C I R M

Numération : mathématiques et informatique

Rencontre organisée par :
Boris Adamczewski, Anne Siegel and Wolfgang Steiner

23-27 mars 2009

Alina Firicel

Subword complexity and finite characteristic numbers

Vol. 1, n° 1 (2009), p. 29-34.

<http://acirm.cedram.org/item?id=ACIRM_2009__1_1_29_0>

Centre international de rencontres mathématiques
U.M.S. 822 C.N.R.S./S.M.F.
Luminy (Marseille) France

Actes des rencontres du C.I.R.M.
Vol. **1** n^o 1 (2009) 29-34

Subword complexity and finite characteristic numbers

Alina Firicel

Abstract

Decimal expansions of classical constants such as $\sqrt{2}$, π and $\zeta(3)$ have long been a source of difficult questions. In the case of finite characteristic numbers (Laurent series with coefficients in a finite field), where no carry-over difficulties appear, the situation seems to be simplified and drastically different. On the other hand, the theory of Drinfeld modules provides analogs of real numbers such as π, e or ζ values. Hence, it became reasonable to enquire how "complex" the Laurent representation of these "numbers" is.

1. Introduction

Let $\mathbf{a} = (a_n)_{n\geq 0}$ be an infinite sequence over a finite alphabet $\mathcal{A}$. The *subword complexity* of $\mathbf{a}$ is the function that associates with each $m \in \mathbb{N}$ the value $p(\mathbf{a}, m)$, equal to the number of distinct factors of length m occurring in $\mathbf{a}$.

For example, let us consider the infinite word $\mathbf{a} = aaa\cdots$, the concatenation of a letter a infinitely many times. It is obvious that $p(\mathbf{a}, m) = 1$ for any $m \in \mathbb{N}$. On the other side, let us consider the infinite word introduced by Champernowne in 1933 ([13]) $\mathbf{a} := 01234567891011\cdots$, that is the concatenation of the sequence of all nonnegative integers ranged in increasing order. Notice that $p(\mathbf{a}, m) = 10^m$ for any $m \in \mathbb{N}$.

However, one can easily prove that for every $m \in \mathbb{N}$ and for every word $\mathbf{a}$ over the alphabet $\mathcal{A}$, we have the following:

$$1 \leq p(\mathbf{a}, m) \leq (\operatorname{card} \mathcal{A})^m,$$

and both inequalities are sharp.

A long standing open question concerns the digits of the real number $\pi = 3.14159\cdots$. The decimal expansion of π has been calculated to billions of digits and unfortunately, there are no evident patterns occurring. Actually, for any $b \geq 2$, the b-ary expansion of π looks like a random sequence (see for example [5]). More concretely, it is widely believed that π is normal, meaning that all blocks of digits of equal length occur in the b-ary representation of π with the same frequency, but current knowledge on this point is scarce.

Let α be a real number and let $\mathbf{a}$ be the representation of α in an integral base $b \geq 2$. The complexity function of α is defined as follows:

$$p(\alpha, b, m) = p(\mathbf{a}, m),$$

for any positive integer m.

Notice that π being normal would imply that its complexity must be maximal, that is $p(\pi, b, m) = b^m$. In this direction, similar questions have been asked about other well-known constants like e, $\log 2$, ζ values or $\sqrt{2}$ and it is widely believed that the following conjecture is true.

Conjecture 1. *Let α be one of the classical constants: π, e, $\log 2$, ζ values and $\sqrt{2}$. The complexity of the real number α satisfies:*

$$p(\alpha, b, m) = b^m,$$

for every positive integer m and every $b \geq 2$.

Text presented during the meeting "Numeration: mathematics and computer science" organized by Boris Adamczewski, Anne Siegel and Wolfgang Steiner. 23-27 mars 2009, C.I.R.M. (Luminy).

If $\alpha \in \mathbb{Q}$, then $p(\alpha, b, m) = O(1)$, for every integer $b \geq 2$. Moreover, there is a famous theorem of Morse and Hedlund ([19]) which states that, if α is an irrational number then $p(\alpha, b, m) \geq m+1$. Concerning irrational algebraic numbers, the main result known to date in this direction is the following.

Theorem 1.1. *[Adamczewski–Bugeaud, 2004] Let α be an irrational algebraic number and $b \geq 2$ be an integer. Then:*

$$\lim_{m\to\infty} \frac{p(\alpha, b, m)}{m} = +\infty.$$

For more details concerning the complexity of real numbers, see for instance [1, 2].

The starting point of this paper is this type of questions, but in the world of Laurent series with coefficients in a finite field. Let us also recall the well-known analogy between integers, rationals and real numbers on one side, and polynomials, rationals functions, and Laurent series with coefficients in a finite field, on the other side.

In this paper, we will focus on Laurent series with coefficients in $\mathbb{F}_q$, where q is a power of a prime number p.

Let $n_0 \in \mathbb{N}$ and consider the Laurent series:

$$f(T) = \sum_{n=-n_0}^{+\infty} a_n T^{-n} \in \mathbb{F}_q((T^{-1})).$$

Let m be a nonnegative integer. We define *the complexity of f*, denoted by $p(f, m)$, as being equal to the complexity of the infinite word $\mathbf{a} = (a_n)_{n\geq 0}$.

2. Complexity of Laurent series

In this section we gather results concerning the complexity of Laurent series over $\mathbb{F}_q$, in function of their arithmetic properties. Notice that, if $f(T) \in \mathbb{F}_q(T)$, then $p(f, m) = O(1)$. First of all, we describe the situation of algebraic power series over $\mathbb{F}_q(T)$.

Secondly, we present the Carlitz' analogs of π, e or ζ values. Many of these finite characteristic numbers were shown to be transcendental over $\mathbb{F}_q(T)$ (see [18, 14, 20, 21, 22]).

Finally, we study the effect of usual operations over Laurent series. More precisely, we give some closure properties of Laurent series of "low" complexity (addition, multiplication, Hadamard product, derivative, Cartier operator).

For the complete proofs of these results, we refer the reader to [17].

2.1. Algebraic Laurent series.

A Laurent series $f(T) = \sum_{n\geq -n_0} a_n T^{-n} \in \mathbb{F}_q((T^{-1}))$ is said to be *algebraic* over the field $\mathbb{F}_q(T)$ if there exist an integer $d \geq 1$ and polynomials $A_0(T), A_1(T), \ldots, A_d(T)$, with coefficients in $\mathbb{F}_q$ and not all zero, such that:

$$A_0 + A_1 f + \cdots + A_d f^d = 0.$$

An important theorem of Christol ([15]) describes precisely the algebraic Laurent series over $\mathbb{F}_q(T)$ as follows.

Theorem 2.1. *[Christol, 1979] Let $f(T) = \sum_{n\geq -n_0} a_n T^{-n}$ be a Laurent series with coefficients over $\mathbb{F}_q$. Then f is algebraic over $\mathbb{F}_q(T)$ if, and only if, the sequence $(a_n)_{n\geq 0}$ is p-automatic.*

For more references on automatic sequences, see for example [4]. Furthermore, Cobham proved that the subword complexity of an automatic sequence is at most linear (see [16]). Hence, an easy consequence of those two results is the following.

Theorem 2.2. *Let $f \in \mathbb{F}_q((T^{-1}))$ algebraic over $\mathbb{F}_q(T)$. Then we have:*

$$p(f, m) = O(m).$$

The reciprocal is obviously not true, since there clearly are uncountable many Laurent series with linear complexity.

An explicit example can be the following transcendental Laurent series:

$$f(T) = \sum_{n\geq 0} f_n T^{-n} \in \mathbb{F}_q((T^{-1})),$$

where $\mathbf{f} := (f_n)_{n\geq 0} = 01001010010010100101\cdots$ is the fixed point of the morphism: $\sigma(0) = 01$ and $\sigma(1) = 0$. The infinite word $\mathbf{f}$ is called the Fibonacci word. It is not difficult to prove that this sequence is not k-automatic, for any positive integer $k \geq 2$. Consequently, according to Christol's theorem, f is transcendental over $\mathbb{F}_q(T)$, for every q. On the other side, it is not difficult to prove that $p(f, m) = O(m)$ (see for instance the Chapter "Sturmian words" from [7]).

In contrast with the real numbers world, the situation is clarified in the case of algebraic Laurent series. Also, notice that Theorems 1.1 and 2.2 give rise to the following corollary, which points out the fact that the situations in $\mathbb{F}_q((T^{-1}))$ and in $\mathbb{R}$ appear to be opposite.

Corollary 2.1. *Let* $f(T) := \sum_{n\geq -n_0} a_n T^{-n} \in \mathbb{F}_p((T^{-1})) \setminus \mathbb{F}_p(T)$ *and* $\alpha := \sum_{n\geq -n_0} a_n p^{-n} \in \mathbb{R} \setminus \mathbb{Q}$. *If the Laurent series f is algebraic over $\mathbb{F}_p(T)$, then α is transcendental. Conversely, if the real number α is algebraic, then the Laurent series f is transcendental over $\mathbb{F}_p(T)$.*

2.2. **Carlitz' analogs of some transcendental constants.** As we have already mentioned in the introduction, we have the following analogies:

$$\begin{aligned} \mathbb{Z} &\approx \mathbb{F}_q[T] \\ \mathbb{Q} &\approx \mathbb{F}_q(T) \\ \mathbb{R} &\approx \mathbb{F}_q((\frac{1}{T})) \\ \mathbb{C} &\approx \mathcal{C} \end{aligned}$$

where $\mathcal{C}$ is the completion of an algebraic closure of $\mathbb{F}_q((\frac{1}{T}))$.

In 1935, Carlitz introduced ([12]) functions in positive characteristic by analogy with the Riemann ζ function, the usual exponential and the logarithm function.

2.2.1. *Definitions.* The Carlitz ζ_q function is analogous of ζ Riemann function and it is defined as follows:

(1) $$\zeta_q : \mathbb{N}^* \to \mathbb{F}_q\left[\left[\frac{1}{T}\right]\right]; \qquad \zeta_q(n) = \sum_{\substack{\mathrm{P} \in \mathbb{F}_q[T] \\ \mathrm{P\ monic}}} \frac{1}{P^n}.$$

Moreover there exists a Laurent series Π_q such that:

(2) $$\forall\, n \equiv 0 \bmod (q-1),\, n \neq 0,\, \exists r_n \in \mathbb{F}_q(X)\ ,\ \zeta_q(n) = \Pi_q^n r_n.$$

Π_q may be defined also by the following infinite product:

(3) $$\Pi_q = \prod_{j=1}^{\infty} \left(1 - \frac{1}{T^{q^j - 1}}\right)^{-1}.$$

Remark 2.1. The property (2) is analogous to the classical results of Euler concerning ζ values for even integers. Indeed, it is well-known that, if s is even, then $\zeta(s) = \pi^s r$, where r is a rational number. This shows a good analogy between the real number π and the finite characteristic number Π_q.

The Carlitz *exponential*, denoted by $e_C(z)$, is defined over $\mathcal{C}$, by the following infinite product:

$$e_C(z) = z \prod_{a\in\mathbb{F}_q[T],\, a\neq 0} (1 - \frac{z}{a\widetilde{\Pi}_q})$$

where

$$\widetilde{\Pi}_q = (-T)^{\frac{q}{q-1}} \prod_{j=1}^{\infty} \left(1 - \frac{1}{T^{q^j - 1}}\right)^{-1}.$$

Remark 2.2. Since $e^z = 1$ if and only if $z \in 2\pi i\mathbb{Z}$ and since $e_C(z)$ was constructed by analogy such that $e_C(z) = 0$ if and only if $z \in \widetilde{\Pi}_q \mathbb{F}_q[T]$ (in other words the kernel of $e_C(z)$ is $\widetilde{\Pi}_q \mathbb{F}_q[T]$), we get a good analog $\widetilde{\Pi}_q$ of $2\pi i$.

Carlitz also defined the inverse of this exponential, the Carlitz *logarithm* by the following formula:

$$\log_C(z) = \sum_{k=0}^{+\infty} z^{q^k} L_k$$

where $L_k = \prod_{i=1}^{k} (T^{q^k} - T)$.

2.2.2. *Complexity of Carlitz' analogs.* In the previous section, we have recalled the definitions of the analogs of real numbers like π, ζ values, e or logarithm values. Now we are interested in the study of the complexity of these analogs in the case of Laurent series.

We begin with the case of the Carlitz's analog of π. If we look for the power series expansion of Π_q, then, using formula (3), one can easily get:

$$\Pi_q = \sum_{n \geq 0} a_n T^{-n},$$

where the sequence $(a_n)_{n\geq 0}$ is defined as following:

(4) $a_n =$ the number of partitions of n whose parts take values in $I = \{q^j - 1, j \geq 1\} \bmod p$.

Computing the complexity of Π_q asks for a closed formula or some recurrence relations for the sequence of partitions $(a_n)_{n\geq 0}$. This seems quite difficult and we are not able to solve at this moment.

However, it was shown in [3] that the inverse of Π_q has the following simple power series expansion:

$$\frac{1}{\Pi_q} = \prod_{j=1}^{\infty} \left(1 - \frac{1}{T^{q^j - 1}}\right) = \sum_{n=0}^{\infty} p_n T^{-n}$$

where the sequence $\mathfrak{p}_q = (p(n))_{n\geq 0}$ is defined as follows:

$$p_n = \begin{cases} 1 \text{ if } n = 0; \\ (-1)^{\operatorname{card} J} \text{ if there exist a set } J \subset \mathbb{N}^* \text{ such that } n = \sum_{j\in J} (q^j - 1); \\ 0 \text{ if there is no set } J \subset \mathbb{N}^* \text{ such that } n = \sum_{j\in J} (q^j - 1). \end{cases}$$

We mention that if such a decomposition exists, it is unique. Besides the fact that this sequence takes only 3 values (for any q), it is possible to find some nice patterns in $\mathfrak{p}_q$. Hence, $\mathfrak{p}_q$ turn out to be easier to study that the sequence defined in (4).

Using the above expression, we obtain the following result ([17]).

We mention that here we use Landau's notations. We write $f(m) = \Theta(g(m))$ if f is bounded both above and below by g asymptotically. In other words, there exist positive integers k_1, k_2, n_0 such that, for every $n > n_0$ we have

$$k_1 |g(n)| < f(n) < k_2 |g(n)|.$$

Theorem 2.3. *Let $q = 2$. The complexity of the inverse of Π_q satisfies:*

$$p(\frac{1}{\Pi_2}, m) = \Theta(m^2).$$

Theorem 2.4. *Let $q \geq 3$. The complexity of the inverse of Π_q satisfies $p(\frac{1}{\Pi_q}, m) = \Theta(m)$.*

Since any algebraic Laurent series has a complexity at most linear (by Theorem 2.2), the following corollary yields.

Corollary 2.2. *Π_2 is transcendental over $\mathbb{F}_2(T)$.*

Thus, the subword complexity of Laurent series can be used to prove transcendence results.

We end this section with some questions related to other Carlitz' analogs.

Question 2.1. Using definition (1), we have for example:

$$\begin{aligned}\zeta_2(1) =& \frac{1}{1} + \frac{1}{T} + \frac{1}{T+1} + \frac{1}{T^2} + \frac{1}{T^2+1} + \frac{1}{T^2+T} + \frac{1}{T^2+T+1} + \cdots \\ =& 1 + \frac{1}{T^2} + \frac{1}{T^3} + \frac{1}{T^4} + \frac{1}{T^5} + \frac{1}{T^9} + \cdots \in \mathbb{F}_q((T^{-1}))\end{aligned}$$

What can we say about the complexity of $\zeta_2(1)$? Or more generally, what can we say about the complexity of $\zeta_q(n)$, for any positive integer n?

In the same spirit, what can we say about the complexity of the Laurent series expansion of $e_C(1)$, the analog of the real number e?

When investigating these problems, we need, in general, the Laurent series expansions of such functions. In this context, one has to mention the work of Berthé [8, 9, 10, 11], where there are illustrated some Laurent series expansions of Carlitz' functions.

Problem 2.1. It would be interesting to investigate the following general question.

Is it true that finite charcteristic analogs of classical constants (see Section 2.2) all have a "low" complexity (i.e. polynomial or subexponential)?

The first clue in this direction are the examples provided by Theorems 2.2, 2.3 and 2.4.

Notice also that a positive answer would reinforce the differences between $\mathbb{R}$ and $\mathbb{F}_q((T^{-1}))$ as hinted in Corollary 2.1. Indeed, recall that classical constants like $\sqrt{2}, e, \pi, \log 2 \in \mathbb{R}$ are expected to be normal, which implies that they have an exponential complexity.

2.3. **Closure properties of a special class of Laurent series.** It is natural to classify Laurent series in function of their complexity. Let us introduce the following set:

$$\mathcal{P} = \{f \in \mathbb{F}_q((T^{-1})), \text{there exists } K \text{ such that } p(f, m) = O(m^K)\}.$$

We have already seen, in Theorem 2.2, that the algebraic Laurent series belong to $\mathcal{P}$. Also, by Theorem 2.3 and 2.4, $\frac{1}{\Pi_q}$ belongs to $\mathcal{P}$. Hence, $\mathcal{P}$ seems to be an important object of interest for this classification. Furthermore, $\mathcal{P}$ is a good candidate to enjoy some nice closure properties.

In this direction, we obtain in [17] the following result.

Theorem 2.5. *$\mathcal{P}$ is a vector space over $\mathbb{F}_q(T)$.*

The proof of this theorem is a straightforward consequence of Propositions 1 and 2.

Proposition 1. *Let f and g be two Laurent series belonging to $\mathbb{F}_q((T^{-1}))$. Then, for every $m \in \mathbb{N}$ we have:*

$$p(f+g, m) \leq p(f, m)p(g, m). \tag{5}$$

Proposition 2. *Let $r(T) \in \mathbb{F}_q(T)$ and $f(T) = \sum_{n \geq -n_0} a_n T^{-n} \in \mathbb{F}_q((T^{-1}))$. Then for every $m \in \mathbb{N}$, there is a positive constant M, depending only on $r(T)$ and n_0, such that:*

$$p(rf, m) \leq M p(f, m).$$

Related to Proposition 1, one can naturally ask the following question.

Question 2.2. Is it possible to saturate the inequality (5) in Proposition 1? In particular, it would be interesting to construct two explicit examples of Laurent series of linear complexity such that their sum has quadratic complexity.

Further, we mention other closure properties of $\mathcal{P}$.

Theorem 2.6. *$\mathcal{P}$ is stable by Hadamard product, formal derivative, Cartier operator.*

The reader may refer to Chapter 12 of the monograph [4] for more details about Hadamard product or Cartier operator.

Problem 2.2. It would be also interesting to study other closure properties of the class $\mathcal{P}$. In particular, is it true that $\mathcal{P}$ is closed under the usual Cauchy product?

Notice that if the answer is positive, this could lead to results of algebraic independence. There are actually some particular cases of Laurent series with low complexity whose product still belongs to $\mathcal{P}$ (see [17]).

On the other hand, if the answer is negative, is it possible to give some explicit examples of Laurent series belonging to $\mathcal{P}$, whose product does not belong to $\mathcal{P}$?

3. Concluding remarks

We conclude this note with some remarks and comments.

- First of all, we precise that another motivation of this work has been also the article [6] of R. Beals and D. Thakur. They proposed a classification of finite characteristic numbers in function of their space or time complexity. This complexity is in fact a characteristic of the (Turing) machine that computes the coefficient a_i, if $f(T) := \sum_i a_i T^{-i}$. They showed that some classes of Laurent series have good algebraic properties (for example the class of Laurent series corresponding to deterministic space class $S(n)$, $S(n) \geq n$, form a field). They also place some finite characteristic numbers like Π_q, $1/\Pi_q$ or e in the computational hierarchy.
- Other applications of complexity results for Laurent series include diophantine properties. In particular, bounds for irrationality measures can be obtained for elements of the class of Laurent series of linear complexity.

References

[1] B. Adamczewski & Y. Bugeaud, On the complexity of algebraic numbers I. Expansions in integer bases, Annals of Mathematics 165 (2007), 547–565.
[2] B. Adamczewski & Y. Bugeaud & F. Luca, Sur la complexité des nombres algébriques. C. R. Math. Acad. Sci. Paris, 339(2004), 11-14.
[3] J.-P. Allouche, Sur la transcendance de la série formelle Π, Journal de Théorie des Nombres de Bordeaux 2 (1990), 103–117.
[4] J.-P. Allouche & J. Shallit, *Automatic Sequences: Theory, Applications, Generalizations*, Cambridge University Press, Cambridge, 2003.
[5] David H. Bailey, Peter B. Borwein and Simon Plouffe, On the Rapid Computation of Various Polylogarithmic Constants, Mathematics of Computation, vol. 66, no. 218 (1997), pg. 903?-913.
[6] R.M. Beals & D.S. Thakur, *Computational classification of numbers and algebraic properties*, Internat. Math. Res. Notices, 15 (1998), 799–818.
[7] J. Berstel and P. Séébold , *Algebraic combinatorics on Words*, chapter Sturmian words, Cambridge University Press (2002).
[8] V. Berthé, Automates et valeurs de transcendance du logarithme de Carlitz Acta Arithmetica LXVI.4 (1994), 369–390.
[9] V. Berthé, Combinaisons linéaires de $\zeta(s)/\Pi^s$ sur $F_q(x)$, pour $1 \leq s \leq q-2$ J. Number Theory 53 (1995), 272–299.
[10] V. Berthé, De nouvelles preuves preuves "automatiques" de transcendance pour la fonction zeta de Carlitz Journées Arithmétiques de Genève, Astérisque 209 (1992), 159–168 .
[11] V. Berthé, Fonction zeta de Carlitz et automates, Journal de Théorie des Nombres de Bordeaux 5 (1993), 53–77.
[12] L. Carlitz, On certain functions connected with polynomials in a Galois field, Duke Math. J., 1 (1935), 137–168.
[13] D. G. Champernowne, The construction of decimals normal in the scale of ten, J. London Math. Soc. (1933), 8, 254?-260.
[14] H. Cherif, B. de Mathan, Irrationality measures of Carlitz zeta values in characteristic p, J. Number Theory 44 (1993), 260–272.
[15] Christol, G.; Kamae, T.; Mendès France, Michel; Rauzy, Gérard, Suites algébriques, automates et substitutions, Bulletin de la Société Mathématique de France, 108 (1980), 401–419.
[16] A. Cobham, Uniform tag sequences, Math. Systems Theory 6 (1972), 164–192.
[17] A. Firicel, Subword Complexity and Laurent series over finite fields, in progress.
[18] D. Goss, *Basic Structures of Function Field Arithmetic*, Springer-Verlag, Berlin, 1996.
[19] M. Morse & G. A. Hedlund, Symbolic dynamics, Amer. J. Math. 60 (1938), 815–866.
[20] D.S. Thakur, *Function Field Arithmetic*, World Scientfic, Singapore, 2004.
[21] L.J. Wade, Certain quantities transcendental over $GF(p_n)(x)$, Duke Math. J. 8 (1941), 701–720.
[22] J. Yu, Transcendence and Special Zeta Values in Characteristic p, Annals of Math. 134 (1991), 1–23.

Université de Lyon, Université Lyon 1, Institut Camille Jordan, UMR 5208 du CNRS, 43, boulevard du 11 novembre 1918, F-69622 Villeurbanne Cedex, France • firicel@math.univ-lyon1.fr • `http://math.univ-lyon1.fr/~firicel/`