max \mathcal{A},\ \forall k \geq 0,\ \overline{A} \leq \sigma^k A \leq A\},$$

$$\Gamma_{strict}(\mathcal{A}) := \{A = (a_n)_{n\geq 0} \in \mathcal{A}^{\mathbb{N}},\ a_0 = \max \mathcal{A},\ \forall k \geq 1,\ \overline{A} < \sigma^k A < A\}.$$

Remark 2 The set $\Gamma(\mathcal{A})$ was introduced by the first author in [1, Part 3, p. 63–90]. Note that there is a misprint in the definition given on p. 66 in [1]: $a_{\beta-i}$ should be changed into $a_{\beta-1-i}$ as confirmed by the rest of the text.

Remark 3 A sequence belongs to $\Gamma_{strict}(\mathcal{A})$ if and only if it belongs to $\Gamma(\mathcal{A})$ and is nonperiodic. Namely, $\sigma^k A = A$ if and only if A is k-periodic; if $\sigma^k A = \overline{A}$, then $\sigma^{2k} A = A$, and the sequence A is $2k$-periodic.

Remark 4 If the set $\mathcal{A}$ is given by $\mathcal{A} := \{i, i+1, \ldots, i+z\}$ where i and z are integers, equipped with the natural order, then for any $x \in \mathcal{A}$, we have $\overline{x} = 2i + z - x$. Namely, following Definition 2 above, we write $\alpha_0 := i, \alpha_1 := i+1, \ldots, \alpha_z := i+z$. Hence, for any $j \in [0, z]$, we have $\overline{\alpha_j} = \alpha_{z-j}$, which can be rewritten $\overline{i+j} = i + z - j$, i.e., for any x in $\mathcal{A}$, we have $\overline{x} = i + z - (x - i) = 2i + z - x$.

A first result is that the sets $\Gamma_{strict}(\mathcal{A})$ are closely linked to the set of admissible sequences whose definition was recalled in the introduction.

Proposition 1 *Let $A = (a_n)_{n\geq 0}$ be a sequence in $\{0, 1, \ldots, b\}^{\mathbb{N}}$, such that $a_0 = t \in [0, b]$. Suppose that the sequence A is not equal to $b\ b\ b\ \ldots$ Then the sequence A is admissible if and only if $2t > b$ and A belongs to the set $\Gamma_{strict}(\{b-t, b-t+1, \ldots, t\})$. (The order on $\{b-t, b-t+1, \ldots, t\}$ is induced by the order on $\mathbb{N}$. From Remark 4 the bar operation is given by $\overline{j} = b - j$.)*

Proof. Let $A = (a_n)_{n \geq 0}$ be a sequence belonging to $\{0, 1, \ldots, b\}^{\mathbb{N}}$ such that $a_0 = t \in [0, b-1]$, and such that $A \neq b\ b\ b \ldots$

$*$ First suppose that $2t > b$ and A belongs to $\Gamma_{strict}(\{b-t, b-t+1, \ldots, t\})$. Then, for all $k \geq 1$, $\overline{A} < \sigma^k A < A$, which clearly implies that A is admissible.

$*$ Now suppose that A is admissible. We thus have

$$\begin{array}{lrcl} \forall k \geq 1 \text{ such that } a_{k-1} < b, & \sigma^k A & < & A, \\ \forall k \geq 1 \text{ such that } a_{k-1} > 0, & \sigma^k A & > & \overline{A}. \end{array}$$

We first prove that, if the sequence A is not a constant sequence, then

$$\forall k \geq 1, \ \overline{A} < \sigma^k A < A.$$

We only prove the inequalities $\sigma^k A < A$. The remaining inequalities are proved in a similar way. If $a_{k-1} < b$, the inequality $\sigma^k A < A$ holds. If $a_{k-1} = b$, there are two cases:

- either $a_0 = a_1 = \ldots = a_{k-1} = b$, then, if $a_k < b$ we clearly have $\sigma^k A < A$; if $a_k = b$, then the sequence $\sigma^k A$ begins with some block of b's followed by a letter $< b$, thus it begins with a block of b's shorter than the initial block of b's of the sequence A itself, hence $\sigma^k A < A$;

- or there exists an index ℓ with $1 < \ell < k$, such that $a_{\ell-1} < b$, and $a_\ell = a_{\ell+1} = \ldots = a_{k-1} = b$. As A is admissible, we have $\sigma^\ell A < A$. It thus suffices to prove that $\sigma^k A \leq \sigma^\ell A$. This is clearly the case if $a_k < b$. On the other hand, if $a_k = b$, the sequence $\sigma^k A$ begins with a block of b's which is shorter than the initial block of b's of the sequence $\sigma^\ell A$, hence $\sigma^k A \leq \sigma^\ell A$.

Now, since $a_0 = t$ and $\sigma^k A < A$ for all $k \geq 1$, we have $a_k \leq t$ for all $k \geq 0$. Similarly, since $\sigma^k A > \overline{A}$ for all $k \geq 1$, we have $a_k \geq b - t$ for all $k \geq 1$. Finally $A > \overline{A}$ implies that $t = a_0 \geq b - t$. We thus have that $2t \geq b$ and A belongs to $\Gamma(\{b-t, b-t+1, \ldots, t\})$. Now, if $b = 2t$, then $\{b-t, b-t+1, ..., t\} = \{t\}$ and $\bar{t} = t$. This implies that $A = t\ t\ t\ \ldots$, which is not an admissible sequence. $\square$

Remark 5 For $b = 1$, this (easy) result is noted without proof in [15] and proved in [4].

We need another definition from [1].

Definition 3 Let b be a positive integer, and $\mathcal{A}$ be a finite ordered set with $b + 1$ elements. Let $\alpha_0 < \alpha_1 < \ldots < \alpha_b$ be the elements of $\mathcal{A}$. We suppose that $\mathcal{A}$ is equipped with a bar operation as in Definition 2. Let $A = (a_n)_{n\geq 0}$ be a periodic sequence of *smallest* period T, and such $a_{T-1} < \max \mathcal{A}$. Let $a_{T-1} = \alpha_j$ (thus $j < b$). Then the sequence $\Phi(A)$ is defined as the $2T$-periodic sequence beginning with $a_0\ a_1\ \ldots\ a_{T-2}\ \alpha_{j+1}\ \overline{a_0}\ \overline{a_1}\ \ldots\ \overline{a_{T-2}}\ \alpha_{b-j-1}$, i.e.,

$$\Phi((a_0\ a_1\ \ldots\ a_{T-2}\ \alpha_j)^\infty := (a_0\ a_1\ \ldots\ a_{T-2}\ \alpha_{j+1}\ \overline{a_0}\ \overline{a_1}\ \ldots\ \overline{a_{T-2}}\ \alpha_{b-j-1})^\infty.$$

We first prove the following easy claim.

Proposition 2 *The smallest element of* $\Gamma(\{b-t, b-t+1, \ldots, t\})$ *(where* $2t > b$*) is the* 2*-periodic sequence* $(t\ (b-t))^\infty = (t\ (b-t)\ t\ (b-t)\ t\ \ldots)$*.*

Proof. Since any sequence $A = (a_n)_{n\geq 0}$ belonging to $\Gamma(\{b-t, b-t+1, \ldots, t\})$ begins in t, and satisfies $\sigma A \geq \overline{A}$, then it must satisfy $a_0 = t$ and $a_1 \geq b-t$. Now if a sequence A belonging to $\Gamma(\{b-t, b-t+1, \ldots, t\})$ is such that $a_0 = t$ and $a_1 = (b-t)$, then it must be equal to the 2-periodic sequence $(t\ (b-t))^\infty$ ([1, Lemma 2, b, p. 73]). Since this periodic sequence trivially belongs to $\Gamma(\{b-t, b-t+1, \ldots, t\})$, it is its smallest element. $\square$

Denoting as usual by Φ^s the s-th iterate of Φ, we state the following theorem which is a particular case of the theorem on pages 72–73 of [1] about the smallest elements in certain subintervals of $\Gamma(\{0, 1, \ldots, b\})$, and of the definition of q-mirror sequences given in [1, Section II, 1, p. 67] (here $q = 2$).

Theorem 1 ([1]) *Define* $P := (t\ (b-t))^\infty = (t\ (b-t)\ t\ (b-t)\ t\ \ldots)$*. The smallest nonperiodic sequence in the set* $\Gamma(\{b-t, b-t+1, \ldots, t\})$ *(i.e., the smallest element of* $\Gamma_{strict}(\{b-t, b-t+1, \ldots, t\})$*) is the sequence* M *defined by*

$$M := \lim_{s\to\infty} \Phi^s(P),$$

that actually takes the (not necessarily distinct) values $b-t$*,* $b-t+1$*,* $t-1$*,* t*. Furthermore, this sequence* $M = (m_n)_{n\geq 0} = t\quad b-t+1\quad b-t\quad t\quad b-t\quad t-1\ \ldots$ *can be recursively defined by*

$$\begin{aligned} &\forall k \geq 0,\ m_{2^{2k}-1} = t, \\ &\forall k \geq 0,\ m_{2^{2k+1}-1} = b+1-t, \\ &\forall k \geq 0,\ \forall j \in [0, 2^{k+1}-2],\ m_{2^{k+1}+j} = \overline{m_j}. \end{aligned}$$

It was proven in [1] that the sequence $\lim_{s\to\infty} \Phi^s((t\ (b-t))^\infty$ is 2-automatic (for more about automatic sequences, see [7]). The second author noted that this sequence is actually a fixed point of a uniform morphism of length 2 as soon as the cardinality of the set $\{b-t, b-t+1, \ldots, b\}$ is at least equal to 4, i.e., $2t \geq b+3$. (Recall that we always have $t \geq b-t$, i.e., $2t \geq b$.) More precisely we have Theorem 2 below, where the Thue-Morse sequence pops up, as in [1] and in [19], but also as in [2] and [18]. Before stating this theorem we give a definition.

Definition 4 The "universal" morphism Θ is defined on $\{e_0, e_1, e_2, e_3\}$ by

$$\Theta(e_3) := e_3e_1,\ \Theta(e_2) := e_3e_0,\ \Theta(e_1) := e_0e_3,\ \Theta(e_0) := e_0e_2.$$

Note that this morphism has an infinite fixed point beginning in e_3

$$\Theta^\infty(e_3) = \lim_{k\to\infty} \Theta^k(e_3) = e_3\ e_1\ e_0\ e_3\ e_0\ e_2\ e_3\ e_1\ e_0\ e_2 \ldots.$$

Theorem 2 *Let $(\varepsilon_n)_{n\geq 0}$ be the Thue-Morse sequence, defined by $\varepsilon_0 = 0$ and for all $k \geq 0$, $\varepsilon_{2k} = \varepsilon_k$ and $\varepsilon_{2k+1} = 1 - \varepsilon_k$. Then the smallest nonperiodic sequence $M = (m_n)_{n\geq 0}$ belonging to $\Gamma(\{b-t, b-t+1, \ldots, t\})$ satisfies*

$$\forall n \geq 0,\ m_n = \varepsilon_{n+1} - (2t-b-1)\varepsilon_n + t - 1.$$

Using the morphism Θ introduced in Definition 4 above we thus have

- *if $2t \geq b+3$, then the sequence M is the fixed point beginning in t of the morphism deduced from Θ by renaming e_0, e_1, e_2, e_3 respectively $b-t, b-t+1, t-1, t$ (note that the condition $2t \geq b+3$ implies that these four numbers are distinct);*

- *if $2t = b+2$ (thus $b-t+1 = t-1$), then the sequence M is the pointwise image of the fixed point beginning in e_3 of the morphism Θ by the map g defined by $g(e_3) := t$, $g(e_2) = g(e_1) := t-1$, $g(e_0) := b-t$;*

- *if $2t = b+1$ (thus $b-t = t-1$ and $b-t+1 = t$), then the sequence M is the pointwise image of the fixed point beginning in e_3 of the morphism Θ by the map h defined by $h(e_3) = h(e_1) := t$, $h(e_2) = h(e_0) := t-1$.*

Proof. Let us first prove that the sequence $M = (m_n)_{n\geq 0}$ is equal to the sequence $(u_n)_{n\geq 0}$, where $u_n := \varepsilon_{n+1} - (2t-b-1)\varepsilon_n + t - 1$. It suffices to prove that the sequence $(u_n)_{n\geq 0}$ satisfies the recursive relations defining $(m_n)_{n\geq 0}$ that are given in Theorem 1. Recall that the sequence $(\varepsilon_n)_{n\geq 0}$ has the property that ε_n is equal to the parity of the sum of the binary digits of the integer n (see [6] for example). Hence, for all $k \geq 0$, $\varepsilon_{2^{2k}-1} = 0$, $\varepsilon_{2^{2k+1}-1} = 1$, and $\varepsilon_{2^{2k}} = \varepsilon_{2^{2k+1}} = 1$. This implies that for all $k \geq 0$, $u_{2^{2k}-1} = t$ and $u_{2^{2k+1}-1} = b+1-t$. Furthermore, for all $k \geq 0$, and for all $j \in [0, 2^{k+1}-2]$, we have $\varepsilon_{2^{k+1}+j} = 1 - \varepsilon_j$ and $\varepsilon_{2^{k+1}+j+1} = 1 - \varepsilon_{j+1}$. Hence $u_{2^{k+1}+j} = b - u_j = \overline{u_j}$.

To show how the "universal" morphism Θ enters the picture, we study the sequence $(v_n)_{n\geq 0}$ with values in $\{0,1\}^2$ defined by: for all $n \geq 0$, $v_n := (\varepsilon_n, \varepsilon_{n+1})$. Since we have, for all $n \geq 0$, $v_{2n} = (\varepsilon_n, 1-\varepsilon_n)$ and $v_{2n+1} = (1-\varepsilon_n, \varepsilon_{n+1})$, we clearly have

$$\begin{array}{llll}
\text{if} & v_n = (0,0), \text{ then} & v_{2n} = (0,1) \text{ and} & v_{2n+1} = (1,0), \\
\text{if} & v_n = (0,1), \text{ then} & v_{2n} = (0,1) \text{ and} & v_{2n+1} = (1,1), \\
\text{if} & v_n = (1,0), \text{ then} & v_{2n} = (1,0) \text{ and} & v_{2n+1} = (0,0), \\
\text{if} & v_n = (1,1), \text{ then} & v_{2n} = (1,0) \text{ and} & v_{2n+1} = (0,1).
\end{array}$$

This exactly means that the sequence $(v_n)_{n\geq 0}$ is the fixed point beginning in $(0,1)$ of the 2-morphism

$$\begin{array}{lcl}
(0,0) & \to & (0,1)(1,0) \\
(0,1) & \to & (0,1)(1,1) \\
(1,0) & \to & (1,0)(0,0) \\
(1,1) & \to & (1,0)(0,1)
\end{array}$$

We may define $e_0 := (1,0)$, $e_1 := (1,1)$, $e_2 := (0,0)$, $e_3 := (0,1)$. Then the above morphism can be written

$$e_3 \to e_3 e_1, \;\; e_2 \to e_3 e_0, \;\; e_1 \to e_0 e_3, \;\; e_0 \to e_0 e_2$$

which is the morphism Θ. The above construction shows that the sequence $(v_n)_{n\geq 0}$ is a fixed point of Θ.

Now, define the map ω on $\{0,1\}^2$ by

$$\omega((x,y)) := y - (2t-b-1)x + t - 1.$$

We have $\omega(v_n) = m_n$ for all $n \geq 0$. Thus

- if $2t \geq b+3$, the sequence $(m_n)_{n\geq 0}$ takes exactly four distinct values, namely $b-t, b-t+1, t-1, t$. This implies that $(m_n)_{n\geq 0}$ is the fixed point beginning in t of the morphism obtained from Θ by renaming the letters, i.e., $e_3 \to t$, $e_2 \to (t-1)$, $e_1 \to (b-t+1)$, $e_0 \to (b-t)$. The morphism can thus be writen $t \to t\ (b-t+1)$, $(t-1) \to t\ (b-t)$, $(b-t+1) \to (b-t)\ t$, $(b-t) \to (b-t)\ (t-1)$;

- if $2t = b+2$ (resp. $2t = b+1$) the sequence $(m_n)_{n\geq 0}$ takes exactly three (resp. two) values, namely $b-t, t-1, t$ (resp. $t-1, t$). It is still the pointwise image by Θ of the sequence $(v_n)_{n\geq 0}$. Renaming Θ as g (resp. h) as in the statement of Theorem 2 only takes into account that the integers $b-t, b-t+1, t-1, t$ are not distinct. $\square$

Remark 6 The reason for the choice of indexes for e_3, e_2, e_1, e_0 is that the order of indexes is the same as the natural order on the integers $t, t-1, b-t+1, b-t$ to which they correspond when $2t \geq b+3$. In particular if $b = t = 3$, the morphism reads: $3 \to 31$, $2 \to 30$, $1 \to 03$, $0 \to 02$. Interestingly enough, though not surprisingly, this morphism also occurs (up to renaming once more the letters) in the study of infinite square-free sequences on a 3-letter alphabet. Namely, in the paper [9], Berstel proves that the square-free Istrail sequence [16], originally defined (with no mention of the Thue-Morse sequence) as the fixed point of the (non-uniform) morphism $0 \to 12$, $1 \to 102$, $2 \to 0$, is actually the pointwise image of the fixed point beginning in 1 of a 2-morphism Θ' on the 4-letter alphabet $\{0, 1, 2, 3\}$ by the map $0 \to 0, 1 \to 1, 2 \to 2, 3 \to 0$. The morphism Θ' is given by

$$\Theta'(0) = 12,\ \Theta'(1) = 13,\ \Theta'(2) = 20,\ \Theta'(3) = 21.$$

The reader will note immediately that Θ' is another avatar of Θ obtained by renaming letters as follows: $0 \to 2, 1 \to 3, 2 \to 0, 3 \to 1$. This, in particular, shows that *the sequence $(m_n)_{n\geq 0}$, in the case where $2t = b + 2$, is the fixed point of the non-uniform morphism $t \to t\ (t-1)\ (b-t)$, $(t-1) \to t\ (b-t)$, $(b-t) \to (t-1)$, i.e., an avatar of Istrail's square-free sequence. Furthermore it results from [9] that this sequence on three letters cannot be the fixed point of a uniform morphism.* A last remark is that the square-free Braunholtz sequence on three letters given in [10] (see also [9, p. 18-07]) is exactly our sequence $(m_n)_{n\geq 0}$ when $t = b = 2$, i.e., the sequence 2 1 0 2 0 1 2 1 0 1 2 0 ...

3 Small admissible sequences and small univoque numbers with given integer part

3.1 Small admissible sequences with values in the set $\{0, 1, \ldots, b\}$

In [19] the authors are interested in the smallest admissible sequence with values in the set $\{0, 1, \ldots, b\}$, where b is an integer ≥ 1. They prove in particular the following result, which is an immediate corollary of our Theorem 2.

Corollary 1 (Theorems 4.3 and 5.1 of [19]) *Let b be an integer ≥ 1. The smallest admissible sequence with values in $\{0, 1, \ldots, b\}$ is the sequence $(z + \varepsilon_{n+1})_{n\geq 0}$ if $b = 2z + 1$, and $(z + \varepsilon_{n+1} - \varepsilon_n)_{n\geq 0}$ if $b = 2z$.*

Proof. Let $A = (a_n)_{n\geq 0}$ be the smallest (non-constant) admissible sequence with values in $\{0, 1, \ldots, b\}$. Since $A > \overline{A}$, we must have $a_0 \geq \overline{a_0} = b - a_0$.

Thus, if $b = 2z + 1$ we have $a_0 \geq z + 1$. We also have, for all $i \geq 0$, $\overline{a_0} \leq a_i \leq a_0$. Now the smallest element of the set $\Gamma(\{b - z - 1, b - z, \ldots, z - 1, z + 1\})$ is the smallest admissible sequence on $\{0, 1, \ldots, b\}$ that begins in $z + 1$. Hence this is the smallest admissible sequence with values in $\{0, 1, \ldots, b\}$. Theorem 2 gives that this sequence is $(m_n)_{n\geq 0}$ with, for all $n \geq 0$, $m_n = \varepsilon_{n+1} + z$.

If $b = 2z$, we have $a_0 \geq z$. But if $a_0 = z$, then $\overline{a_0} = z$, and the conditions of admissibility implies that $a_n = z$ for all $n \geq 0$ and $(a_n)_{n\geq 0}$ would be the constant sequence $(z\ z\ z \ldots)$. Hence we must have $a_0 \geq z + 1$. Now the smallest element of the set $\Gamma(\{b-z-1, b-z, \ldots, z-1, z+1\})$ is the smallest admissible sequence on $\{0, 1, \ldots, b\}$ that begins in $z + 1$. Hence this is the smallest admissible sequence with values in $\{0, 1, \ldots, b\}$. Theorem 2 gives that this sequence is $(m_n)_{n\geq 0}$ with, for all $n \geq 0$, $m_n = \varepsilon_{n+1} - \varepsilon_n + z$. □

3.2 Small univoque numbers with given integer part

We are interested here in the univoque numbers λ in an interval $(b, b + 1]$ with b a positive integer. This set was studied in [17], where it was proven of Lebesgue measure 0. Since $1 = \sum_{j\geq 0} a_j \lambda^{-(j+1)}$, $\lambda \in (b, b+1]$ and $a_0 \leq b$, the fact that the expansion of 1 is unique, hence equal to the greedy expansion, implies that $a_0 = b$. In other words, we study the admissible sequences with values in $\{0, 1, \ldots, b\}$ that begin in b, i.e., the set $\Gamma_{strict}(\{0, 1, \ldots, b\})$. We prove here, as a corollary of Theorem 2, that, for any positive integer b, there exists a smallest univoque number belonging to $(b, b + 1]$. This result was obtained in [14] (see the penultimate remark in that paper); it generalizes the result obtained for $b = 1$ in [18].

Corollary 2 *For any positive integer b, there exists a smallest univoque number in the interval $(b, b+1]$. This number is the solution of the equation $1 = \sum_{n\geq 0} d_n \lambda^{-n-1}$, where the sequence $(d_n)_{n\geq 0}$ is given by, for all $n \geq 0$, $d_n := \varepsilon_{n+1} - (b - 1)\varepsilon_n + b - 1$.*

Proof. It suffices to apply Theorem 2 with $t = b$. □

4 Transcendence results

We prove here, mimicking the proof given in [3], that numbers such that the expansion of 1 is given by the sequence $(m_n)_{n\geq 0}$ are transcendental. This generalizes the transcendence results of [3] and [19].

Theorem 3 *Let b be an integer ≥ 1 and $t \in [0, b]$ be an integer such that $2t \geq b+1$. Define the sequence $(m_n)_{n\geq 0}$ as in Theorem 2 by, for all $n \geq 0$, $m_n := \varepsilon_{n+1} - (2t-b-1)\varepsilon_n + t - 1$, thus the sequence $(m_n)_{n\geq 0}$ begins with $t \quad b-t+1 \quad b-t \quad t \quad b-t \quad t-1 \ \ldots$ Then the number λ belonging to $(1, b+1)$ defined by $1 = \sum_{n\geq 0} m_n \lambda^{-n-1}$ is transcendental.*

Proof. Define the ± 1 Thue-Morse sequence (r_n) by $r_n := (-1)^{\varepsilon_n}$. We clearly have $r_n = 1 - 2\varepsilon_n$ (recall that ε_n is 0 or 1). It is also immediate that the function F defined for the complex numbers X such that $|X| < 1$ by $F(X) = \sum_{n\geq 0} r_n X^n$ satisfies $F(X) = \prod_{k\geq 0}(1 - X^{2^k})$ (see, e.g., [6]). Since

$$2m_n = 2\varepsilon_{n+1} - 2(2t-b-1)\varepsilon_n + 2t - 2 = b - r_{n+1} + (2t-b-1)r_n$$

we have, for $|X| < 1$,

$$2X \sum_{n\geq 0} m_n X^n = ((2t-b-1)X - 1)F(X) + 1 + \frac{bX}{1-X}.$$

Taking $X = 1/\lambda$ where $1 = \sum_{n\geq 0} m_n \lambda^{-n-1}$, we get the equation

$$2 = ((2t-b-1)\lambda^{-1} - 1)F(1/\lambda) + 1 + \frac{b}{\lambda - 1}.$$

Now, if λ were algebraic, then this equation shows that $F(1/\lambda)$ would be an algebraic number. But, since $1/\lambda$ would be an algebraic number in $(0, 1)$, the quantity $F(1/\lambda)$ would be transcendental from a result of Mahler [20], giving a contradiction. □

Remark 7 In particular the $\{0, 1, \ldots, b\}$-univoque number corresponding to the smallest admissible sequence with values in $\{0, 1, \ldots, b\}$ is transcendental, as proved in [19] (Theorems 4.3 and 5.9). Also the smallest univoque number belonging to $(b, b+1)$ is transcendental.

5 Conclusion

There are many papers dealing with univoque numbers. We will just mention here the study of univoque Pisot numbers. The authors together with K. G. Hare determined in [5] the smallest univoque Pisot number, which happens to have algebraic degree 14. Note that the number corresponding to the sequence of Proposition 2 is the larger real root of the polynomial $X^2 - tX - (b-t+1)$, hence a Pisot number (which is unitary if $t = b$). Also note that for any $b \geq 2$, the real number β such that the β-expansion of 1 is $b1^\infty$ is a univoque Pisot number belonging to the interval $(b, b+1)$.

It would be interesting to determine the smallest univoque Pisot number belonging to $(b, b+1)$: the case $b = 1$ was addressed in [5], but the proof uses heavily the fine structure of Pisot numbers in the interval $(1, 2)$ (see [8, 21, 22]). A similar study of Pisot numbers in $(b, b+1)$ would certainly help.

Acknowledgments. The authors thank M. de Vries and V. Komornik for their remarks on a previous version of this paper.

References

[1] J.-P. Allouche, *Théorie des Nombres et Automates*, Thèse d'État, Bordeaux, 1983.

[2] J.-P. Allouche, M. Cosnard, Itérations de fonctions unimodales et suites engendrées par automates, *C. R. Acad. Sci. Paris, Sér. 1* **296** (1983) 159–162.

[3] J.-P. Allouche, M. Cosnard, The Komornik-Loreti constant is transcendental, *Amer. Math. Monthly* **107** (2000) 448–449.

[4] J.-P. Allouche, M. Cosnard, Non-integer bases, iteration of continuous real maps, and an arithmetic self-similar set, *Acta Math. Hung.* **91** (2001) 325–332.

[5] J.-P. Allouche, Ch. Frougny, K. Hare, On univoque Pisot numbers, *Math. Comp.* **76** (2007) 1639–1660.

[6] J.-P. Allouche, J. Shallit, The ubiquitous Prouhet-Thue-Morse sequence, in C. Ding, T. Helleseth, H. Niederreiter (Eds.), *Sequences and their applications, Proceedings of SETA'98*, Springer, 1999, pp. 1–16.

[7] J.-P. Allouche, J. Shallit, *Automatic sequences. Theory, applications, generalizations*, Cambridge University Press, Cambridge, 2003.

[8] M. Amara, Ensembles fermés de nombres algébriques, *Ann. Sci. École Norm. Sup.* **83** (1966) 215–270.

[9] J. Berstel, Sur la construction de mots sans carré, *Sém. Théor. Nombres, Bordeaux* (1978–1979), Exposé 18, 18-01–18-15.

[10] C. H. Braunholtz, An infinite sequence on 3 symbols with no adjacent repeats, (Solution to Problem 439 posed by H. Noland), *Amer. Math. Monthly* **70** (1963) 675–676.

[11] M. Cosnard, Étude de la classification topologique des fonctions unimodales, *Ann. Inst. Fourier* **35** (1985) 59–77.

[12] Z. Daróczy, I. Kátai, Univoque sequences, *Publ. Math. Debrecen* **42** (1993) 397–407.

[13] Z. Daróczy, I. Kátai, On the structure of univoque numbers, *Publ. Math. Debrecen* **46** (1995) 385–408.

[14] M. de Vries, V. Komornik, Unique expansions of real numbers, Preprint, 2007, `http://arxiv.org/abs/math/0609708v3`

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

[16] S. Istrail, On irreductible [sic] languages and nonrational numbers, *Bull. Math. Soc. Sci. Math. R. S. Roumanie* **21** (1977) 301–308.

[17] I. Kátai, G. Kallós, On the set for which 1 is univoque, *Publ. Math. Debrecen* **58** (2001) 743–750.

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

[19] V. Komornik, P. Loreti, Subexpansions, superexpansions and uniqueness properties in non-integer bases, *Period. Math. Hungar.* **44** (2002) 197–218.

[20] K. Mahler, Arithmetische Eigenschaften der Lösungen einer Klasse von Funktionalgleichungen, *Math. Annalen* **101** (1929) 342–266. Corrigendum **103** (1930) 532.

[21] F. L. Talmoudi, Sur les nombres de $S \cap [1, 2]$, *C. R. Acad. Sci. Paris, Sér. Math.* **285** (1977) 969–971.

[22] F. L. Talmoudi, Sur les nombres de $S \cap [1, 2[$, *C. R. Acad. Sci. Paris, Sér. Math.* **287** (1978) 739–741.