

On the Representation of Numbers in a Rational Base

Shigeki Akiyama* Christiane Frougny† Jacques Sakarovitch‡

DEDICATED TO OUR FRIEND ANTONIO RESTIVO
ON THE OCCASION OF HIS 60TH BIRTHDAY

Summary A new method for representing positive integers and real numbers in a rational base is considered. It amounts to computing the digits from right to left, least significant first. Every integer has a unique such expansion. The set of expansions of the integers is not a regular language but nevertheless addition can be performed by a letter-to-letter finite right transducer. Every real number has at least one such expansion and a countable infinite set of them have more than one. We explain how these expansions can be approximated and characterize the expansions of reals that have two expansions.

These results are not only developed 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 paper¹, we introduce and study a new method for representing positive integers and real numbers in the base $\frac{p}{q}$, where $p > q \geq 2$ are coprime integers.² 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 standpoint as well as for improving computation algorithms. It is worth first (briefly) recalling 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.

Many non-standard numeration systems have been considered, [13, Vol. 2, Chap. 4] or [14, Chap. 7], for instance, give extensive references. Representation in integer base with signed digits was popularized in computer arithmetic by Avizienis [3] and can be found earlier in a work of Cauchy [6]. When the base is a real number $\beta > 1$, any non-negative real number is given an expansion on the canonical alphabet $\{0, 1, \dots, \lfloor \beta \rfloor\}$ by the greedy algorithm of Rényi [19]; a number may have several β -representations on the canonical alphabet, but the greedy one is the greatest in the lexicographical order. The set of greedy β -expansions of numbers of $[0, 1[$ is shift-invariant, 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 [17, 14].

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

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

‡LTCl, UMR 5141, CNRS / ENST, sakarovitch@enst.fr

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

²To keep the length into reasonable bounds for the proceedings of a conference, proofs are only sketched or even sometimes omitted. A full paper will soon be submitted to a journal.

When β is a Pisot number³, the β number system shares many properties with the integer base case: the set of greedy representations is recognizable by a finite automaton; the conversion between two alphabets of digits (in particular addition) is realized by a finite transducer [10].

Here, we first define the $\frac{p}{q}$ -*expansion* of an integer N : it is a way of writing N in the base $\frac{p}{q}$ by an algorithm which produces the digits from right to left, that is *least significant digits first*. We prove:

THEOREM 1 *Every non-negative integer N has a $\frac{p}{q}$ -expansion which is an integer representation. It is the unique finite $\frac{p}{q}$ -representation of N .*

The $\frac{p}{q}$ -expansions are not the $\frac{p}{q}$ -representations that would be obtained by the classical “greedy algorithm” in base $\frac{p}{q}$. They are written on the alphabet $A = \{0, 1, \dots, p-1\}$, but not every word of A^* is admissible. These $\frac{p}{q}$ -expansions share some properties with the expansions in an integer base — digit set conversion is realized by a finite transducer for instance — and are completely different as far as other 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). To some extent, the study and understanding of this set of words $L_{\frac{p}{q}}$ is what this paper is about.

By construction, the set $L_{\frac{p}{q}}$ is prefix-closed and any of its elements can be extended (to the right) in $L_{\frac{p}{q}}$. Hence, $L_{\frac{p}{q}}$ is the set of labels of the finite paths in an *infinite subtree* $T_{\frac{p}{q}}$ of the infinite full p -ary tree of the free monoid A^* . The tree $T_{\frac{p}{q}}$ contains a maximal infinite word $\mathbf{t}_{\frac{p}{q}}$ — maximal in the lexicographic ordering — whose numerical value is $\omega_{\frac{p}{q}}$. We define the *admissible $\frac{p}{q}$ -expansions* of real numbers to be the set $W_{\frac{p}{q}}$ of infinite words that label the *infinite paths* of $T_{\frac{p}{q}}$ and we prove:

THEOREM 2 *Every real in $[0, \omega_{\frac{p}{q}}]$ has exactly one $\frac{p}{q}$ -expansion, but for an infinite countable number of them which have more than one such expansion.*

If $p \geq 2q - 1$, then no real has more than two $\frac{p}{q}$ -expansions. It is noteworthy as well that no $\frac{p}{q}$ -expansion is eventually periodic and thus in particular — and in contrast with the expansion of reals in an integer base — no $\frac{p}{q}$ -expansion ends with 0^ω or, which is the same, is finite. This is a very remarkable feature of the $\frac{p}{q}$ number system for reals and we explain how the $\frac{p}{q}$ -expansion of a real number can be computed (in fact approximated).

We shall give here two examples of the relations of the $\frac{p}{q}$ -expansions of reals with other problems in combinatorics and number theory. The first one is the so-called “Josephus problem” in which a certain constant $K(p)$ is defined (*cf.* [16, 11, 22]) which is a special case of our constant $\omega_{\frac{p}{q}}$ (with $q = p - 1$) and this definition yields a new method for computing $K(p)$.

The connection with the second problem, namely the distribution of the powers of a rational number modulo 1, is even more striking. It requires to be presented that the framework of this long standing and deeply intriguing problem be set.⁴

Koksma proved that for almost every real number $\theta > 1$ the sequence $\{\theta^n\}$ is uniformly distributed in $[0, 1]$, but very few results are known for specific values of θ .

³An algebraic integer whose Galois conjugates are all less than 1 in modulus

⁴This presentation is based on the introduction of [5]. The fractional part of a number x is denoted by $\{x\}$.

One of these is that if θ is a Pisot number, then the above sequence converges to 0 if we identify $[0, 1)$ with \mathbb{R}/\mathbb{Z} .

The distribution of $\left\{ \left(\frac{p}{q} \right)^n \right\}$ for coprime positive integers $p > q \geq 2$ remains an unsolved problem. Experimental results show 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. [24]). Vijayaraghavan [23] 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

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

according to the value of the real number ξ . Once again, the sequence $f_n(\xi)$ is uniformly 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 uniformly distributed*, Mahler considered those for which the sequence is eventually contained in $[0, \frac{1}{2}[$. Mahler's notation is generalized as follow: let I be a (strict) subset of $[0, 1[$ — indeed I will be a finite union of semi-closed intervals — and write:

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

Mahler [15] proved that $\mathbf{Z}_{\frac{3}{2}}([0, \frac{1}{2}[$) is at most countable but left open the problem to decide whether it is empty or not. Mahler's work has been developed in two directions: the search for subsets I as large as possible such that $\mathbf{Z}_{\frac{p}{q}}(I)$ is empty — which amounts to proving that $f_n(\xi)$ is never too much unevenly distributed — and conversely the search for subsets I as small as possible such that $\mathbf{Z}_{\frac{p}{q}}(I)$ is non-empty — which corresponds to establish the existence of ξ for which $f_n(\xi)$ is as unevenly distributed as possible.

Along the first line, a remarkable progress has been made by Flatto *et al.* ([8]) who proved that the set of reals s such that $\mathbf{Z}_{\frac{p}{q}}([s, s + \frac{1}{p}[$) is empty is *dense* in $[0, 1 - \frac{1}{p}]$, and this has been even improved by Bugeaud [5] who proved that its complement is of Lebesgue measure 0. Along the other line, Pollington [18] showed that $\mathbf{Z}_{\frac{3}{2}}([\frac{4}{65}, \frac{61}{65}[$) is non-empty.

Our contribution to the problem can be seen as an improvement of this result.

THEOREM 3 *If $p \geq 2q - 1$, there exists a subset $Y_{\frac{p}{q}}$ of $[0, 1[$, of Lebesgue measure $\frac{q}{p}$, such that $\mathbf{Z}_{\frac{p}{q}}(Y_{\frac{p}{q}})$ is countable infinite.*

The elements of $\mathbf{Z}_{\frac{p}{q}}(Y_{\frac{p}{q}})$ are indeed the reals which have *two $\frac{p}{q}$ -expansions* and this is the reason why the consideration of the $\frac{p}{q}$ number system allowed to make some progress in Mahler's problem.

*

In conclusion, we have introduced and studied here a fascinating family of sets of words which can be seen from many sides, which raises still many difficult questions and

whose further study will certainly mix techniques from word combinatorics, automata theory, and number theory.

In order to keep the proceedings into reasonable length bounds we do not include in this extended abstract any preliminary for definition nor notation on (infinite) words and automata but rather follow [14, 7, 12] and we only give sketch of proofs, when we give any. A more complete version will be to be found elsewhere in a journal ([2]).

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

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

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

We will say that U , together with the alphabet $A = \{0, \dots, p-1\}$, 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 digits in D indexed by a section of \mathbb{Z} : $(d_i)_{k \geq i \geq -\infty}$, such that:

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

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* π :

$$\pi: D^{\mathbb{Z}} \longrightarrow \mathbb{R}, \quad \mathbf{d} = \{d_i\}_{k \geq i \geq -\infty} \longmapsto \pi(\mathbf{d}) = \sum_{-\infty}^{i=k} d_i u_i.$$

It is important to remark that this definition is *not* the classical one for the numeration system in base $\frac{p}{q}$: U is *not* the sequence of powers of $\frac{p}{q}$ but rather these powers *divided by* q and the digits *are not* the integers smaller than $\frac{p}{q}$ but rather the integers *whose quotient by* q is smaller than $\frac{p}{q}$. These two differences compensate each other and make the developments that follow possible.

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 $i \geq 0$, write

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

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 of N from right to left, that is to say *least significant digit first*, and stops for some k when $N_{k+1} = 0$. It holds $N = \sum_{i=0}^k a_i u_i$ and thus the word $a_k \cdots a_0$ is a $\frac{p}{q}$ -representation of N ; it will be called *the $\frac{p}{q}$ -expansion* of N and written $\langle N \rangle_{\frac{p}{q}}$. By convention the $\frac{p}{q}$ -expansion of 0 is the empty word.

It can be further proved that $\langle N \rangle_{\frac{p}{q}}$ is the unique *finite* $\frac{p}{q}$ -representation of N (under the condition that $a_k \neq 0$). We have thus established:

THEOREM 1 Every non-negative integer N has a $\frac{p}{q}$ -expansion which is an integer representation. It is the unique finite $\frac{p}{q}$ -representation of N . \blacksquare

EXAMPLE 1 Let $p = 3$ and $q = 2$, then $A = \{0, 1, 2\}$ — this will be our main running example. Table 1 in Fig.1 gives the $\frac{3}{2}$ -expansions of the twelve first integers (cf. Appendix A as well). \diamond

REMARK 1 If $q = 1$, i.e. the base is an integer, the above algorithm gives the same representation as the one given by the classical greedy (left-to-right) algorithm (see [19] or [14, Chapter 7]). On the contrary, if $q \neq 1$, the $\frac{p}{q}$ -expansion of n is not the representation obtained by the greedy algorithm but for $n = 0, 1, \dots, \lfloor \frac{p}{q} \rfloor$. It follows then from Theorem 1 that no integer (but $0, 1, \dots, \lfloor \frac{p}{q} \rfloor$) is given a finite representation by the greedy algorithm.

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.

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 . Let us call *word label* of a node s , and write $w(s)$, the label of *the* path from the root to s . By construction the label of s is $\pi(w(s))$. Let us denote $I_{\frac{p}{q}}$ the subtree of $T_{\frac{p}{q}}$ made of nodes whose word label does not begin with a 0. See Fig. 1 for (a part of) $T_{\frac{3}{2}}$ and $I_{\frac{3}{2}}$.

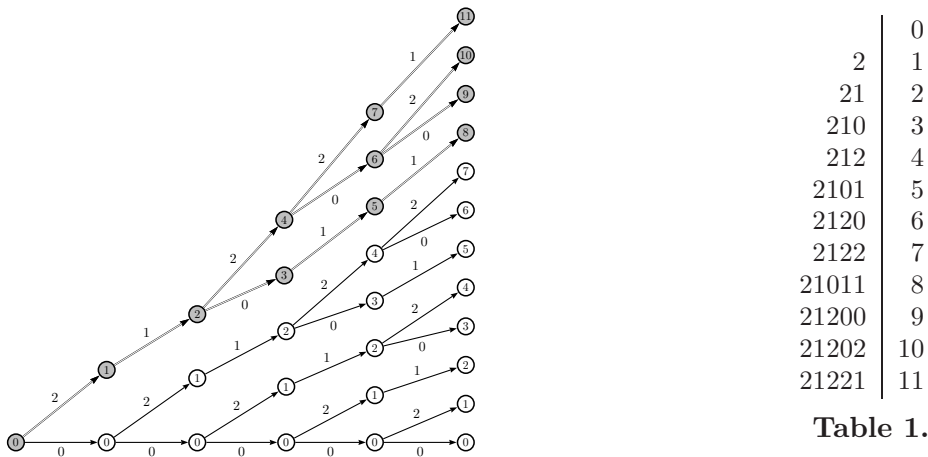


Figure 1: The tree $T_{\frac{3}{2}}$, the tree $I_{\frac{3}{2}}$ in grey and double edge, and a table

Among the digits that label *the edges* that start from a node with label N , there is a minimum one, $\text{minDigit}(N)$, which belongs to $\{0, \dots, q-1\}$ — and this is characteristic of a minimum digit — and there is a maximal one, $\text{MaxDigit}(N)$, which belongs to $\{p-q, \dots, p-1\}$ — and this is characteristic of a maximal digit. If $D = \text{MaxDigit}(N)$ and $d = \text{minDigit}(N+1)$, then $D = d + (p-q)$ and $\tau_d(N+1) = \tau_D(N) + 1$.

It follows that for every integer k , there exists an integer M_k such that the nodes of depth k in $T_{\frac{p}{q}}$ are labelled by the integers from 0 to M_k (see Section 3 for the computation of the M_k 's). And 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 4 *No two subtrees of $I_{\frac{p}{q}}$ are isomorphic.*

COROLLARY 5 *$L_{\frac{p}{q}}$ is not a regular language.*

LEMMA 6 *For any words x and y in $A^* \setminus 0^*$, $y \neq \varepsilon$, there exists an integer constant $K(x, y)$ such that, if xy^n belongs to $L_{\frac{p}{q}}$, then $n \leq K(x, y)$.*

COROLLARY 7 *$L_{\frac{p}{q}}$ is not a context-free language.*

PROPOSITION 8

Every w in A^k is the suffix of the $\frac{p}{q}$ -expansion of a unique integer n , $0 \leq n < p^k$.

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 $\chi_D: D^* \rightarrow A^*$ which commutes to the evaluation map π , that is a map which preserves the numerical value:

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

PROPOSITION 9 *For any alphabet D the conversion χ_D is realizable by a finite letter-to-letter sequential right 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) , with d in D and a in A , goes from h to k if and only if

$$qh + d = pk + a . \tag{1}$$

The set of accessible states is finite and this establishes Proposition 9. ■

The *integer addition* may be seen — after digit-wise addition — as a particular case of a digit-set conversion χ_D with $D = \{0, 1, \dots, 2(p-1)\}$ and Figure 2 shows the converter that realizes addition in the $\frac{3}{2}$ -system.

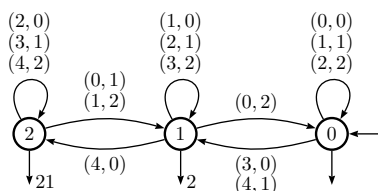


Figure 2: A converter for addition in the $\frac{3}{2}$ number system

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 [14, Ch. 7]).

REMARK 4 The conversion χ_D may also be extended to left infinite words and it follows from Proposition 9 that it is a continuous function. In particular, the odometer is continuous.

3 A remarkable set of infinite words

Let $W_{\frac{p}{q}}$ be the set of labels of infinite paths starting from the root 0 in $T_{\frac{p}{q}}$. It follows from the construction of $T_{\frac{p}{q}}$ by the partial functions $\{\tau_a \mid a \in A\}$ that two nodes with the same label are the root of the same subtree and from Proposition 4 that these subtrees are characteristic of the label (in fact, *any* infinite path from a node is characteristic of the label of the node).

This set $W_{\frac{p}{q}}$ will be used in the next section in order to define the representations of real numbers. We first try to describe it and to present some of its properties. In the previous sections, digits in a $\frac{p}{q}$ -representation were indexed from left to right by decreasing nonnegative integers for the “integer” part and by decreasing negative integers for the “decimal” part; as we shall now deal mainly with the “decimal” part of the representations, we find it much more convenient *to change the convention of indexing* and use the positive indices *after* the decimal point, in the increasing order.

We denote by $\text{MaxWord}(N)$ (resp. $\text{minWord}(N)$) the label of the infinite path that starts from a node with label N and that follows always the edges with the maximal (resp. minimal) digit label. From our observation in Section 2.2, it follows that for every N , the digit-wise difference between $\text{minWord}(N+1)$ and $\text{MaxWord}(N)$ is $(p-q)^\omega$.

Let us note $\mathbf{t}_{\frac{p}{q}} = \text{MaxWord}(0)$ and $\mathbf{g}_{\frac{p}{q}} = \text{minWord}(1)$. The word $\mathbf{t}_{\frac{p}{q}} = (t_i)_{i \geq 1}$ belongs to $\{p-q, \dots, p-1\}^{\mathbb{N}}$ and is the maximal word in $T_{\frac{p}{q}}$ (or $I_{\frac{p}{q}}$) in the lexicographic ordering. The word $\mathbf{g}_{\frac{p}{q}} = (g_i)_{i \geq 1}$ belongs to $\{0, \dots, q-1\}^{\mathbb{N}}$ and the word $q\mathbf{g}_{\frac{p}{q}}$ is the minimal word of $I_{\frac{p}{q}}$ in the lexicographic ordering. Remark that when $q=1$, $\text{MaxWord}(N) = (p-1)^\omega$, and $\text{minWord}(N) = 0^\omega$ for every N .

EXAMPLE 2 For $\frac{p}{q} = \frac{3}{2}$, $\mathbf{g}_{\frac{3}{2}} = 101100011010011010100110\dots$ ◇

For $n \geq 1$ let $G_n = \pi(qg_1 \cdots g_{n-1})$ (if $n=1$, $G_1 = \pi(q) = 1$), and $M_n = \pi(t_1 \cdots t_n)$. Of course $M_n = G_{n+1} - 1$. Let $\omega_{\frac{p}{q}} = \pi(\cdot \mathbf{t}_{\frac{p}{q}})$ the numerical value of the maximal infinite word and it holds $\gamma_{\frac{p}{q}} = \pi(\cdot q\mathbf{g}_{\frac{p}{q}}) = \pi(\cdot 0\mathbf{t}_{\frac{p}{q}}) = \frac{q}{p}\omega_{\frac{p}{q}}$. We then have the following results.

PROPOSITION 10 The sequence $(G_n)_{n \geq 1}$ satisfies the recurrence $G_n = \lceil \frac{p}{q} G_{n-1} \rceil$ with $G_1 = 1$, and for $n \geq 1$ there exists an integer e_n , $0 \leq e_n < (q-1)/(p-q)$, such that

$$G_n = \lfloor \gamma_{\frac{p}{q}} \left(\frac{p}{q}\right)^n \rfloor - e_n.$$

COROLLARY 11 If $p \geq 2q-1$ then, for $n \geq 1$, $G_n = \lfloor \gamma_{\frac{p}{q}} \left(\frac{p}{q}\right)^n \rfloor$.

REMARK 5 If $p \geq 2q-1$ then for each $n \geq 1$, the digit g_n is obtained as follows:

- (i) compute $G_{n+1} = \lceil \frac{p}{q} G_n \rceil$
- (ii) $g_n = qG_{n+1} \bmod p$.

The definition of the sequence G_n and the computation of $\mathbf{g}_q^{\frac{p}{q}}$ have been developed not only because they are important for the description of $T_q^{\frac{p}{q}}$ but also as they relate to a classical problem in combinatorics.

Inspired by the so-called ‘‘Josephus problem’’, Odlyzko and Wilf consider, for a real $\alpha > 1$, the iterates of the function $f(x) = \lceil \alpha x \rceil$: $f_0 = 1$ and $f_{n+1} = \lceil \alpha f_n \rceil$ for $n \geq 0$. They show (in [16]) that if $\alpha \geq 2$, or $\alpha = 2 - 1/q$ for some integer $q \geq 2$, then there exists a constant $H(\alpha)$ such that $f_n = \lfloor H(\alpha) \alpha^n \rfloor$ for all $n \geq 0$.

We have thus obtain the same result as in [16] for rational $\alpha = \frac{p}{q}$, with $p \geq 2q - 1$, and we find $H(\frac{p}{q}) = \frac{p}{q} \gamma_{\frac{p}{q}} = \omega_{\frac{p}{q}}$. Our method does not yield an ‘‘independent’’ way of computing this constant, as was called for in [16], but the writing of $\omega_{\frac{p}{q}}$ in the $\frac{p}{q}$ -system gives at least an easy algorithm.

In the case where $q = p - 1$ (the Josephus case), the constant $\omega_{\frac{p}{q}}$ is the constant $K(p)$ in [16]. In this case the integer e_n of Proposition 10 is less than $p - 2$, and this is the same bound as in [16].

EXAMPLE 3 For $\frac{p}{q} = \frac{3}{2}$, the constant $\omega_{\frac{3}{2}}$ is the constant $K(3)$ already discussed in [16, 11, 22]. Its decimal expansion $1.622270502884767315956950982 \dots$ is recorded as Sequence A083286 in [21]. Observe that, in the same case, the sequence $(G_n)_{n \geq 1}$ is Sequence A061419 in [21]. \diamond

As a consequence of Lemma 6, we have:

PROPOSITION 12 If $q > 1$ then no element of $W_q^{\frac{p}{q}}$ is eventually periodic, but 0^ω .

4 Representation of the reals

In order to distinguish between the integer and the real numerical values of some finite word, we put a decimal point \cdot indicating the position of the index 0.

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

We take as a definition that the $\frac{p}{q}$ -expansions of the real numbers are the elements of $W_q^{\frac{p}{q}}$. That is, an infinite word $\mathbf{a} = \{a_i\}_{i \geq 1}$ in $W_q^{\frac{p}{q}}$ is a $\frac{p}{q}$ -expansion of the real number x :

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

Let $X_q^{\frac{p}{q}} = \pi(W_q^{\frac{p}{q}})$. The elements of $X_q^{\frac{p}{q}}$ are non-negative real numbers less than or equal to $\omega_{\frac{p}{q}}$. Note that $\omega_{\frac{p}{q}} \leq \frac{p-1}{p-q}$. The fact that the $\frac{p}{q}$ number system may be used for representing the reals is expressed by the following statement.

THEOREM 2 Every real in $[0, \omega_{\frac{p}{q}}]$ has exactly one $\frac{p}{q}$ -expansion, but for an infinite countable number of them which have more than one such expansion.

The proof of the first part of Theorem 2, that is to say the proof that $X_q^{\frac{p}{q}} = [0, \omega_{\frac{p}{q}}]$, relies on three facts. First, $W_q^{\frac{p}{q}}$ is closed in the compact set $A^{\mathbb{N}}$, hence is compact. Second, the map $\pi: W_q^{\frac{p}{q}} \rightarrow X_q^{\frac{p}{q}}$ is continuous and order-preserving. Hence $X_q^{\frac{p}{q}}$ is a closed subset of the interval $[0, \omega_{\frac{p}{q}}]$. And finally, properties of the tree $T_q^{\frac{p}{q}}$ imply that

$[0, \omega_{\frac{p}{q}}] \setminus X_{\frac{p}{q}}$ cannot contain any non-empty open interval. From the same properties we deduce that real numbers having more than one expansion correspond to the branching nodes of $T_{\frac{p}{q}}$, hence the second part of Theorem 2.

REMARK 6 *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 $W_{\frac{p}{q}}$.*

COROLLARY 13 *If $p \geq 2q - 1$ then no real number can have three different expansions.*

4.2 The companion $\frac{p}{q}$ -representation and the co-converter

A 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 for integers, 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*⁵. This difficulty is somewhat overcome with the definition of *another* $\frac{p}{q}$ -representation for the reals; it can be computed with any prescribed precision (provided we can compute in \mathbb{Q} with the same precision) and somehow *from left to right*. The price we have to pay for this is that we use a larger alphabet of digits, containing *negative digits*, exactly as the Avizienis representation of reals allows to perform sequentially addition from left to right [3].

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

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

The function h is periodic of period q and for all z in \mathbb{R}_+ , $h(z)$ belongs to the digit alphabet

$$C = \{-(q-1), \dots, 0, 1, \dots, (p-1)\} .$$

(If $q = 1$, then $C = A$; $C = A \cup \{-(q-1), \dots, -1\}$ otherwise.)

Let us write now, for every n in \mathbb{N} , $c_n = h\left(\left(\frac{p}{q}\right)^{n-1}z\right)$ which, in turn, defines a map $\varphi(z): \mathbb{R}_+ \rightarrow C^{\mathbb{N}}$ by $\varphi(z) = \mathbf{c} = .c_1c_2 \cdots c_n \cdots$. If $q = 1$, c_n is precisely the n -th digit after the decimal point in the expansion of z in base p .

We call the sequence $\varphi(z)$ *the companion representation* of z , and we have:

PROPOSITION 14 *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 . ■*

Let x be in $[0, \omega_{\frac{p}{q}}]$. Let $\langle x \rangle_{\frac{p}{q}} = \mathbf{a} = .a_1a_2 \cdots$ be a $\frac{p}{q}$ -expansion of x and let $\varphi(x) = \mathbf{c} = .c_1c_2 \cdots$ its companion representation. Let us denote by $\rho_n(x)$ the integer part $\lfloor \pi(.a_{n+1}a_{n+2} \cdots) \rfloor$; easy calculation then shows:

$$c_n + p\rho_{n-1}(x) = a_n + q\rho_n(x) . \tag{2}$$

There are a finite number of possible values for $\rho_n(x)$ since $0 \leq \rho_n(x) < \frac{p-1}{p-q}$, and (2) can be seen as the definition of a (left) 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 (1), 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:

⁵As W. Allen said: "The infinite is pretty far, especially towards the end".

PROPOSITION 15 *Let x be a real in $[0, \omega_{\frac{p}{q}}]$, \mathbf{c} its companion representation and \mathbf{a} a $\frac{p}{q}$ -expansion 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 $\mathcal{A}_{\frac{p}{q}}$. \blacksquare*

If $p \geq 2q - 1$, the interesting case which we have already considered, $\mathcal{A}_{\frac{p}{q}}$ has then *only two states*. The transducer $\mathcal{A}_{\frac{3}{2}}$ is drawn at Figure 3.

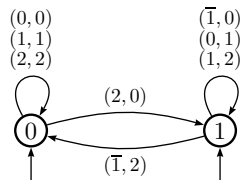


Figure 3: The transducer $\mathcal{A}_{\frac{3}{2}}$

The computation of the companion representation is the first step of the “algorithm” for the computation of $\frac{p}{q}$ -expansions of the real numbers.

Let x be in $[0, \omega_{\frac{p}{q}}]$, and let \mathbf{c} be its companion representation. Let n be a fixed (large) positive integer and w be the prefix of length n of \mathbf{c} . When w is read *from right to left* by the converter \mathcal{C}_C — which is the transposed of $\mathcal{A}_{\frac{p}{q}}$ — and taking a state s as initial state, the output is a word $f^{(s)}$ of length n on the alphabet A and which depends upon s . The maximal common prefix of all these words $f^{(s)}$ 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 n' larger than n , but it is not possible to know in advance how large has to be this n' in order to get a better approximation.

5 On the fractional part of the powers of rational numbers

We are now in a position to give at least a sketch of the proof of the results we have announced in the introduction. In what follows, we suppose, once again, that $p \geq 2q - 1$.

For a fixed rational $\frac{p}{q}$ we define the subset $Y_{\frac{p}{q}}$ of $[0, 1[$ to be the union of q intervals of length $\frac{1}{p}$ by

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

where the k_c are such that $k_c \in \{0, \dots, p-1\}$ and $qk_c \equiv c \pmod{p}$. For instance:

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

Theorem 3 is then a direct consequence of the following:

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.*

As one can consider arbitrarily large rationals $\frac{p}{q}$, it then comes:

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 smaller than ε such that $\mathbf{Z}_{\frac{p}{q}}(Y_{\frac{p}{q}})$ is infinite countable. ■

The proof of Theorem 16 first relies on the characterization of reals with double $\frac{p}{q}$ -expansions (Lemma 18).

LEMMA 18 Let x be in $[0, \omega_{\frac{p}{q}}]$. 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 $\{0, \dots, q-1\}$;
- (iv) x has an expansion which is an eventually maximal word;
- (v) x has an expansion which is eventually written on the alphabet $\{p-q, \dots, p-1\}$.

The next step in the proof of Theorem 16 is a characterization of the companion representation of the reals that have multiple $\frac{p}{q}$ -expansions (and thus two $\frac{p}{q}$ -expansions because of the assumption on p and q).

Let us write the digit alphabet $C = \{-(q-1), \dots, 0, 1, \dots, (p-1)\}$, the image of the function h , as the union $C = C_1 \cup C_2 \cup C_3$ with $C_1 = \{-(q-1), \dots, -1\}$, $C_2 = \{0, \dots, q-1\}$ and $C_3 = \{q, \dots, p-1\}$.

PROPOSITION 19 A real x has two $\frac{p}{q}$ -expansions if and only if its companion representation is eventually in $C_2^{\mathbb{N}}$.

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. Indeed, if \mathcal{A} is an automaton over an alphabet C , two distinct paths in \mathcal{A} with the same label give a path in the square of \mathcal{A} that goes outside of the “diagonal”. If \mathcal{T} is a transducer, the square \mathcal{T}^2 is obtained by constructing the square of the underlying automaton of \mathcal{T} and by giving as output label of each transition of \mathcal{T}^2 the pairs of corresponding output labels in \mathcal{T} , see [4] or [20].

This implies, as $\mathcal{A}_{\frac{p}{q}}$ has only two states under the current hypothesis, that \mathbf{c} is eventually in C_2 — this can be easily seen on Figure 4 for the case $\frac{p}{q} = \frac{3}{2}$.

Let \mathbf{c} and \mathbf{a} be the companion representation and a $\frac{p}{q}$ -representation respectively of a real x . By Proposition 15, (\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 $0 \leq i < n$, $a'_n = a_n - q$, 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 . ■

The final step consists in the description of the inverse of the function h . For every c in $C_2 = \{0, \dots, q-1\}$ let us define the integer k_c in A , i.e. $0 \leq k_c \leq p-1$, by $qk_c = c \pmod{p}$.

LEMMA 20 For every c in C_2 , $h(x) = c$ if and only if $\left\{\frac{x}{q}\right\} \in \left[\frac{1}{p}k_c, \frac{1}{p}(k_c+1)\right[$.

From Proposition 19 follows that a real x has two $\frac{p}{q}$ -expansions if and only if there exists $M > 0$ such that for any $n > M$,

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

and this concludes the proof of Theorem 16.

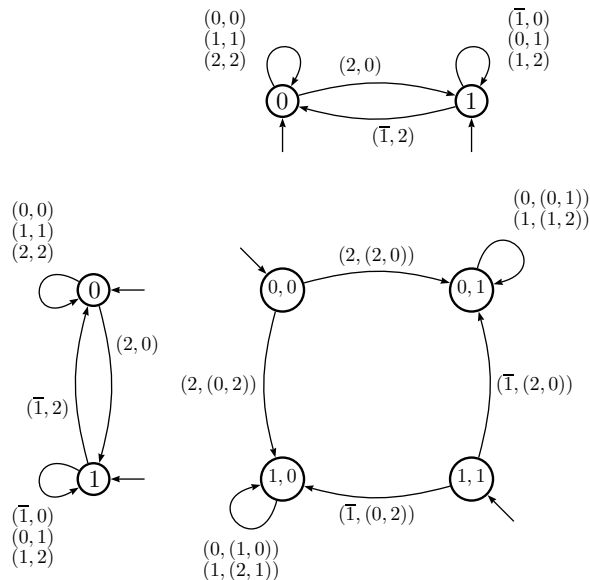


Figure 4: The square of $\mathcal{A}_{\frac{3}{2}}$ (outside of the diagonal)

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] S. Akiyama, Ch. Frougny and J. Sakarovitch, Powers of rationals modulo 1 and representations in a rational base, *to appear*.
- [3] A. Avizienis, Signed-digit number representations for fast parallel arithmetic, *IRE Transactions on electronic computers* **10** (1961) 389–400.
- [4] M.-P. Béal, O. Carton, C. Prieur, and J. Sakarovitch, Squaring transducers: An efficient procedure for deciding functionality and sequentiality, *Theoret. Comput. Sci.* **292** (2003) 45–63.
- [5] Y. Bugeaud, Linear mod one transformations and the distribution of fractional parts $\{\xi(\frac{p}{q})^n\}$, *Acta Arith.* **114** (2004) 301–311.
- [6] A. Cauchy, Sur les moyens d’éviter les erreurs dans les calculs numériques, *C.R. Acad. Sc. Paris série I* **11** (1840) 789–798.
- [7] S. Eilenberg, *Automata, Languages and Machines*, Vol. A, Academic Press (1974).
- [8] 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.

- [9] A.S. Fraenkel, Systems of numeration, *Amer. Math. Monthly* **92** (1985) 105–114.
- [10] Ch. Frougny, Representation of numbers and finite automata, *Math. Sys. Th.* **25** (1992) 37–60.
- [11] L. Halbeisen and N. Hungerbühler, The Josephus problem,
<http://citeseer.nj.nec.com/23586.html>.
- [12] J.E. Hopcroft and J.D. Ullman, *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesley (1979).
- [13] D. Knuth, *The Art of Computer Programming*, Addison Wesley (1969).
- [14] M. Lothaire, *Algebraic Combinatorics on Words*, Cambridge University Press (2002).
- [15] K. Mahler, An unsolved problem on the powers of $3/2$, *J. Austral. Math. Soc.* **8** (1968) 313–321.
- [16] A. Odlyzko and H. Wilf, Functional iteration and the Josephus problem, *Glasgow Math. J.* **33** (1991) 235–240.
- [17] W. Parry, On the β -expansions of real numbers, *Acta Math. Acad. Sci. Hung.*, **11** 401–416.
- [18] A.D. Pollington, Progressions arithmétiques généralisées et le problème des $(3/2)^n$. *C.R. Acad. Sc. Paris série I* **292** (1981) 383–384.
- [19] A. Rényi, Representations for real numbers and their ergodic properties, *Acta Math. Acad. Sci. Hung.* **8** (1957) 477–493.
- [20] J. Sakarovitch, *Eléments de théorie des automates*, Vuibert (2003). English translation: *Elements of Automata Theory*, Cambridge University Press, to appear.
- [21] N.J.A. Sloane, *The On-Line Encyclopedia of Integer Sequences*,
<http://www.research.att.com/~njas/sequences/>.
- [22] R. Stephan, On a sequence related to the Josephus problem,
<http://arxiv.org/abs/math.CO/0305348> (2003).
- [23] T. Vijayaraghavan, On the fractional parts of the powers of a number, I, *J. London Math. Soc.* **15** (1940), 159–160.
- [24] E. Weisstein, Power fractional parts, from MathWorld,
<http://mathworld.wolfram.com/PowerFractionalParts.html>

APPENDIX

A The first words in $L_{\frac{3}{2}}$

	0
2	1
21	2
210	3
212	4
2101	5
2120	6
2122	7
21011	8
21200	9
21202	10
21221	11
210110	12
210112	13
212001	14
212020	15
212022	16
212211	17
2101100	18
2101102	19
2101121	20
2120010	21
2120012	22
2120201	23
2120220	24
2120222	25
2122111	26
21011000	27
21011002	28
21011021	29
21011210	30
21011212	31
21200101	32
21200120	33
21200122	34
21202011	35
21202200	36
21202202	37
21202221	38
21221110	39
21221112	40

3/2-expansions of the 41 first integers

B Another view on $T_{\frac{3}{2}}$

