On Number Representation in a Rational Base

Shigeki Akiyama[*] Christiane Frougny[†] Jacques Sakarovitch[‡]

November 4, 2004

Summary In this paper[1] we introduce and study a new method for representing positive integers and real numbers in the base $\frac{p}{q}$, where $p > q \geqslant 2$ are coprime integers. It differs from the classical "greedy" algorithm and computes the digits from the least significant to the most significant one. Every integer has a unique such expansion. The set of expansions of the integers is not a regular language. However addition can be performed (least significant digit first) by a letter-to-letter finite transducer.

We show that every real number has at least one such expansion and, under the hypothesis that $p \geqslant 2q - 1$, that every real number has exactly one such expansion but for a countable infinite set of them which have two expansions. We explain how these expansions can be approximated and characterize the expansions of reals that have two expansions.

These results which are mainly obtained by means of techniques from theoretical computer science are not developped only for their own sake but also as they relate to other problems in combinatorics and number theory. A first example is a new interpretation and expansion of the constant $K(p)$ from the so-called "Josephus problem". More important, these expansions in the base $\frac{p}{q}$ allow us to make some progress in the problem of the distribution of the fractional part of the powers of rational numbers.

EXTENDED ABSTRACT

In this paper, we introduce and study a new method for representing positive integers and real numbers in the base $\frac{p}{q}$, where $p > q \geqslant 2$ are coprime integers. The idea of non-standard representation systems of numbers is far from being original and there have been extensive studies of these, from a theoretical point of view as well as for improving computation algorithms. It is worth (briefly) recalling first the main features of these systems in order to clearly put in perspective and in contrast the results we have obtained on rational base systems.

There exists many non-standard numeration systems; the base can be a negative integer, or a real number, or a complex number, etc; the reader may consult [12, Chap. 4, Vol. 2]. Representation in integer base with signed digits was popularized in computer arithmetic by Avizienis [2] and can be found earlier in a work of Cauchy [4]. When the base is a real number $\beta > 1$, by the greedy algorithm of Rényi [18] any non-negative real number is given an expansion on the canonical alphabet $\{0, 1, \ldots, \lfloor\beta\rfloor\}$. Note that a number may have several representations on the same alphabet, and that the greedy one is the greatest in the lexicographical order. The set of greedy β-expansions of numbers of $[0, 1[$ is shift-invariant,

[*]Dept. of Mathematics, Niigata University, akiyama@math.sc.niigata-u.ac.jp

[†]LIAFA, UMR 7089, and Université Paris 8, Christiane.Frougny@liafa.jussieu.fr

[‡]LTCI, UMR 5141, CNRS / ENST, sakarovitch@enst.fr

[1]Work partially supported by the CNRS/JSPS contract 13 569.

and its closure forms a symbolic dynamical system called the β-*shift*. The properties of the β-shift are well understood, using the so-called β-expansion of 1, see [16, 13].

When β is a Pisot number, that is to say an algebraic integer such that all its Galois conjugates are less than 1 in modulus, the β-representation shares several properties with the standard case where the base is an integer: the set of greedy representations is recognizable by a finite automaton; the conversion between two alphabet of digits (in particular addition) is realized by a finite automaton [8]. When β is a Pisot number which is a unit, then self-similar tilings can be associated with β-expansions, see in particular [1].

It is also possible to represent integers with respect to an increasing sequence of integers such as the sequence of Fibonacci numbers, see [7]. In particular the link between Fibonacci expansions and $\frac{1+\sqrt{5}}{2}$-expansions has been proved to be computable by a finite automaton in [9].

In this work we introduce number representation in a rational base, call "p/q-*expansion*" of the number n, which is a way of writing an integer n in the *base* p/q by an adaptation of the Bezout algorithm which produces the digits least significant first. This expansion is unique for every n but it has to be stressed that *it is not* the expansion that would be obtained by the classical "greedy algorithm" in that base. The $\frac{p}{q}$-expansion are written on the alphabet $A = \{0, 1, \ldots, p-1\}$, but not every word is admissible. These $\frac{p}{q}$-expansions share some properties with the expansion in an integer base — digit set conversion is realized by a finite automaton — and are completely different as far as some of their aspects are concerned. Above all, the set $L_{\frac{p}{q}}$ of all $\frac{p}{q}$-expansions is not a regular language (not even a context-free one).

By construction, the set $L_{\frac{p}{q}}$ is prefix-closed and any element can be extended (to the right) in $L_{\frac{p}{q}}$. Hence, $L_{\frac{p}{q}}$ can naturally be seen as a *subtree* of the full tree of the free monoid A^*, and this tree is naturally associated with a set of infinite words $W_{\frac{p}{q}}$, subset of $A^{\mathbb{N}}$. It is noteworthy that this set is not shift-invariant, and that the smallest symbolic dynamical system containing it is the full shift $A^{\mathbb{N}}$. These infinite words may in turn be seen as a $\frac{p}{q}$-expansion of a real number and we prove that conversely any real within an interval $[0, \boldsymbol{\omega}_{\frac{p}{q}}]$ has exactly one such expansion but for a countable number of them which have two such expansions. Note that no expansion is eventually periodic. The constant $\boldsymbol{\omega}_{\frac{p}{q}}$ is the numerical value of a special infinite word, called the *maximal word*. The use of the $\frac{p}{q}$ number system gives an easy method to compute this constant. It happens that the constant $K(p)$ of [15, 10, 20] related to the Josephus problem is a special case of our constant $\boldsymbol{\omega}_{\frac{p}{q}}$ (with $q = p - 1$).

We explain how the expansions of real numbers can be computed (in fact approximated). Using tools from automata theory we are able to present a characterization of reals having two expansions that is formulated in terms that are remarkably similar to the statement of the celebrated Mahler's problem on fractional part of the powers of a rational [14]. Recall that Hardy and Littlewood proved that the sequence of fractional parts of the powers of a real number θ is equidistributed for almost all real numbers $\theta > 1$; the integers and the Pisot numbers are exceptional. The problem is still open for a rational number, see [21, 14, 17, 6, 3].

In conclusion, we have introduced and studied here a fascinating object which can be seen from many sides, which raises still many difficult questions and whose further study will certainly mix techniques from combinatorics of words, automata theory, and number theory.

1 Preliminaries

1.1 Words and automata

An *alphabet* A is a finite set. A finite sequence of elements of A is called a *word*, and the set of words on A is the free monoid A^*. The *length* of a word v is equal to the number of its letters, and is denoted by $|v|$. The *empty word* is denoted by ε. A word u is a *factor* of a word v if $v = xuy$. If x (resp. y) is the empty word, u is a *prefix* (resp. *suffix*) of v. If the set A is ordered then A^* is ordered by the radix order, see [13] for instance for definition.

An *infinite word* over A is an infinite sequence of elements of A. In this work, infinite words can be indexed by positive integers, or by negative integers, depending on the context. Let $A^{\mathbb{N}}$ denote the set of infinite words on A. This set is a compact metric space. An infinite word is said to be *eventually periodic* if it is of the form $uv^\omega = uvvvvv\cdots$ where u and v belong to A^*.

For definitions and results on automata theory the reader is refered to [5] and to [11]. A *two-tape finite automaton* (or *transducer*) is a finite automaton labelled by elements of $A \times B^*$. A two-tape finite automaton is *sequential* when the projection on the input tape A is a deterministic automaton. It is said to be *letter-to-letter* if the labels of edges belong to $A \times B$. A terminal function ω is defined on the set of states. In this work we consider *right* automata, where words are processed from right to left (*i.e.* from least significant to most significant digit). A function $\varphi : A^* \to B^*$ is *realizable by a right finite automaton* if there exists such an automaton such that the graph of φ is the set $\{(u,v) \in A^* \times B^* \bigm|$ there exists a path from the initial state s_0 to a state s labelled by (u, v'') and such that $v = v'v''$, $v' = \omega(s)\}$.

1.2 The $\frac{p}{q}$ number system

Let $p > q \geqslant 1$ be two co-prime integers and let U be the sequence defined by:

$$U = \{u_i = \frac{1}{q}\left(\frac{p}{q}\right)^i \bigm| i \in \mathbb{Z}\}.$$

We will say that U, *together with* the alphabet $A = \{0, \ldots, p-1\}$, is the $\frac{p}{q}$ *number system*. If $q = 1$, it is exactly the classical number system in base p.

A *representation in the system* U of a non-negative real number x on a finite alphabet of digits D is an infinite sequence of elements of D indexed by a section of $\mathbb{Z}$: $(d_i)_{k \geqslant i \geqslant -\infty}$, such that:

$$x = \sum_{-\infty}^{i=k} d_i u_i\,, \quad \text{an equation that is written as} \quad \langle x \rangle_{\frac{p}{q}} = d_k \cdots d_0 {\scriptstyle\bullet} d_{-1} d_{-2} \cdots\,,$$

most significant digit first. When a representation ends in infinitely many zeroes, it is said to be *finite*, and the trailing zeroes are omitted. When all the d_i with negative index are zeroes, the representation is said to be an *integer representation*. Conversely, the numerical value in the system U of a word on an alphabet of digits D is given by the *evaluation map*

$$\pi\colon D^{\mathbb{Z}} \longrightarrow \mathbb{R} \qquad \mathbf{d} = \{d_i\}_{k \geqslant i \geqslant -\infty} \longmapsto \pi(\mathbf{d}) = \sum_{-\infty}^{i=k} d_i u_i$$

2 Representation of the integers

2.1 The $\frac{p}{q}$-expansion of an integer

Let N be any positive integer. Write $N_0 = N$ and, for $i \geqslant 0$, write

$$q\,N_i = p\,N_{i+1} + a_i \tag{1}$$

where a_i is the remainder of the Euclidean division of $q\,N_i$ by p, and thus belongs to A. This is an algorithm that produces the digits least significant first, that is to say, from right to left, and stops for some k when $N_{k+1} = 0$. Thus $N = \sum_{i=0}^{k} a_i u_i$. The word $a_k \cdots a_0$ is a $\frac{p}{q}$-representation of N and it can be proved that it is the unique *finite* one (under the condition that $a_k \neq 0$). We have thus established:

Theorem & Definition 1 *Every non-negative integer N has an* integer *representation in the $\frac{p}{q}$ number system. It is the* unique finite $\frac{p}{q}$-representation *of N. It will be called* the $\frac{p}{q}$-expansion *of N and written $\langle N \rangle_{\frac{p}{q}}$.* ■

Example 1 *Let $p = 3$ and $q = 2$, then $A = \{0, 1, 2\}$ — this will be our main running example. Table 1 gives the $\frac{3}{2}$-expansions of the ten first integers.* ◇

Remark 1 *This representation is not — if $q \neq 1$ — the representation obtained by the greedy (left-to-right) algorithm, (see [18] or [13, Chapter 7]), which gives representations on the alphabet $\{0, 1, \cdots, \lfloor \frac{p}{q} \rfloor\}$. Moreover, it follows from Theorem 1 that no integer (but 1) is given a finite representation by the greedy algorithm.*

If $q = 1$ on the contrary, the above algorithm gives the same representation as the one given by the classical greedy algorithm.

0	0
2	1
21	2
210	3
212	4
2101	5
2120	6
2122	7
21011	8
21200	9
21202	10

Table 1.

2.2 The set of $\frac{p}{q}$-expansions

Let us denote by $L_{\frac{p}{q}}$ the set of $\frac{p}{q}$-*expansions* of the non-negative integers. If $q = 1$ then $L_{\frac{p}{q}}$ is the set of all words of A^* which do not begin with a 0; if we release this last condition, we then get the whole A^*. If $q \neq 1$, $L_{\frac{p}{q}}$ is prefix-closed by construction and the observation of Table 1 shows that it is not suffix-closed. To some extend, the description and the understanding of this set $L_{\frac{p}{q}}$ is all what this work is about.

For each a in A, we define a partial map τ_a from $\mathbb{N}$ into itself: for z in $\mathbb{N}$, $\tau_a(z) = \frac{1}{q}(pz + a)$ if the latter is an integer, $\tau_a(z)$ is undefined otherwise. The labelled tree $T_{\frac{p}{q}}$ is then constructed as follows: The root is labelled by 0, the children of a node labelled by z are nodes labelled by the (defined) $\tau_a(z)$, the edge from z to $\tau_a(z)$ being labelled by a. Figure 1 shows (a part) of $T_{\frac{3}{2}}$. Let us call *access label* of a node s, and write $w(s)$, the label of *the* path from the root to s. Let us denote by $I_{\frac{p}{q}}$ the "upper part" of $T_{\frac{p}{q}}$, that is the subtree made of nodes whose

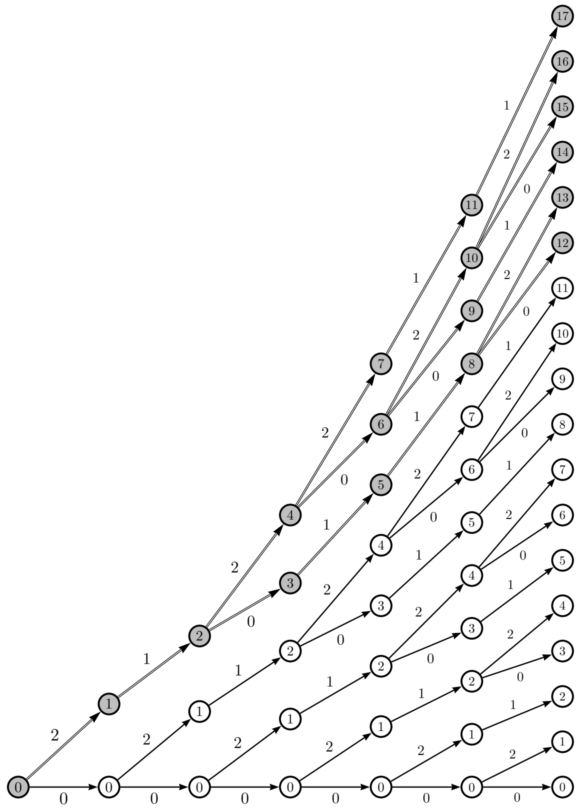

Figure 1: The tree $T_{\frac{3}{2}}$

access label does not begin with a 0 (in Figure 1, nodes of $I_{\frac{3}{2}}$ are in grey and edges are drawn with double lines).

The following properties of $L_{\frac{p}{q}}$, $T_{\frac{p}{q}}$ and $I_{\frac{p}{q}}$ are established:

(i) For any node s of $T_{\frac{p}{q}}$, $\lambda(s) = \pi(w(s))$.

(ii) The nodes of depth k in $T_{\frac{p}{q}}$ are labelled by the integers from 0 to $\left\lfloor (\frac{p}{q})^k \right\rfloor$.

(iii) The labelling of nodes (in $\mathbb{N}$) gives the ordering in the radix order on $I_{\frac{p}{q}}$.

Closer investigations give the following (under the hypothesis that $q \neq 1$):

Proposition **2** *No two subtrees of $I_{\frac{p}{q}}$ are isomorphic.*

Corollary **3** *$L_{\frac{p}{q}}$ is not a regular language.*

Any word of A^k is a factor of a word of $L_{\frac{p}{q}}$, and thus a suffix of $L_{\frac{p}{q}}$; more precisely it holds:

Proposition **4** *For any $w \in A^k$ there exists a unique n, $0 \leqslant n \leqslant p^k$, such that there exists a $v \in A^*$ with $\langle n \rangle_{\frac{p}{q}} = v\,w$.*

Proposition **5** *$L_{\frac{p}{q}}$ is not a context-free language.*

2.3 Conversion between alphabets

Let D be a finite alphabet of (positive or negative) digits that contains A. The *digit-set conversion* is a map $\chi_D \colon D^* \to A^*$ that commutes to the evaluation map π, that is which preserves the numerical value:

$$\forall w \in D^* \qquad \pi(\chi_D(w)) = \pi(w) \ .$$

The *integer addition* may be seen — after digit-wise addition — as a particular case of a digit-set conversion χ_D with $D = \{0, 1, \ldots, 2(p-1)\}$.

PROPOSITION **6** *For any alphabet D the conversion χ_D is realizable by a letter-to-letter right sequential finite transducer $\mathcal{C}_D$.*

The states of $\mathcal{C}_D$ are *integers*, the state 0 is *initial*, and the final function of a state h (with h positive) is the $\frac{p}{q}$-expansion of h. A transition labelled by (d, a) — d in D, a in A — goes from h to k if and only if

$$q\,h + d = p\,k + a \ . \tag{2}$$

One then verifies that the set of accessible states is finite and this establishes Proposition 6. Figure 2 shows the converter that realizes addition in the $\frac{3}{2}$-system.

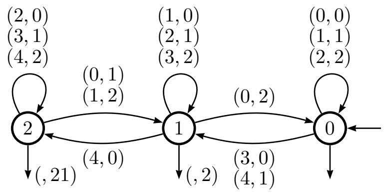

Figure 2: A converter for addition in the $\frac{3}{2}$ number system

REMARK **2** *Let us stress that χ_D is defined on the whole set D^* even for word v such that $\pi(v)$ is not an integer, and also that, if $\pi(v)$ is in $\mathbb{N}$, then $\chi_D(v)$ is the unique $\frac{p}{q}$-expansion of $\pi(v)$.*

REMARK **3** *As $\frac{p}{q}$ is not a Pisot number (when $q \neq 1$), the conversion from any representation onto the expansion computed by the greedy algorithm is not realized by a finite transducer (see [13, Ch. 7]).*

3 Representation of the reals

The tree $T_{\frac{p}{q}}$ that contains all $\frac{p}{q}$-expansions of the integers will now be used to define representations of real numbers.

3.1 The $\frac{p}{q}$-expansions of reals

Let $W_{\frac{p}{q}}$ be the set of labels of infinite paths starting from the root 0 in $T_{\frac{p}{q}}$. In the previous sections digits where indexed by decreasing nonnegative integers from left to right; as we shall now deal mainly with the "decimal" part of the expansions, we find it much more convenient *to reverse the direction of indexing* and use the positive indices *after* the decimal point.

DEFINITION **7** *Let* $\mathbf{a} = \{a_i\}_{i\geqslant 1}$ *be in* $W_{\frac{p}{q}}$*. The infinite word* $\mathbf{a}$ *is a* $\frac{p}{q}$-expansion *of the real number* x*:*

$$x = \pi(\boldsymbol{.}\mathbf{a}) = \frac{1}{q}\sum_{i\geqslant 1} a_i \left(\frac{q}{p}\right)^i .$$

The set $W_{\frac{p}{q}}$ clearly contains a maximal element with respect to the lexicographical order, an infinite word denoted by $\mathbf{t}_{\frac{p}{q}}$ (or $\mathbf{t}$ for short). For instance:

$$\mathbf{t}_{\frac{3}{2}} = 21221112212112212121211221\cdots$$

Let $X_{\frac{p}{q}} = \pi(W_{\frac{p}{q}})$. Let $\boldsymbol{\omega}_{\frac{p}{q}} = \pi(\boldsymbol{.}\mathbf{t}_{\frac{p}{q}})$ be the numerical value of the maximal infinite word. Clearly the elements of $X_{\frac{p}{q}}$ are non-negative real numbers less than or equal to $\boldsymbol{\omega}_{\frac{p}{q}}$. Note that $\boldsymbol{\omega}_{\frac{p}{q}} \leqslant \frac{p-1}{p-q}$. The fact that the $\frac{p}{q}$ number system may be used for representing the reals is expressed by the following statement.

THEOREM **8** *Any real in* $[0, \boldsymbol{\omega}_{\frac{p}{q}}]$ *has a* $\frac{p}{q}$*-expansion, that is to say,* $X_{\frac{p}{q}} = [0, \boldsymbol{\omega}_{\frac{p}{q}}]$ *.*

The proof of Theorem 8 relies on three facts. First, $W_{\frac{p}{q}}$ is closed in the compact set $A^{\mathbb{N}}$, hence is compact. Second, the map $\pi\colon W_{\frac{p}{q}} \to X_{\frac{p}{q}}$ is continuous and order-preserving. Hence $X_{\frac{p}{q}}$ is a closed subset of the interval $[0, \boldsymbol{\omega}_{\frac{p}{q}}]$. And finally, properties of the tree $T_{\frac{p}{q}}$ imply that $[0, \boldsymbol{\omega}_{\frac{p}{q}}] \setminus X_{\frac{p}{q}}$ cannot contain any non-empty open interval. From the same properties (see Appendix A) we deduce:

PROPOSITION **9** *The set of real numbers having more than one expansion is infinite countable.*

REMARK **4** *In contrast with the classical representations of reals, the finite prefixes of a* $\frac{p}{q}$*-expansion of a real number, completed by zeroes,* are not $\frac{p}{q}$*-expansions of real numbers (though they can be given a value by the function* π *of course), that is to say, if a finite word* w *is in* $L_{\frac{p}{q}} \setminus 0^*$*, then the word* $w0^\omega$ *does not belong to* $W_{\frac{p}{q}}$*.*

PROPOSITION **10** *If* $q > 1$ *then no element of* $W_{\frac{p}{q}}$ *is eventually periodic, but* 0^ω*.*

COROLLARY **11** *If* $p \geqslant 2q - 1$ *then no real number can have three different expansions.*

3.2 Computation of the $\frac{p}{q}$-expansion of the real numbers

A striking feature of the $\frac{p}{q}$-expansion of the integers is that it is computed least significant digit first, or from right to left. This is quite an accepted process, that becomes problematic when it comes to the reals and that you have to compute from right to left a representation which is *infinite to the right*[2]. Let x be in $[0, \boldsymbol{\omega}_{\frac{p}{q}}]$ and let $\langle x\rangle_{\frac{p}{q}} = \mathbf{a} = \boldsymbol{.}a_1a_2\cdots$ be a $\frac{p}{q}$-expansion

[2] As W. Allen said: "The infinity is pretty far, especially when you reach the end".

of x. Let M be a (large) positive integer. By definition $\langle (\frac{p}{q})^M x\rangle_{\frac{p}{q}} = a_1 \cdots a_M {\bullet} a_{M+1} a_{M+2} \cdots$ is a $\frac{p}{q}$-expansion of $(\frac{p}{q})^M x$. By definition also, $a_1 \cdots a_M$ is *the* $\frac{p}{q}$-expansion of the integer $\pi(a_1 \cdots a_M)$. Let us write $\rho_M(x)$ the integral part $\lfloor \pi({\bullet} a_{M+1} a_{M+2} \cdots)\rfloor$. It then holds

$$\left\lfloor \left(\frac{p}{q}\right)^M x \right\rfloor = \pi(a_1 \cdots a_M) + \rho_M(x) \qquad \text{with} \qquad 0 \leqslant \rho_M(x) < \frac{p-1}{p-q} .$$

From this we deduce an approximation process for the computation of the $\frac{p}{q}$-expansion of x:

(i) Take a large integer M, such that $\lfloor (\frac{p}{q})^M x\rfloor >> 1$.

(ii) Compute the $\frac{p}{q}$-expansion of $\lfloor (\frac{p}{q})^M x\rfloor - h$, for every integer h, $0 \leqslant h < \frac{p-1}{p-q}$.

(iii) The maximal common prefix of all these expansions is the beginning of all the $\frac{p}{q}$-expansions of x.

To get longer prefixes one has to make again the computation with an M' larger than M, but it is not possible to know in advance how large has to be this M' in order to get a better approximation. In the case that $p \geqslant 2q - 1$, the only possible values of h are 0 and 1, and one can keep track of the two approximations, see the Appendix.

3.3 On the computation and the value of $\boldsymbol{\omega}_{\frac{p}{q}}$

Let us come back to the partial functions $\{\tau_a \mid a \in A\}$. If we start from 0 and, at each step, the greatest a such that $\tau_a(z) \in \mathbb{N}_+$ is chosen, we obtain the maximal word in the lexicographic order for each length. If at each step we choose the smallest a such that $\tau_a(z) \in \mathbb{N}_+$, we obtain the minimal word in the lexicographic order. By construction, the maximal (resp. minimal) word of length n is a prefix of the maximal (resp. minimal) word of length $n+1$, so when n goes to infinity we obtain two infinite words, the maximal one $\mathbf{t}_{\frac{p}{q}}$ and the minimal one $\mathbf{g}_{\frac{p}{q}} = (g_i)_{i \geqslant 1}$, which belong to $I_{p/q}$.

Example **2** *For* $\frac{p}{q} = \frac{3}{2}$, $\mathbf{g}_{\frac{p}{q}} = 2101100011010011010100110 \cdots$. ◇

Indeed the two words $\mathbf{t}_{\frac{p}{q}}$ and $\mathbf{g}_{\frac{p}{q}}$ are very similar and $\mathbf{g}_{\frac{p}{q}}$ is a plain "transliteration" of $\mathbf{t}_{\frac{p}{q}}$ as we see now. Denote by $\bar{d}$ the signed digit $-d$.

Proposition **12** *The digit-wise difference between the shifted infinite maximal word* $0\mathbf{t}_{\frac{p}{q}}$ *and the infinite minimal word* $\mathbf{g}_{\frac{p}{q}}$ *is equal to* $\bar{q}(p-q)^\omega$ *.*

For $n \geqslant 1$ let $G_n = \pi(g_1 \cdots g_n)$. Note that, for $n \geqslant 2$, $g_n \in \{0, \ldots, q-1\}$. We then have the following result.

Proposition **13** *The sequence* $(G_n)_{n \geqslant 1}$ *satisfies the following recurrence*

$$G_n = \lceil \frac{p}{q} G_{n-1} \rceil$$

with $G_1 = 1$*.*

Observe that, for $\frac{p}{q} = \frac{3}{2}$, the sequence $(G_n)_{n \geqslant 1}$ is exactly Sequence A061419 in [19].

Let $\boldsymbol{\gamma}_{\frac{p}{q}} = \pi({\bullet}\mathbf{g}_{\frac{p}{q}}) = \frac{p}{q}\boldsymbol{\omega}_{\frac{p}{q}}$. Note that the constant $\boldsymbol{\gamma}_{\frac{p}{q}}$ has two $\frac{p}{q}$-expansions, namely $\mathbf{g}_{\frac{p}{q}}$ and $0\mathbf{t}_{\frac{p}{q}}$. The following result is straightforward.

Proposition **14** *For $n \geqslant 1$,*

$$G_n = \lfloor \boldsymbol{\gamma}_{\frac{p}{q}} \Big(\frac{p}{q}\Big)^n \rfloor - e_n$$

where e_n is an integer $0 \leqslant e_n < (q-1)/(p-q)$.

Corollary **15** *If $p \geqslant 2q-1$ then, for $n \geqslant 1$, $G_n = \lfloor \boldsymbol{\gamma}_{\frac{p}{q}} \big(\frac{p}{q}\big)^n \rfloor$.*

Remark **5** *If $p \geqslant 2q-1$ then for each $n \geqslant 1$, the digit g_n of the $\frac{p}{q}$-expansion of $\boldsymbol{\gamma}_{\frac{p}{q}}$ can be easily obtained as follows:*

1. *compute $G_n = \lceil \frac{p}{q} G_{n-1} \rceil$*
2. *$g_n = qG_n \mod p$.*

In [15] the authors study the following problem. Let $\alpha > 1$ and let $f(x) = \lceil \alpha x \rceil$. The iterates $f_n = f_n(\alpha)$ of f are defined by $f_{n+1} = \lceil \alpha f_n \rceil$, $f_0 = 1$ for $n \geqslant 0$. They show that if $\alpha \geqslant 2$ or $\alpha = 2 - 1/q$ for some integer $q \geqslant 2$ then for all $n \geqslant 0$, $f_n = \lfloor H(\alpha)\alpha^n \rfloor$ where $H(\alpha)$ is a constant.

Taking $\alpha = \frac{p}{q}$, with $p \geqslant 2q-1$, we obtain the same result as in [15], with $H(\frac{p}{q}) = \boldsymbol{\omega}_{\frac{p}{q}}$. Our method is not an "independent" way of computing this constant, as was called for in [15], but, using the $\frac{p}{q}$-expansion of $\boldsymbol{\omega}_{\frac{p}{q}}$ gives an easier computation.

In the case where $q = p-1$ (the Josephus case), the constant $\boldsymbol{\omega}_{\frac{p}{q}}$ is the constant $K(p)$ in [15]. In this case the integer e_n of Proposition 14 is less than $p-2$, and this is the same bound as in [15].

Example **3** *For $\frac{p}{q} = \frac{3}{2}$, our constant $\boldsymbol{\omega} = \boldsymbol{\omega}_{\frac{p}{q}}$ is the constant $K(3)$ already discussed in [15, 10, 20]. Its decimal expansion $\langle \boldsymbol{\omega} \rangle_{10} = 1{\cdot}62227050288476731595695098 2\cdots$ is Sequence A083286 in [19]. We have of course $\langle \boldsymbol{\omega} \rangle_{\frac{3}{2}} = {\cdot}21222111221211221212 11221\cdots$.* ◇

4 On the fractional part of the powers of rational numbers

The distribution of the fractional part of the powers of a rational number is a long standing and deeply intriguing problem. It is well known[3] that for allmost all real number $\theta > 1$ the sequence $\{\theta^n\}$ is uniformely distributed in $[0,1]$, but very few results are known for specific value of θ. One of these is that *if θ is a Pisot number*, then the above sequence has only 0 and 1 as limit points. The distribution of $\left\{(\frac{p}{q})^n\right\}$ for coprime positive integers $p > q \geqslant 2$ remains an unsolved problem. Experimental results shows that this distribution looks more "chaotic" than the distribution of the fractional part of the powers of a transcendental number like e or π (*cf.* [22]). Vijayaraghavan [21] showed that the sequence has infinitely many limits points.

The next step in attacking this problem has been *to fix the rational* $\frac{p}{q}$ and to study the distribution of the sequence

$$f_n(\xi) = \left\{ \xi \left(\frac{p}{q}\right)^n \right\}$$

[3]This presentation of the problem is based on the introduction of [3].

according to the value of the real number ξ. Once again, the sequence $f_n(\xi)$ is uniformely distributed for almost all $\xi > 0$, but nothing is known for specific value of ξ.

In the search for ξ's for which the sequence $f_n(\xi)$ is *not uniformely distributed*, Mahler considered those for which the sequence is eventually contained in $[0, \frac{1}{2}[$. Let us generalize Mahler's notation as follow: let I be a (strict) subset of $[0,1[$ — indeed I will be a finite union of semi-closed intervals — and write:

$$\mathbf{Z}_{\frac{p}{q}}(I) = \{\xi \in \mathbb{R} \mid \left\{\xi \left(\frac{p}{q}\right)^n\right\} \quad \text{stays eventually in } I\ \} \ .$$

Mahler [14] proved that $\mathbf{Z}_{\frac{3}{2}}\left([0, \frac{1}{2}[\right)$ is at most countable but left open the problem to decide whether it is empty or not. Mahler's work has been developped in two directions: the search for subsets I as large as possible such that $\mathbf{Z}_{\frac{p}{q}}(I)$ is empty and conversely the search for subsets I as small as possible such that $\mathbf{Z}_{\frac{p}{q}}(I)$ is non-empty.

Along the first line, a remarkable progress has been made by Flatto *et al.* ([6]) who proved that the set of reals s such that $\mathbf{Z}_{\frac{p}{q}}\left([s, s+\frac{1}{p}[\right)$ is empty is *dense* in $[0, 1-\frac{1}{p}]$, and recently Bugeaud [3] proved that its complement is of Lebesgue measure 0. Along the other line, Pollington [17] showed that $\mathbf{Z}_{\frac{3}{2}}\left([\frac{4}{65}, \frac{61}{65}[\right)$ is non-empty. Our contribution to the problem, that we can eventually present, yields an improvement of this result.

In what follows, we suppose, once again, that $p \geqslant 2q-1$. For every $\frac{p}{q}$ we define — see the Appendix — a subset $Y_{\frac{p}{q}}$ of $[0,1[$ which is the union of q intervals of length $\frac{1}{p}$ as

$$Y_{\frac{p}{q}} = \bigcup_{0 \leqslant c \leqslant q-1} [\frac{1}{p}k_c, \frac{1}{p}(k_c+1)[$$

where $k_c \in \{0, \ldots, p-1\}$ and $qk_c = c \mod p$. For instance:

$$Y_{\frac{3}{2}} = [0, \frac{1}{3}[\cup[\frac{2}{3}, 1[$$

Theorem **16** *A positive real ξ belongs to $\mathbf{Z}_{\frac{p}{q}}\left(Y_{\frac{p}{q}}\right)$ if and only if ξ has two $\frac{p}{q}$-expansions.*

This implies in particular that $\mathbf{Z}_{\frac{p}{q}}\left(Y_{\frac{p}{q}}\right)$ is infinite countable. It implies also the following statement:

Corollary **17** *For any $\varepsilon > 0$, there exists a rational $\frac{p}{q}$ and a subset $Y_{\frac{p}{q}} \subseteq [0,1[$ of Lebesgue measure ε such that $\mathbf{Z}_{\frac{p}{q}}\left(Y_{\frac{p}{q}}\right)$ is infinite countable.*

The proof of Theorem 16 is sketched in the Appendix. It relies of course on the characterization of double expansions by Theorem 19 but it is mainly based on two supplementary ideas. The first one is the definition of a representation that we call the *companion* representation, and which can be computed with any prescribed precision (provided we can compute in $\mathbb{Q}$ with the same precision). It amounts indeed to the incremental use of the algorithm described in the previous section. The price we have to pay for the fact we indeed compute the companion representation *from left to right* is that we use a larger alphabet of digits, containing *negative digits*, exactly as the Avizienis representation of reals allows to perform addition from left to right [2]. The second idea is to put into correspondance any $\frac{p}{q}$-expansion of a real and its companion representation in the (infinite) computations of a transducer that proves to be the transposed automaton of the converter that we have defined in Section 3.

References

[1] S. Akiyama, Self affine tiling and Pisot numeration system, in *Number theory and its applications*, K. Györy and S. Kanemitsu editors, Kluwer (1999) 7–17.

[2] A. Avizienis, Signed-digit number representations for fast parallel arithmetic, *IRE Transactions on electronic computers* **10** (1961) 389–400.

[3] Y. Bugeaud, Linear mod one transformations and the distribution of fractional parts $\{\xi(\frac{p}{q})^n\}$, *Acta Arith.* **114** (2004) 301–311.

[4] A. Cauchy, Sur les moyens d'éviter les erreurs dans les calculs numériques, *C.R. Acad. Sc. Paris* série I **11** (1840) 789–798.

[5] S. Eilenberg, *Automata, Languages and Machines*, Vol. A, Academic Press (1974).

[6] L. Flatto, J.C. Lagarias and A.D. Pollington, On the range of fractional parts $\{\xi(\frac{p}{q})^n\}$, *Acta Arith.* **70** (1995) 125–147.

[7] A.S. Fraenkel, Systems of numeration, *Amer. Math. Monthly* **92** (1985) 105–114.

[8] Ch. Frougny, Representation of numbers and finite automata, *Math. Sys. Th.* **25** (1992) 37–60.

[9] Ch. Frougny and J. Sakarovitch, Automatic conversion from Fibonacci representation to representation in base phi, and a generalization, *Internat. J. Algebra Comput.* **9** (1999) 351–384.

[10] L. Halbeisen and N. Hungerbüler, The Josephus problem, `http://citeseer.nj.nec.com/23586.html`.

[11] J.E. Hopcroft and J.D. Ullman, *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesley (1979).

[12] D. Knuth, *The Art of Computer Programming*, Addison Wesley (1969).

[13] M. Lothaire, *Algebraic Combinatorics on Words*, Cambridge University Press (2002).

[14] K. Mahler, An unsolved problem on the powers of 3/2, *J. Austral. Math. Soc.* **8** (1968) 313–321.

[15] A. Odlyzko and H. Wilf, Functional iteration and the Josephus problem, *Glasgow Math. J.* **33** (1991) 235–240.

[16] W. Parry, On the β-expansions of real numbers, *Acta Math. Acad. Sci. Hung.*, **11** 401–416.

[17] A.D. Pollington, Progressions arithmétiques généralisées et le problème des $(3/2)^n$. *C.R. Acad. Sc. Paris* série I **292** (1981) 383–384.

[18] A. Rényi, Representations for real numbers and their ergodic properties, *Acta Math. Acad. Sci. Hung.* **8** (1957) 477–493.

[19] N.J.A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, `http://www.research.att.com/~njas/sequences/`.

[20] R. Stephan, On a sequence related to the Josephus problem, `http://arxiv.org/abs/math.CO/0305348` (2003).

[21] T. Vijayaraghavan, On the fractional parts of the powers of a number, I, *J. London Math. Soc.* **15** (1940), 159–160.

[22] E. Weisstein, Power fractional parts, from MathWorld, `http://mathworld.wolfram.com/PowerFractionalParts.html`

A Minimal and maximal paths in $T_{\frac{p}{q}}$

Let v be any node in $T_{\frac{p}{q}}$. The *maximal* (resp. *minimal*) path rooted in v is the infinite path in $T_{\frac{p}{q}}$ obtained by choosing at each step the maximal (resp. minimal) digit possible. (see Section 2.2). It is denoted by Maxpath(v) (resp. Minpath(v)). The label of a maximal (resp. minimal) path is said to be a *maximal* (resp. *minimal*) word. A word in $W_{\frac{p}{q}}$ is said to be *eventually maximal* (resp. *eventually minimal*) if it has a suffix which is a maximal (resp. minimal) word. The following is somehow a characterization of minimal and maximal words.

LEMMA **18** *Let v be any node in $T_{\frac{p}{q}}$, which is branching. Suppose that there are two infinite branches $v \xrightarrow{b_1} v_1 \xrightarrow{b_2} v_2 \cdots$ and $v \xrightarrow{d_1} w_1 \xrightarrow{d_2} w_2 \cdots$. Then the following are equivalent:*
(i) *for every $i \geqslant 1$, $\lambda(v_i) = \lambda(w_i) + 1$;*
(ii) *$b_1 - d_1 = q$ and for each $i \geqslant 2$, $b_i - d_i = q - p$;*
(iii) *$b_1 - d_1 = q$, $v_1 v_2 \cdots =$ Minpath(v_1), for each $i \geqslant 2$, $b_i \in \{0, \ldots, q-1\}$, $w_1 w_2 \cdots =$ Maxpath(w_1) and for each $i \geqslant 2$, $d_i \in \{p-q, \ldots, p-1\}$.*
In that case

$$\pi(.b_1 b_2 \cdots) = \pi(.d_1 d_2 \cdots).$$

THEOREM **19** *Let x be in $[0, \boldsymbol{\omega}_{\frac{p}{q}}]$. The following are equivalent:*
(i) *x has more than one expansion;*
(ii) *x has an expansion which is an eventually minimal word;*
(iii) *x has an expansion which is eventually written on the alphabet $\{0, \ldots, q-1\}$;*
(iv) *x has an expansion which is an eventually maximal word;*
(v) *x has an expansion which is eventually written on the alphabet $\{p-q, \ldots, p-1\}$.*

B Sketch of the proof of Theorem 16

B.1 The companion representation

Let $C = \{-(q-1), \ldots, 0, 1, \ldots, (p-1)\}$ be an augmented digit alphabet ($C = A$ if $q = 1$, $C = A \cup \{-(q-1), \ldots, \bar{1}\}$ otherwise).

Let $h\colon \mathbb{R}_+ \to \mathbb{Z}$ be the function defined by:

$$h(z) = q\left\lfloor (\frac{p}{q}) z \right\rfloor - p \left\lfloor z \right\rfloor$$

The function h is periodic of period q and it holds:

PROPERTY **20** *For all z in $\mathbb{R}_+$, $h(z)$ belongs to C.*

Let us write now, for every n in $\mathbb{N}$:

$$h_n(z) = h\left((\frac{p}{q})^{n-1} z \right) = c_n$$

which, in turn, defines a map: $\varphi(z) \colon \mathbb{R}_+ \to C$ by:

$$\varphi(z) = \mathbf{c} = \mathbf{.}c_1 c_2 \cdots c_n \cdots \quad .$$

The sequence $\varphi(z) = \mathbf{c} = \mathbf{.}c_1 c_2 \ldots c_n \cdots$ is called the *companion representation* of z. If $q = 1$, $h_n(z)$ is precisely the n-th digit after the decimal point in the expansion of z in base p. A property that can be generalized in the $\frac{p}{q}$-system by easy calculations and stated as the following:

PROPERTY **21** *For all z in $\mathbb{R}_+$, $\varphi(z)$ is a $\frac{p}{q}$-representation of $\{z\} = z - \lfloor z \rfloor$, the fractional part of z.*

And this implies that for every k in $\mathbb{N}$, $\mathbf{.}c_k c_{k+1} c_{k+2} \ldots$ is a $\frac{p}{q}$-representation of $\{(\frac{p}{q})^{k-1} z\}$.

B.2 The co-converter

Let x be in $[0, \boldsymbol{\omega}_{\frac{p}{q}}]$. Let $\langle x \rangle_{\frac{p}{q}} = \mathbf{a} = \mathbf{.}a_1 a_2 \cdots$ be a $\frac{p}{q}$-expansion of x and $\varphi(z) = \mathbf{c} = \mathbf{.}c_1 c_2 \cdots$ its companion representation. Let us evaluate $h_n(z)$ using the techniques and notation of Section 3.2. If we bring

$$\left\lfloor (\frac{p}{q})^{n-1} x \right\rfloor = \pi(a_1 \cdots a_{n-1}) + \rho_{n-1}(x) \qquad \text{and} \qquad \left\lfloor (\frac{p}{q})^{n} x \right\rfloor = \pi(a_1 \cdots a_{n}) + \rho_{n}(x)$$

$$\text{into} \qquad c_n = h\left((\frac{p}{q})^{n-1} z \right) = q \left\lfloor (\frac{p}{q})^{n} x \right\rfloor - p \left\lfloor (\frac{p}{q})^{n-1} x \right\rfloor$$

we get, since by (1) we have $q\,\pi(a_1 \cdots a_n) = p\,\pi(a_1 \cdots a_{n-1}) + a_n$, the equality:

$$c_n + p\,\rho_{n-1}(x) = a_n + q\,\rho_n(x) \tag{3}$$

There are a finite number of possible values for $\rho_n(x)$ (indeed $\left\lfloor \frac{p-1}{p-q} \right\rfloor$), and (3) can be seen as the definition of a transducer $\mathcal{A}_{\frac{p}{q}}$: a transition labelled by (c_n, a_n) goes from the state $\rho_{n-1}(x)$ to the state $\rho_n(x)$. We recognize, by comparison with (2), that $\mathcal{A}_{\frac{p}{q}}$ is the transposed automaton of the converter $\mathcal{C}_C$ that we have described at Section 2.3. The transducer $\mathcal{A}_{\frac{p}{q}}$ is *co-sequential* (that is *input co-deterministic*) and in substance we have proved:

PROPOSITION **22** *Let x be a real in $[0, \boldsymbol{\omega}_{\frac{p}{q}}]$, $\mathbf{c}$ its companion representation and $\mathbf{a}$ a $\frac{p}{q}$-representation of x. Then $(\mathbf{c}, \mathbf{a})$ is the label of an infinite path that begins in the state $\rho_0(x)$ in the transducer $\mathcal{A}_{\frac{p}{q}}$.*

A case that is especially interesting is when $p \geqslant 2q - 1$: $\mathcal{A}_{\frac{p}{q}}$ has then 2 states. Let us write C as the union $C = C_1 \cup C_2 \cup C_3$ with

$$C_1 = \{-(q-1), \ldots, -1\}, \qquad C_2 = \{0, \ldots, q-1\}, \qquad \text{and} \quad C_3 = \{q, \ldots, p-1\} \quad .$$

The transducer $\mathcal{A}_{\frac{p}{q}}$ is drawn at Figure 3.

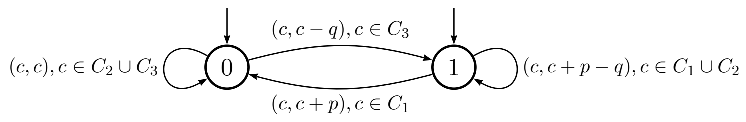

Figure 3: The transducer $\mathcal{A}_{\frac{p}{q}}$ when $p \geqslant 2q-1$

B.3 The companion representation of reals with two $\frac{p}{q}$ expansions

Let us now suppose for the rest of the appendix that $p \geqslant 2q-1$. The next step in the proof of Theorem 16 is the following:

PROPOSITION **23** *A real x has two $\frac{p}{q}$-expansions if and only if its companion representation is eventually in $C_2 = \{0, \ldots, q-1\}$.*

*Proof.*The condition is necessary for if x has two $\frac{p}{q}$-expansions $\mathbf{a}'$ and $\mathbf{a}''$, then $(\mathbf{c}, (\mathbf{a}', \mathbf{a}''))$ must be the label of an infinite path in the *square of the transducer* $\mathcal{A}_{\frac{p}{q}}$ that goes *outside of the diagonal*; this implies — as seen easily on Figure 4 — that $\mathbf{c}$ is eventually in C_2.

Let $\mathbf{c}$ and $\mathbf{a}$ be the companion representation and a $\frac{p}{q}$-representation respectively of a real x. By Proposition 22, $(\mathbf{c}, \mathbf{a})$ is the label of an infinite path starting in s in $\mathcal{A}_{\frac{p}{q}}$. Suppose that c_n is the last digit of $\mathbf{c}$ not in C_2 and, by way of example, that it belongs to C_3. Then (c_n, a_n) is the label of a transition that leaves state 0. If $a_n = c_n$, then the infinite word $\mathbf{a}'$ defined by $a'_i = a_i$ for $0 \leqslant i < n$, $a'_n = a_n - q$, and $a'_i = a_i + p - q$ for $n < i$, is such that $(\mathbf{c}, \mathbf{a}')$ is the label of an infinite path in $\mathcal{A}_{\frac{p}{q}}$ with s as initial state — which implies that $\mathbf{a}'$ is a $\frac{p}{q}$-representation of x — and it can be verified that $\mathbf{a}'$ belongs to $W_{\frac{p}{q}}$, which shows that it is a second $\frac{p}{q}$-expansion of x. ∎

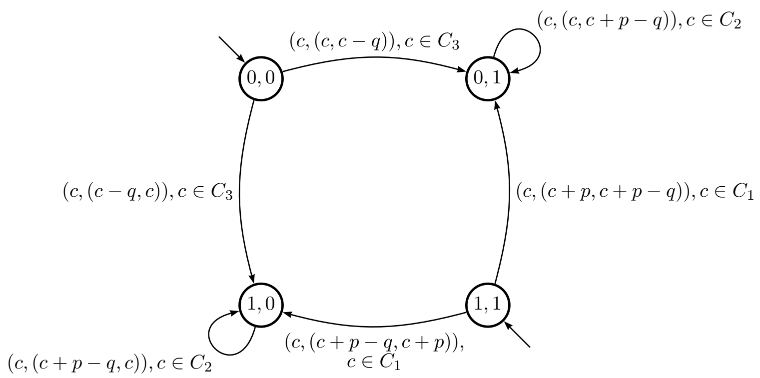

Figure 4: The square of $\mathcal{A}_{\frac{p}{q}}$ (outside of the diagonal)

B.4 Inversion of the function h

The final step consists in the description of the inverse of the function h:

LEMMA **24** *Let c be in $C_2 = \{0, \dots, q-1\}$. Then $h(x) = c$ if and only if $\left\{\frac{x}{q}\right\} \in [\frac{1}{p}k_c, \frac{1}{p}(k_c + 1)[$ where $0 \leqslant k_c < p$ and $q\, k_c = c \mod p$.*

From Proposition 23 follows that a real x has two $\frac{p}{q}$-expansions if and only if there exists $M > 0$ such that for any $n > M$,

$$\{\left(\frac{p}{q}\right)^n \frac{x}{q}\} \in Y_{\frac{p}{q}} = \bigcup_{0 \leqslant c \leqslant q-1} [\frac{1}{p}k_c, \frac{1}{p}(k_c + 1)[$$

where $k_c \in \{0, \dots, p-1\}$ and $qk_c = c \mod p$, and this concludes the proof of Theorem 16.