Real and p-adic expansions involving symmetric patterns

Boris Adamczewski & Yann Bugeaud

Abstract. *This paper is motivated by the non-Archimedean counterpart of a problem raised by Mahler and Mendès France, and by questions related to the expected normality of irrational algebraic numbers. We introduce a class of sequence enjoying a particular combinatorial property: the precocious occurrences of infinitely many symmetric patterns. Then, we prove several transcendence statements involving both real and p-adic numbers associated with these palindromic sequences.*

1. Introduction

One motivation for the present paper comes from the following question concerning the expansion of algebraic numbers in integer bases. It appears at the end of a paper of Mendès France [8], but in conversation he attributes the paternity of this problem to Mahler (see the discussion in [4], page 403). Though we do not find any trace of it in Mahler's work, we will refer to it as the Mahler–Mendès France problem. It can be stated as follows: For an arbitrary infinite sequence $\mathbf{a} = (a_k)_{k\geq 1}$ of 0's and 1's, prove that the real numbers

$$\sum_{k=1}^{+\infty} \frac{a_k}{2^k} \quad \text{and} \quad \sum_{k=1}^{+\infty} \frac{a_k}{3^k}$$

are algebraic if and only if both are rational.

We raise here a non-Archimedean version of this conjecture.

Problem 1. *Let p be a prime number, $\mathbf{a} = (a_k)_{k\geq 1}$ be an infinite sequence on $\{0, 1, \ldots, p-1\}$, and set*

$$\alpha = \sum_{k=1}^{+\infty} \frac{a_k}{p^k} \quad \textit{and} \quad \alpha_p = \sum_{k=1}^{+\infty} a_k p^k.$$

Then, prove that the real number α and the p-adic number α_p are algebraic if and only if both are rational.

Another formulation of Problem 1 is that for every non-eventually periodic sequence $\mathbf{a}$ at least one number among α and α_p is transcendental. To the best of our knowledge, no example is known of such a sequence for which we are able to reach this conclusion

2000 *Mathematics Subject Classification : 11J81, 11J61.*

without previously determinating the exact status of one of the corresponding numbers. Our main result (Theorem 1 below) is of a different nature. It answers positively Problem 1 for a large class of sequences, without determining the exact status of the corresponding numbers α and α_p.

Some results in the same vein are already known. Indeed, things become easier when considering addition and multiplication without carry. This yields the following analog of the Mahler–Mendès France problem in positive characteristic, as proved in [6]. For simplicity, the next two results are stated here up to an injection of a finite set $\mathcal{A}$ with cardinality at most q into the finite field $\mathbf{F}_q$.

Theorem A. *Let q_1 and q_2 with $q_1 < q_2$ be integer powers of two distinct prime numbers. Let $\mathbf{a} = (a_k)_{k\geq 1}$ be an infinite sequence on $\{0, 1, \ldots, q_1 - 1\}$, and set*

$$f(X) = \sum_{k=1}^{+\infty} a_k X^k \in \mathbf{F}_{q_1}((X)) \quad \textit{and} \quad g(X) = \sum_{k=1}^{+\infty} a_k X^k \in \mathbf{F}_{q_2}((X)).$$

Then, f and g are two algebraic functions (resp. over $\mathbf{F}_{q_1}(X)$ and over $\mathbf{F}_{q_2}(X)$) if and only if both are rational.

Theorem A is derived from two important results, namely, Christol's theorem [5] and Cobham's theorem [7].

In a previous paper [1], we proved the following mixed-characteristic analog of the Mahler–Mendès France problem. This is a consequence of the main transcendence result of [1] and of Christol's theorem.

Theorem B. *Let p be a prime, $\mathbf{a} = (a_k)_{k\geq 1}$ be an infinite sequence on $\{0, 1, \ldots, p - 1\}$ and set*

$$\alpha = \sum_{k=1}^{+\infty} \frac{a_k}{p^k} \quad \textit{and} \quad f(X) = \sum_{k=1}^{+\infty} a_k X^k \in \mathbf{F}_p((X)).$$

Then, α and f are algebraic (resp. over $\mathbf{Q}$ and over $\mathbf{F}_p(X)$) if and only if both are rational.

The proofs of Theorems A and B heavily rest on the fact that an algebraic power series defined over a finite field has a quite simple Laurent series expansions that can be precisely described in terms of finite automata (this is Christol's theorem). Both Mahler–Mendès France problem and Problem 1 seem to be much more difficult since we do not know many things about the b-adic (resp. Hensel) expansion of irrational algebraic real (resp. p-adic) numbers. Moreover, it is expected that these numbers have chaotic expansions (see Section 4). If true, this would likely be a source of serious difficulty to tackle these questions.

Our paper is organized as follows. In Section 2, we state our main result regarding Problem 1, namely Theorem 1. More precisely, we confirm our conjecture in the case where we can detect an excess of symmetry in the sequence $\mathbf{a}$. The purpose of Section 3 is to introduce the Monna map and to show how it naturally gives rise to an interesting reformulation of Problem 1. In Section 4, we present another motivation for considering palindromic real numbers that relies on the expected normality of irrational algebraic real numbers. We establish generalizations of Theorem 1 and several related results, the proofs of which are postponed to Sections 5 and 6. We conclude our paper in Section 7 by an extension of Theorem 1. The main tool for the proofs of all our results is a version of the Schmidt Subspace Theorem, due to Schlickewei [10].

2. Main result

In this Section, we introduce a class of sequences, called palindromic sequences, enjoying the precocious occurrences of symmetric patterns. This combinatorial property is described thanks to the notion of reversal. Note that continued fractions involving similar sequences were previously considered in [2].

We use the terminology from combinatorics on words. Let $\mathcal{A}$ be a finite set. The length of a finite word W on the alphabet $\mathcal{A}$, that is, the number of letters composing W, is denoted by $|W|$. The reversal (or the mirror image) of $W := a_1 \dots a_n$ is the word $\overline{W} := a_n \dots a_1$. In particular, W is a palindrome if and only if $W = \overline{W}$. We identify any sequence $\mathbf{a} = (a_n)_{n\geq 1}$ of elements from $\mathcal{A}$ with the infinite word $a_1 a_2 \dots a_n \dots$

An infinite sequence $\mathbf{a}$ is called a *palindromic sequence* if there exist real numbers w, w' and three sequences of finite words $(U_n)_{n\geq 1}$, $(V_n)_{n\geq 1}$, and $(W_n)_{n\geq 1}$ such that:

(i) For any $n \geq 1$, the word $W_n U_n V_n \overline{U}_n$ is a prefix of the word $\mathbf{a}$;
(ii) The sequence $(|V_n|/|U_n|)_{n\geq 1}$ is bounded from above by w;
(iii) The sequence $(|W_n|/|U_n|)_{n\geq 1}$ is bounded from above by w';
(iv) The sequence $(|U_n|)_{n\geq 1}$ is increasing.

In other words, a palindromic sequence has the property that infinitely many symmetric patterns (that is, the words $U_n V_n \overline{U}_n$) occur not too far from its beginning. Note that numerous examples of classical sequences in word combinatorics, such as Sturmian sequences, the Thue–Morse sequence, or Paperfolding sequences, turn out to be palindromic. The palindromic sequences should be compared with the stammering sequences introduced in [1] (see Section 6 for a definition): in the former case, we have some excess of symmetry, while, in the latter case, we have some excess of periodicity.

Our main result answers Problem 1 for palindromic sequences.

Theorem 1. *Let p be a prime number and let $\mathbf{a} = (a_k)_{k\geq 1}$ be a palindromic sequence on the alphabet $\{0, 1, \dots, p-1\}$. Then, the numbers*

$$\alpha := \sum_{k=1}^{+\infty} \frac{a_k}{p^k} \quad \text{and} \quad \alpha_p := \sum_{k=1}^{+\infty} a_k p^k$$

are algebraic if and only if both are rationals.

In transcendence theory, many results asserts that at least one number among some finite list is transcendental. Apparently, Theorem 1 is the first result of this type which deals with an Archimedean number and a non-Archimedean one.

Theorem 1 is an immediate consequence of Theorem 2 below. It is also a very particular case of Theorem 5, stated and proved in Section 7.

3. The Monna map

Monna [9] introduced a map $\mathcal{M}$ defined from the set of positive real numbers to the one of p-adic numbers. We call $\mathcal{M}$ the Monna map and recall its definition:

$$\mathcal{M} : \mathbf{R}_+^* \longrightarrow \mathbf{Q}_p$$
$$\alpha := \sum_{k\geq -k_0} a_k/p^k \longmapsto \alpha_p = \sum_{k\geq -k_0} a_k p^k.$$

Thus, $\mathcal{M}$ maps the positive real number α whose expansion in base p is the sequence $\mathbf{a} = (a_k)_{k\geq -k_0}$ to the p-adic number having the same Hensel expansion. Note that a small difficulty occurs when defining $\mathcal{M}$. Indeed, we recall that the real numbers which belong to the set $\mathbf{N}_p := \{a/p^n, \ \ a \in \mathbf{Z}_{\geq 1}, n \in \mathbf{Z}_{\geq 0}\}$ have two distinct p-adic expansions. The first one (usually called the proper one) is finite while, for the second one, we have $a_n = p-1$ for every integer n large enough. In order to well-define the map $\mathcal{M}$, Monna used the improper expansion for the elements of $\mathbf{N}_p$. In the sequel, we identify $\mathbf{N}_p$ and $\mathbf{Q}$ with their natural injections in $\mathbf{Q}_p$.

The Monna map has several interesting properties. Some of them, which can be easily checked, are displayed below:

- It is a bijection from $\mathbf{R}_+^*$ to $\mathbf{Q}_p \setminus \mathbf{N}_p$;
- It is continuous at all points of $\mathbf{R}_+^* \setminus \mathbf{N}_p$;
- The map $\mathcal{M}^{-1}$ is continuous and is even a 1-Lipschitz function (that is, $|\mathcal{M}^{-1}(x) - \mathcal{M}^{-1}(y)| \leq |x-y|_p$ for every pair $(x,y) \in (\mathbf{Q}_p \setminus \mathbf{N}_p)^2$);
- Every positive rational number is mapped by $\mathcal{M}$ on a rational number (more precisely, we have $\mathcal{M}(\mathbf{Q} \cap \mathbf{R}_+^*) = \mathbf{Q} \setminus \mathbf{N}_p$).

At the end of his paper, Monna wrote: 'Par exemple, la question se pose si les nombres réels transcendants sont transformés en nombres P-adiques transcendants et nombres algébriques en nombres algébriques.' This latter sentence seems to suggest that $(\alpha, \mathcal{M}(\alpha))$ is either a pair of algebraic numbers or a pair of transcendental numbers. Obviously, if true, this would provide a contradiction with the conclusion expected for Problem 1, which can be nicely reformulated thanks to the introduction of the map $\mathcal{M}$.

Problem 1 (alternative formulation). *Prove that the Monna map $\mathcal{M}$ takes transcendental values at all irrational algebraic points.*

4. Normal and abnormal real numbers

The aim of this Section is to provide another motivation for the present work. It is related to questions about the expected normality of irrational algebraic numbers with respect to their representation in integer bases.

Let $b \geq 2$ be an integer. Not much is known on the b-adic expansions of classical numbers, like π, $\log 2$, and $\sqrt{2}$. It is, however, widely believed that these numbers are normal in base b, that is, that every string of k letters from $\{0, 1, \ldots, b-1\}$ occurs in their b-adic expansion with the same frequency $1/b^k$. Despite some recent progress for $\sqrt{2}$ (and, more generally, for every irrational algebraic number), we are still very far away from confirming this guess.

In this Section, we are interested in the combinatorial properties of the b-adic expansions of irrational algebraic numbers. We are motivated by the following principle, which, if true, would be a first step towards a proof that these numbers are normal in base b:

If a real irrational number ξ is clearly abnormal, in the sense that, for some $b \geq 2$, its b-adic expansion strongly differs from a normal sequence, then it is transcendental.

In a previous work [3] (see also Theorem D in Section 6), we established a combinatorial criterion stating that if large blocks of digits do repeat unusually close to the beginning of the sequence of digits, then ξ must be either rational or transcendental. Thus, an excess of repetitions in a non-eventually periodic sequence does imply transcendence. We suggest here to consider another combinatorial property, namely an excess of symmetry. From now on, we say that a real (resp. p-adic) number is a *palindromic number* if its expansion in some base $b \geq 2$ (resp. its Hensel expansion) is a palindromic sequence (as defined in Section 2) and we investigate the following problem:

Problem 2. *Prove that any irrational (real or p-adic) palindromic number is transcendental.*

Following our general principle, it is likely that such numbers are transcendental. Although we are not able at this point to give a positive answer to Problem 2, our Theorem 1 provides a first step in this direction. Actually, Theorem 1 is a consequence of more general results that we state below. Let $b \geq 2$ be an integer and $(a_k)_{k \geq 1}$ be a sequence on $\{0, 1, \ldots, b-1\}$. Let $\mathcal{S}$ be the set of prime divisors of b. For any $p \in \mathcal{S}$, the integer sequence $(d_n)_{n \geq 1}$, where

$$d_n = \sum_{k=1}^{n} a_k b^k,$$

converges in $\mathbf{Z}_p$ to a p-adic number, that we denote by α_p. Observe that, if $b = p$ (as it is the case in Theorem 1), then the Hensel expansion of α_p is given by $\sum_{k=1}^{+\infty} a_k p^k$.

Theorem 2. *Let $b \geq 2$ be an integer and $\mathbf{a} = (a_k)_{k \geq 1}$ be a palindromic sequence on the alphabet $\{0, 1, \ldots, b-1\}$, with parameters w and w'. Assume that $\mathbf{a}$ is not ultimately* *periodic. Let p be a prime divisor of b and let p^u be the greatest power of p dividing b. Assume that*

$$\frac{\log p^u}{\log b} > \frac{w'}{1+w'}. \tag{4.1}$$

Then, at least one of the numbers

$$\alpha := \sum_{k=1}^{+\infty} a_k/b^k, \quad \alpha_p := \sum_{k=1}^{+\infty} a_k b^k$$

is transcendental.

Since condition (4.1) is satisfied whenever b is a prime number, our Theorem 1 is an immediate consequence of Theorem 2.

If the mild condition (4.1) is not satisfied, our method yields a weaker conclusion than in Theorem 2.

Theorem 3. *Let $b \geq 2$ be an integer and $\mathbf{a} = (a_k)_{k\geq 1}$ be a palindromic sequence on the alphabet $\{0, 1, \ldots, b-1\}$. Let $\mathcal{S}$, α and α_p be as above. Then, the set $\{\alpha\} \cup \{\alpha_p : p \in \mathcal{S}\}$ is either composed solely of rational numbers, or it contains at least one transcendental number.*

Condition (4.1) of Theorem 2 is trivially satisfied if $w' = 0$, that is, if W_n is the empty word for any $n \geq 1$ in the definition of a palindromic sequence. This motivates the introduction of the notion of *initially palindromic* sequences. An infinite sequence $\mathbf{a}$ is initially palindromic if there exist a real number w and two sequences of finite words $(U_n)_{n\geq 1}$, $(V_n)_{n\geq 1}$ such that:

(i) For any $n \geq 1$, the word $U_n V_n \overline{U}_n$ is a prefix of the word $\mathbf{a}$;
(ii) The sequence $(|V_n|/|U_n|)_{n\geq 1}$ is bounded from above by w;
(iii) For any $n \geq 1$, $|U_{n+1}| \geq (w+2)|U_n|$.

The factor $w+2$ occurring in (iii) above is motivated by Theorem 4 below. Needless to say, to achieve (iii), we may if needed extract a subsequence from $(U_n)_{n\geq 1}$. We display a particular case of Problem 2:

Problem 3. *Prove that any irrational (real or p-adic) initially palindromic number is transcendental.*

A partial answer to Problem 3 is given by the following immediate corollary to Theorem 2.

Corollary 1. *Let $b \geq 2$ be an integer and $\mathbf{a} = (a_k)_{k\geq 1}$ be an initially palindromic sequence on the alphabet $\{0, 1, \ldots, b-1\}$. Assume that $\mathbf{a}$ is not ultimately periodic. Then, either the real number $\alpha := \sum_{k=1}^{+\infty} a_k/b^k$ is transcendental, or, for any prime divisor p of b, the p-adic number $\alpha_p := \sum_{k=1}^{+\infty} a_k b^k$ is transcendental.*

To complement Corollary 1, we establish that if the expansion of α is initially palindromic, then, under some additional assumption on the density of the symmetric prefixes, it is possible to establish the transcendence of α and of all the α_p's.

Theorem 4. *Let $b \geq 2$ an integer and p be a prime divisor of b. Let $\mathbf{a}$ be an initially palindromic sequence, and let $(U_n)_{n\geq 1}$ be the corresponding sequence of prefixes satisfying conditions (i) to (iii). If we have*

$$\liminf_{n\to+\infty} \frac{|U_{n+1}|}{|U_n|} < +\infty, \tag{4.2}$$

then the real number $\alpha := \sum_{k=1}^{+\infty} a_k/b^k$ and the p-adic number $\alpha_p := \sum_{k=1}^{+\infty} a_k b^k$ are either both rational or both transcendental.

The proof of Theorem 4 relies on results from [3] and [1]. We show that any initially palindromic sequence satisfying (4.2) has an excess of repetitions.

5. Proofs of Theorems 1 to 3

The main tool for the proofs of Theorems 1 to 3 is the following version of the Schmidt Subspace Theorem, established by Schlickewei [10]. Throughout the paper, for any prime number p, the p-adic absolute value $|\cdot|_p$ is normalized in such a way that $|p|_p = p^{-1}$.

Theorem C. *Let $m \geq 2$ be an integer. Let $L_{1,\infty}, L_{2,\infty}, \ldots, L_{m,\infty}$ be m linearly independent linear forms in the variable $\mathbf{x} = (x_1, x_2, \ldots, x_m)$, with real algebraic coefficients. Let $\mathcal{S}$ be a finite set of prime numbers. For any prime p in $\mathcal{S}$, let $L_{1,p}, L_{2,p}, \ldots, L_{m,p}$ be linear forms in the same variable $\mathbf{x} = (x_1, x_2, \ldots, x_m)$ that are linearly independent and whose coefficients are algebraic p-adic numbers. Let ε be a positive real number. Then, all the solutions $\mathbf{x} = (x_1, x_2, \ldots, x_m)$ in $\mathbf{Z}^m$ of the inequality*

$$\prod_{i=1}^{m} |L_{i,\infty}(\mathbf{x})| \prod_{p\in\mathcal{S}} \prod_{i=1}^{m} |L_{i,p}(\mathbf{x})|_p < (\max\{|x_1|, \ldots, |x_m|\})^{-\varepsilon}$$

are contained in a finite union of proper subspaces of $\mathbf{Q}^m$.

We begin with the proof of Theorem 3.

Proof of Theorem 3. Let $b \geq 2$ be an integer. Consider a palindromic sequence $\mathbf{a} = (a_k)_{k\geq 1}$ defined over the alphabet $\{0, 1, \ldots, b-1\}$. We assume that the parameters w, w', $(U_n)_{n\geq 1}$, $(V_n)_{n\geq 1}$ and $(W_n)_{n\geq 1}$, are fixed, and we set $r_n = |U_n|$, $s_n = |V_n|$ and $t_n = |W_n|$ for any $n \geq 1$. Recall that

$$\mathbf{a} = W_n U_n V_n \overline{U}_n \ldots$$

and that

$$\alpha = \sum_{k=1}^{+\infty} \frac{a_k}{b^k}, \quad \alpha_p := \sum_{k=1}^{+\infty} a_k b^k,$$

for any prime divisor p of b.

We assume that α and all the α_p's are algebraic, and we aim at proving that all are rational numbers. In order to do this, we will apply Theorem C.

Let n be a positive integer. Define the integer polynomial $P_n(X)$ by

$$P_n(X) = \sum_{k=1}^{t_n} a_k\, X^k + \sum_{k=t_n+1}^{2r_n+2t_n+s_n} a_{2r_n+2t_n+s_n-k+1}\, X^k.$$

Since the rational $P_n(b)/b^{2r_n+2t_n+s_n+1}$ has the b-adic expansion

$$\frac{P_n(b)}{b^{2r_n+2t_n+s_n+1}} = 0.W_n U_n V_n \overline{U}_n\, \overline{W}_n,$$

we get that

$$\left|\alpha - \frac{P_n(b)}{b^{2r_n+2t_n+s_n+1}}\right| \leq \frac{1}{b^{2r_n+s_n+t_n}}. \tag{5.1}$$

Now, observe that

$$P_n(X) = \sum_{k=1}^{r_n+t_n} a_k\, X^k + \sum_{k=r_n+t_n+1}^{2r_n+2t_n+s_n} a_{2r_n+2t_n+s_n-k+1}\, X^k.$$

Consequently, for any prime divisor p of b, we have

$$|P_n(b) - \alpha_p|_p \leq |b|_p^{r_n+t_n+1}. \tag{5.2}$$

Consider the linearly independent linear forms:

$$L_{1,\infty} = \alpha X - Y, \quad L_{2,\infty} = X, \quad L_{3,\infty} = Z.$$

Let $p_1, \ldots, p_\ell$ be the prime divisors of b. For $j = 1, \ldots, \ell$, consider the linearly independent linear forms

$$L_{1,j} = X, \quad L_{2,j} = Y - \alpha_{p_j} Z, \quad L_{3,j} = Z.$$

By assumption, all these linear forms have algebraic (real or p-adic) coefficients. We evaluate the product of these linear forms at the integer points

$$\mathbf{x}_n = \big(b^{2t_n+2r_n+s_n+1}, P_n(b), 1\big),$$

and we infer from (5.1) and (5.2) that

$$\Pi_n := \prod_{i=1}^{3} |L_{i,\infty}(\mathbf{x_n})| \prod_{j=1}^{l} \prod_{i=1}^{3} |L_{i,j}(\mathbf{x_n})|_p \ll b^{t_n} \prod_{j=1}^{\ell} |b|_{p_j}^{r_n+t_n} = b^{-r_n}.$$

By Conditions (ii) and (iii), we easily get that

$$\Pi_n \ll (\max\{b^{2t_n+2r_n+s_n+1}, P_n(b), 1\})^{-\varepsilon},$$

for some positive real number ε. Here and below, the constant implied by $\ll$ does not depend on n. Then, by Theorem C, there exist a non-zero integer triple (z_1, z_2, z_3) and an infinite set of distinct positive integers $\mathcal{N}_1$ such that

$$z_1 b^{2t_n+2r_n+s_n+1} + z_2 P_n(b) + z_3 = 0, \tag{5.3}$$

for any n in $\mathcal{N}_1$. Dividing (5.3) by $b^{2t_n+2r_n+s_n+1}$ and letting n tend to infinity along $\mathcal{N}_1$, we get that α is rational. Consequently, the sequence $\mathbf{a}$ is ultimately periodic and all the α_p's are rationals. Thus, we have proved that if α and the α_p's are algebraic, then all are rationals. This establishes Theorem 3. $\square$

Now, we show how to modify the proof of Theorem 3 in order to establish Theorem 2.

Proof of Theorem 2. We keep the same notation as in the proof of Theorem 3, but we apply Theorem C with a slightly different set of linear forms. Let p be a prime divisor of b and denote by p^u the greatest power of p dividing b. We consider the linearly independent linear forms with algebraic coefficients

$$L_{1,\infty} = \alpha X - Y, \quad L_{2,\infty} = X, \quad L_{3,\infty} = Z,$$

$$L_{1,p} = X, \quad L_{2,p} = Y - \alpha_p Z, \quad L_{3,p} = Z,$$

and, for every prime number $p' \neq p$ dividing b, we set

$$L_{1,p'} = X, \quad L_{2,p'} = Y, \quad L_{3,p'} = Z.$$

We evaluate the product Π_n of the norms of these linear forms at the integer points

$$\mathbf{x}_n = \big(b^{2t_n+2r_n+s_n+1}, P_n(b), 1\big).$$

Recall that, by assumption, there is a real number w' such that $t_n \leq w' r_n$ for any $n \geq 1$. Consequently, we infer from (5.1) and (5.2) that

$$\Pi_n \ll b^{t_n} \, |b|_p^{r_n+t_n} = b^{t_n} p^{-u(r_n+t_n)} \leq (b^{w'} p^{-u(1+w')})^{r_n}.$$

Then, it follows from (4.1) that there exists a positive real number $C < 1$ satisfying $\Pi_n \ll C^{r_n}$. Thus, by Conditions (ii) and (iii) we get that

$$\Pi_n \ll (\max\{b^{2t_n+2r_n+s_n+1}, P_n(b), 1\})^{-\varepsilon},$$

for some positive real number ε. We then apply Theorem C and we conclude exactly as in the proof of Theorem 3. $\square$

6. Proof of Theorem 4

We first recall a result from [3] and [1]. For any positive integer ℓ, we write W^ℓ for the word $W \ldots W$ (ℓ times repeated concatenation of the word W). More generally, for any positive real number x, we denote by W^x the word $W^{[x]}W'$, where W' is the prefix of W of length $\lceil (x - [x])|W| \rceil$. Here, $[y]$ and $\lceil y \rceil$ denote, respectively, the integer part and the upper integer part of the real number y. A sequence $\mathbf{a}$ is called stammering if there exist a real number $w > 1$ and two sequences $(X_n)_{n\geq 1}$ and $(Y_n)_{n\geq 1}$ of finite words such that:

(i) For any $n \geq 1$, the word $Y_n X_n^w$ is a prefix of the word $\mathbf{a}$;

(ii) The sequence $(|Y_n|/|X_n|)_{n\geq 1}$ is bounded;

(iii) The sequence $(|X_n|)_{n\geq 1}$ is increasing.

Note that, unlike in [1], we do not assume here that a stammering sequence is non-eventually periodic. We recall now the following transcendence criterion for stammering real and p-adic numbers.

Theorem D. *Let $b \geq 2$ be an integer. Let $\mathbf{a}$ be a stammering sequence. Then, the real number $\alpha := \sum_{k=1}^{+\infty} a_k/b^k$ is either rational or transcendental. Furthermore, for any prime number p dividing b, the p-adic number $\alpha_p := \sum_{k=1}^{+\infty} a_k b^k$ is either rational or transcendental.*

The last assertion of Theorem D is proved in [1] only when b is a prime number. However, it is easily seen that the same arguments can be used to establish Theorem D for any integer $b \geq 2$.

We go on with the proof of Theorem 4.

Proof of Theorem 4. Let $\mathbf{a}$, $(U_n)_{n\geq 1}$, $(V_n)_{n\geq 1}$ and w be as in the definition of an initially palindromic sequence. By assumption, there are infinitely many indices n and an integer M such that $|U_{n+1}| \leq M|U_n|$. Moreover, both words $U_n V_n \overline{U}_n$ and $U_{n+1}V_{n+1}\overline{U}_{n+1}$ are prefixes of $\mathbf{a}$, and they satisfy $|U_n V_n \overline{U}_n| \leq |U_{n+1}|$, as follows from (iii).

Thus, for each such n, there exists a finite word W_n (which may be empty) such that $U_{n+1} = U_n V_n \overline{U}_n W_n$. Consequently, $\mathbf{a}$ begins with

$$U_n V_n \overline{U}_n W_n V_{n+1} \overline{U_n V_n \overline{U}_n W_n} = U_n V_n \overline{U}_n W_n V_{n+1} \overline{W}_n U_n \overline{V}_n \overline{U}_n.$$

Set $X_n = U_n V_n \overline{U}_n W_n V_{n+1} \overline{W}_n$. Then, $\mathbf{a}$ begins with $X_n^{1+\varepsilon_n}$, where $\varepsilon_n = |U_n|/|X_n|$. Since $|X_n| \leq (2 + w)|U_{n+1}| \leq M(2 + w)|U_n|$, the sequence $\mathbf{a}$ begins in particular with $X_n^{1+\varepsilon}$, where $\varepsilon = 1/(M(2+w))$. It follows that $\mathbf{a}$ is a stammering sequence (here, Y_n is the empty word), and we thus derive Theorem 4 by applying Theorem D. □

7. An extension of Theorem 1

If the conclusion of Problem 1 turns out to be true, the we have the following situation: a sequence $\mathbf{a}$ cannot represent at once a real algebraic number in the integer base

p and a p-adic algebraic number, except in the trivial case where $\mathbf{a}$ is eventually periodic. As mentioned in Section 4, it is also believed that the expansions of irrational algebraic numbers are, in some sense, 'random'. Actually, we could reasonably (or, we should rather say unreasonably!) expect the following deeper statement: the expansion of an algebraic irrational number in the integer base p and the Hensel expansion of a p-adic algebraic irrational number are at once 'random and independent'.

We end this paper with an extension of Theorem 1 which fits into such a general philosophy.

Let $\mathbf{a} = (a_k)_{k\geq 1}$ and $\mathbf{a}' = (a'_k)_{k\geq 1}$ be sequences of elements from a finite set $\mathcal{A}$. We say that the pair $(\mathbf{a}, \mathbf{a}')$ satisfies Condition $(*)$ if there exist three sequences of finite words $(U_n)_{n\geq 1}$, $(U'_n)_{n\geq 1}$, and $(V_n)_{n\geq 1}$ such that:

(i) For any $n \geq 1$, the word $U_n V_n$ is a prefix of the word $\mathbf{a}$;

(ii) For any $n \geq 1$, the word $U'_n \overline{V}_n$ is a prefix of the word $\mathbf{a}'$;

(iii) The sequences $(|U_n|/|V_n|)_{n\geq 1}$ and $(|U'_n|/|V_n|)_{n\geq 1}$ are bounded from above;

(iv) The sequence $(|V_n|)_{n\geq 1}$ is increasing.

Note that if the pair $(\mathbf{a}, \mathbf{a})$ satisfies Condition $(*)$, then the sequence $\mathbf{a}$ is palindromic. Our extension of Theorem 1 is as follows.

Theorem 5. *Let p be a prime number. Let $\mathbf{a} = (a_k)_{k\geq 1}$ and $\mathbf{a}' = (a'_k)_{k\geq 1}$ be sequences of integers from $\{0, 1, \ldots, p-1\}$. If the pair $(\mathbf{a}, \mathbf{a}')$ satisfies Condition $(*)$, then*

$$\alpha := \sum_{k=1}^{+\infty} \frac{a_k}{p^k}, \quad \textit{and} \quad \alpha'_p := \sum_{k=1}^{+\infty} a'_k p^k,$$

are algebraic if and only if both are rational.

It follows from Theorem 5 that if we suitably perturb the sequence of digits of an irrational algebraic number written in base p to get a sequence $(a'_k)_{k\geq 1}$, then the p-adic number $\sum_{k=1}^{+\infty} a'_k p^k$ is transcendental.

Before proving Theorem 5, we need the following auxiliary result.

Lemma 1. *Let $(\mathbf{a}, \mathbf{a}')$ be a pair of sequences satisfying Condition $(*)$. If $\mathbf{a}$ is eventually periodic, then $\mathbf{a}'$ is a stammering sequence.*

Proof. Since $\mathbf{a}$ is eventually periodic, there exist two finite words U and V such that $\mathbf{a} = UVV\ldots V\ldots$ Set $M = |U| + |V|$. For any sufficiently large factor W of $\mathbf{a}$, there exists two (possibly empty) finite words A_W and B_W, with $|A_W| < M$ and $|B_W| < M$, and a positive integer s such that

$$W = A_W V^s B_W.$$

There thus exist two (possibly empty) finite words A and B, with $|A| < M$ and $|B| < M$, and two increasing sequences of positive integers $(n_k)_{k\geq 1}$ and $(s_k)_{k\geq 1}$ such that

$$V_{n_k} = AV^{s_k}B.$$

Recall that the sequences $(V_n)_{n\geq 1}$, $(U_n)_{n\geq 1}$ and $(U'_n)_{n\geq 1}$ are given by Condition $(*)$. It follows that, for any positive integer k, the sequence $\mathbf{a}'$ begins with the words $U'_{n_k}\overline{B}\,\overline{V}^{s_k}$. Set $Y_k = U'_{n_k}\overline{B}$ and $X_k = \overline{V}^{\lfloor s_k/2\rfloor}$. Then, the word $Y_k X_k^2$ is a prefix of $\mathbf{a}'$. Moreover, the sequence $(|X_k|)_{k\geq 1}$ is increasing since by assumption $(|V^{s_k}|)_{k\geq 1}$ is increasing, and the sequence $(|Y_k|/|X_k|)_{k\geq 1}$ is bounded since by assumption $(|U'_n|/|V_n|)_{n\geq 1}$ is bounded. This proves that $\mathbf{a}'$ is a stammering sequence, which ends the proof. □

Proof of Theorem 5. We assume that α and α'_p are algebraic and we aim at proving that both are rational.

We first note that it is sufficient to prove that α is rational. Indeed, if α is rational, then $\mathbf{a}$ is eventually periodic and, thanks to Lemma 1, we obtain that $\mathbf{a}'$ is a stammering sequence. We then infer from Theorem D that α_p is either rational or transcendental, and since α_p is assumed to be algebraic, this implies that α_p is rational.

For any $n\geq 1$, set $r_n = |U_n|$, $r'_n = |U'_n|$ and $s_n = |V_n|$. Let n be a positive integer. Let α_n be the rational number having the following expansion in the integer base p:

$$\alpha_n := 0.U_n V_n \overline{U'}_n.$$

Define the integer polynomials $P_n(X)$ by

$$P_n(X) := \sum_{k=1}^{r'_n} a'_k\, X^k + \sum_{k=1}^{r_n+s_n} a_{r_n+s_n-k+1} X^{k+r'_n}.$$

Note that by assumption we also have

$$P_n(X) := \sum_{k=1}^{r'_n+s_n} a'_k\, X^k + \sum_{k=s_n+1}^{r_n+s_n} a_{r_n+s_n-k+1} X^{k+r'_n}. \tag{7.1}$$

The definition of $P_n(X)$ ensures that

$$\frac{P_n(p)}{p^{r_n+s_n+r'_n+1}} = \alpha_n$$

and, by (i), we have

$$\left|\alpha - \frac{P_n(p)}{p^{r_n+s_n+r'_n+1}}\right| \leq p^{-r_n-s_n}. \tag{7.2}$$

Now, define the integer polynomials $Q_n(X)$ by

$$Q_n(X) := \sum_{k=1}^{r'_n+s_n} a'_k\, X^k.$$

We infer from (7.1) and from the definition of α_p that

$$|P_n(p) - Q_n(p)|_p < p^{-r'_n-s_n} \quad \text{and} \quad |Q_n(p) - \alpha'_p|_p < p^{-r'_n-s_n}. \tag{7.3}$$

We consider the following systems of independent linear forms with algebraic (real and p-adic) coefficients in the four variables (X, Y, Y', Z)

$$L_{1,\infty} = \alpha X - Y, \quad L_{2,\infty} = X, \quad L_{3,\infty} = Y', \quad L_{4,\infty} = Z,$$

and

$$L_{1,p} = X, \quad L_{2,p} = Y - \alpha'_p Z, \quad L_{3,p} = Y - Y', \quad L_{4,p} = Z.$$

We evaluate the product Π_n of the norms of these linear forms at the integer points

$$\mathbf{x}_n = \big(p^{r_n+s_n+r'_n+1}, P_n(p), Q_n(p), 1\big).$$

We derive from (7.2) and (7.3) that

$$\Pi_n \ll p^{r'_n}\, p^{r_n+r'_n+s_n}\, p^{r'_n+s_n}\, p^{-r_n-r'_n-s_n}\, p^{-r'_n-s_n}\, p^{-r'_n-s_n} \ll p^{-s_n}.$$

Thus, Theorem C implies the existence of a non-zero integer quadruple (z_1, z_2, z_3, z_4) and an infinite set of distinct positive integers $\mathcal{N}_1$ such that

$$z_1 p^{r_n+s_n+r'_n+1} + z_2 P_n(p) + z_3 Q_n(p) + z_4 = 0, \tag{7.4}$$

for any n in $\mathcal{N}_1$.

We have now to distinguish two cases.

Let us first assume that $z_3 = 0$. Then, dividing (7.4) by $p^{r_n+s_n+r'_n+1}$ and letting n tend to infinity along $\mathcal{N}_1$, we get that $z_1 + z_2\alpha = 0$. Since (z_1, z_2, z_3, z_4) is a non-zero quadruple, we easily check that $z_2 \neq 0$ and α is thus rational.

Now let us assume that $z_3 \neq 0$. Then, dividing (7.4) by $p^{r_n+s_n+r'_n+1}$ and letting n tend to infinity along $\mathcal{N}_1$, we get that

$$\beta := \lim_{\mathcal{N}_1 \ni n \to +\infty} \frac{Q_n(p)}{p^{r_n+s_n+r'_n+1}} = -\frac{z_1 + z_2\alpha}{z_3}.$$

Note that it follows from our assumption that β is algebraic. Furthermore, we infer from (7.2) that, for any n in $\mathcal{N}_1$, we have

$$\begin{aligned} \left|\beta - \frac{Q_n(p)}{p^{r_n+s_n+r'_n+1}}\right| &= \left|\frac{z_1 + z_2\alpha}{z_3} - \frac{z_1 + z_2 P_n(p)/p^{r_n+s_n+r'_n+1} + z_4/p^{r_n+s_n+r'_n+1}}{z_3}\right| \\ &\ll \frac{1}{p^{r_n+s_n}}. \end{aligned} \tag{7.5}$$

Consider now the independent linear forms with algebraic (real and p-adic) coefficients

$$L_{1,\infty} = \beta X - Y, \quad L_{2,\infty} = X, \quad L_{3,\infty} = Z,$$

and

$$L_{1,p} = X, \quad L_{2,p} = Y - \alpha'_p Z, \quad L_{3,p} = Z.$$

We evaluate the product Π'_n of the norms of these linear forms at the integer points

$$\mathbf{x}_n = \bigl(p^{r_n+s_n+r'_n+1}, Q_n(p), 1\bigr).$$

We derive from (7.3) and (7.5) that

$$\Pi'_n \ll p^{-s_n}.$$

Thus, Theorem C implies the existence of a non-zero integer triple (z'_1, z'_2, z'_3) and an infinite set of distinct positive integers $\mathcal{N}_2 \subset \mathcal{N}_1$ such that

$$z'_1 p^{r_n+s_n+r'_n+1} + z'_2 Q_n(p) + z'_3 = 0, \tag{7.6}$$

for any n in $\mathcal{N}_2$. Dividing (7.6) by $p^{r_n+s_n+r'_n+1}$ and letting n tend to infinity along $\mathcal{N}_2$, we get that β is rational since it is easily verified that $z'_2 \neq 0$. We thus infer from the definition of β that α is rational since we can also easily check that $z_2 \neq 0$. This ends the proof of Theorem 5. □

References

[1] B. Adamczewski & Y. Bugeaud, *On the complexity of algebraic numbers I. Expansions in integer bases*, Ann. of Math. To appear.

[2] B. Adamczewski & Y. Bugeaud, *Palindromic continued fractions.* Preprint 2005.

[3] B. Adamczewski, Y. Bugeaud & F. Luca, *Sur la complexité des nombres algébriques*, C. R. Acad. Sci. Paris 339 (2004), 11–14.

[4] J.-P. Allouche & J. O. Shallit, Automatic Sequences: Theory, Applications, Generalizations, Cambridge University Press, Cambridge, 2003.

[5] G. Christol, Ensembles presques périodiques k-reconnaissables, *Theor. Comput. Sci.* **9** (1979), 141–145.

[6] G. Christol, T. Kamae, M. Mendès France & G. Rauzy, *Suites algébriques, automates et substitutions*, Bull. Soc. Math. France 108 (1980), 401–419.

[7] A. Cobham, On the base-dependence of sets of numbers recognizable by finite automata, *Math. Systems Theory* **3** (1969), 186–192.

[8] M. Mendès France, *Sur les décimales des nombres algébriques réels*, in Seminar on Number Theory, 1979–1980, Exp. No. 28, 7 pp., Univ. Bordeaux I, Talence, 1980.

[9] A. F. Monna, *Sur une transformation simple des nombres P-adiques en nombres réels*, Indagationes Math. 14, (1952). 1–9.

[10] H. P. Schlickewei, *The $\mathfrak{p}$-adic Thue–Siegel–Roth–Schmidt theorem*, Arch. Math.(Basel) 29 (1977), 267–270.

Boris Adamczewski
CNRS, Institut Camille Jordan
Université Claude Bernard Lyon 1
Bât. Braconnier, 21 avenue Claude Bernard
69622 VILLEURBANNE Cedex (FRANCE)

Boris.Adamczewski@math.univ-lyon1.fr

Yann Bugeaud
Université Louis Pasteur
U. F. R. de mathématiques
7, rue René Descartes
67084 STRASBOURG Cedex (FRANCE)

bugeaud@math.u-strasbg.fr