

# FUNCTION FIELDS IN POSITIVE CHARACTERISTIC: EXPANSIONS AND COBHAM'S THEOREM

BORIS ADAMCZEWSKI AND JASON BELL

ABSTRACT. In the vein of Christol, Kamae, Mendès France and Rauzy, we consider the analogue of a problem of Mahler for rational functions in positive characteristic. To solve this question, we prove an extension of Cobham's theorem for quasi-automatic functions and use the recent generalization of Christol's theorem obtained by Kedlaya.

## 1. INTRODUCTION

One motivation for the present work comes from a number theoretical problem concerning the expansion of algebraic numbers in integer bases. It appears at the end of a paper of Mendès France [11], but in conversation he attributes the paternity of this problem to Mahler (see for instance the discussion in Allouche and Shallit [1]). It can be stated as follows. Let  $\mathbf{a} = (a_n)_{n \geq 0}$  be a binary sequence and consider the two real numbers

$$\alpha = \sum_{n \geq 0} \frac{a_n}{2^n} \quad \text{and} \quad \beta = \sum_{n \geq 0} \frac{a_n}{3^n}.$$

Then, the problem is to show that these numbers are both algebraic if and only if they are both rational. At first glance, this problem seems contrived, but behind it hides the more fundamental question of the structure of representations of real numbers in two multiplicatively independent integer bases. Unfortunately, problems of this type are difficult and we seem to be far from answering Mahler's question.

However, when considering addition and multiplication without carry things become easier. In particular, we have a nice result of Christol, Kamae, Mendès France and Rauzy [4]: a sequence of coefficients represents two algebraic power series in distinct characteristics if and only if these power series are rational functions. In more concrete terms, Christol *et al.* give the following result.

**Theorem 1.1** (Christol *et al.*). *Let  $p_1$  and  $p_2$  be distinct prime numbers and let  $q_1$  and  $q_2$  be powers of prime  $p_1$  and  $p_2$  respectively. Let  $(a_n)_{n \geq 0}$  be a sequence with values in a finite set  $\mathcal{A}$  with cardinality at most  $\min\{q_1, q_2\}$ . Let  $i_1$  and  $i_2$  be two injections from  $\mathcal{A}$  into  $\mathbb{F}_{q_1}$  and  $\mathbb{F}_{q_2}$  respectively. Then, the formal power series*

$$f(t) = \sum_{n \geq 0} i_1(a_n)t^n \in \mathbb{F}_{q_1}((t)) \quad \text{and} \quad g(t) = \sum_{n \geq 0} i_2(a_n)t^n \in \mathbb{F}_{q_2}((t))$$

*are both algebraic (respectively over  $\mathbb{F}_{q_1}(t)$  and  $\mathbb{F}_{q_2}(t)$ ) if and only if they are rational functions.*

As was remarked by Christol *et al.* [4], Theorem 1.1 is a straightforward consequence of two important results. On one side, Christol's theorem [3] describes precisely in terms of automata the algebraic closure of  $\mathbb{F}_q(t)$  in  $\mathbb{F}_q((t))$  ( $q$  being a power of a prime  $p$ ). On the other side, one finds Cobham's theorem [5] proving that for multiplicatively independent positive integers  $k$  and  $l$ , a function  $h : \mathbb{N} \rightarrow \mathbb{F}_q$  that is both  $k$ - and  $l$ -automatic is eventually periodic (see Section 3 for a definition of an automatic function and a more precise statement of this result).

Christol's theorem gives a very concrete description of the elements of  $\mathbb{F}_q((t))$  that are algebraic over  $\mathbb{F}_q(t)$ ; it shows in fact that being an algebraic power series is equivalent to the sequence of coefficients being  $p$ -automatic. As Kedlaya [9] points out, this result does not give the complete picture, as the field  $\mathbb{F}_q((t))$  is far from being algebraically closed. Indeed, for an algebraically closed field  $\mathbb{K}$  of characteristic 0, the field

$$\bigcup_{i=1}^{\infty} \mathbb{K}((t^{1/i}))$$

is itself algebraically closed and contains in particular the algebraic closure of  $\mathbb{K}(t)$ ; but in positive characteristic, things are rather different. The algebraic closure of  $\mathbb{F}_q((t))$  is more complicated, due to the existence of wildly ramified field extensions. For instance, Chevalley remarked [2] that the Artin-Schreier

polynomial  $x^p - x - 1/t$  does not split in the field  $\bigcup_{n=1}^{+\infty} \mathbb{F}_q((t^{1/n}))$ .

It turns out that the appropriate framework to describe the algebraic closure of  $\mathbb{F}_p(t)$  is provided by the fields of generalized power series  $\mathbb{F}_q((t^{\mathbb{Q}}))$  introduced by Hahn [7]; the construction of which is the object of Section 4. The work of Kedlaya [10] is precisely devoted to a description of the algebraic closure of  $\mathbb{F}_p(t)$  in such fields of generalized power series. For this purpose, Kedlaya introduces the notion of a  $p$ -quasi-automatic function over the rationals (see Section 5 for a definition). His extension of Christol's theorem is that if  $q$  is a power of a prime  $p$ , then a generalized power series

$$\sum_{\alpha \in \mathbb{Q}} h(\alpha) t^\alpha \in \mathbb{F}_q((t^{\mathbb{Q}}))$$

is algebraic if and only if the function  $h : \mathbb{Q} \rightarrow \mathbb{F}_q$  is  $p$ -quasi-automatic.

Motivated by this recent work of Kedlaya, we propose to give an analogue of the result of Christol *et al.* [4] for generalized power series. That is,

- first prove an extension of Cobham's theorem to quasi-automatic functions (this is Theorem 2.1);
- then, put it together with Kedlaya's result to derive an analogue of Theorem 1.1 (this corresponds to Theorem 2.2).

Actually, the full-power of Theorem 2.1 is not needed for proving Theorem 2.2 and the necessary part of it relies more straightforwardly on the original version of Cobham's theorem.

The paper is organized as follows. Our main results are presented in Section 2. We recall the definition of an automatic sequence and the classical

version of Cobham's theorem in Section 3. We describe Hahn's construction of generalized power series rings in Section 4. The notion of quasi-automatic functions and Kedlaya's extension of Christol theorem are recalled in Section 5. We introduce the notion of Saguaro sets in Section 6 and prove a Cobham's theorem analogue for these sets in Section 7. The proofs of our main results are postponed to Section 8.

## 2. MAIN RESULTS

Our main result is the following extension of Cobham's theorem to quasi-automatic functions. We refer the reader to Section 5 for a definition of the notion of a quasi-automatic function.

**Theorem 2.1.** *Let  $k$  and  $\ell$  be multiplicatively independent positive integers. A function  $h : \mathbb{Q} \rightarrow \Delta$  is both  $k$ - and  $\ell$ -quasi-automatic if and only if there exist integers  $a$  and  $b$  with  $a > 0$  such that:*

- (1) *the sequence  $\{h((n - b)/a)\}_{n \in \mathbb{N}}$  is eventually periodic;*
- (2)  *$h((x - b)/a) = 0$  for  $x \in \mathbb{Q} \setminus \mathbb{N}$ .*

Note that our approach relies on the classical version of Cobham's theorem and we thus do not derive an independent proof of that result.

For a definition of the notions of generalized power series and of well-ordered sets, the reader is referred to Section 4. Thanks to Kedlaya's extension of Christol's theorem (see Section 5), we then prove that for a generalized power series, being algebraic in two distinct characteristics reduces to triviality (i.e., "almost rationality"); providing in this framework an analogue of Theorem 1.1. The main interest for such a generalization is to present a complete picture, in the sense that no algebraic function escapes this statement.

**Theorem 2.2.** *Let  $p_1$  and  $p_2$  be distinct primes and let  $q_1$  and  $q_2$  be powers of  $p_1$  and  $p_2$  respectively. Let  $(r_\alpha)_{\alpha \in \mathbb{Q}}$  be sequence with well-ordered support and with values lying in a finite set  $\mathcal{A}$  with cardinality at most  $\min\{q_1, q_2\}$ . Let  $i_1$  and  $i_2$  be injections from  $\mathcal{A}$  into  $\mathbb{F}_{q_1}$  and  $\mathbb{F}_{q_2}$  respectively. Then the generalized power series*

$$f(t) = \sum_{\alpha \in \mathbb{Q}} i_1(r_\alpha) t^\alpha \in \mathbb{F}_{q_1}((t^{\mathbb{Q}})) \quad \text{and} \quad g(t) = \sum_{\alpha \in \mathbb{Q}} i_2(r_\alpha) t^\alpha \in \mathbb{F}_{q_2}((t^{\mathbb{Q}}))$$

*are both algebraic (respectively over  $\mathbb{F}_{q_1}(t)$  and  $\mathbb{F}_{q_2}(t)$ ) if and only if there exists a positive integer  $n$  such that  $f(t^n)$  and  $g(t^n)$  are both rational functions.*

Note that it would be unreasonable to expect in this context that, as in the conclusion of Theorem 1.1,  $f$  and  $g$  are two rational functions. Indeed, the function  $a(t) = t^{1/2}$  is clearly algebraic over  $\mathbb{F}_p(t)$  for every prime  $p$ , without being itself rational. However, in this case we have that  $a(t^2)$  is a rational function.

### 3. FINITE AUTOMATA AND COBHAM'S THEOREM

In this Section, we recall the definition of an automatic sequence and the classical version of Cobham's theorem.

Let  $k \geq 2$  be an integer. An infinite sequence  $\mathbf{a} = (a_n)_{n \geq 0}$  is said to be  $k$ -automatic if  $a_n$  is a finite-state function of the base- $k$  representation of  $n$ . This means that there exists a finite automaton starting with the  $k$ -ary expansion of  $n$  as input and producing the term  $a_n$  as output. A nice reference on this topic is the book of Allouche and Shallit [1].

More precisely,  $k$ -automatic sequences can be given as follows. Denote by  $\Sigma_k$  the set  $\{0, 1, \dots, k-1\}$ . By definition, a  $k$ -automaton is a 6-tuple

$$A = (Q, \Sigma_k, \delta, q_0, \Delta, \tau),$$

where  $Q$  is a finite set of states,  $\delta : Q \times \Sigma_k \rightarrow Q$  is the transition function,  $q_0$  is the initial state,  $\Delta$  is the output alphabet and  $\tau : Q \rightarrow \Delta$  is the output function. For a state  $q$  in  $Q$  and for a finite word  $W = w_1 w_2 \dots w_n$  on the alphabet  $\Sigma_k$ , we define recursively  $\delta(q, W)$  by  $\delta(q, W) = \delta(\delta(q, w_1 w_2 \dots w_{n-1}), w_n)$ . Let  $n \geq 0$  be an integer and let  $w_r w_{r-1} \dots w_1 w_0$  in  $(\Sigma_k)^{r+1}$  be the  $k$ -ary expansion of  $n$ ; thus,  $n = \sum_{i=0}^r w_i k^i$ . We denote by  $W_n$  the word  $w_0 w_1 \dots w_r$ . Then, a sequence  $\mathbf{a} = (a_n)_{n \geq 0}$  is said to be  $k$ -automatic if there exists a  $k$ -automaton  $A$  such that  $a_n = \tau(\delta(q_0, W_n))$  for all  $n \geq 0$ .

A classical example of a 2-automatic sequence is given by the binary Thue-Morse sequence  $\mathbf{a} = (a_n)_{n \geq 0} = 0110100110010\dots$ . This sequence is defined as follows:  $a_n$  is equal to 0 (resp. to 1) if the sum of the digits in the binary expansion of  $n$  is even (resp. is odd). It is easy to check that it can be generated by the 2-automaton

$$A = (\{q_0, q_1\}, \{0, 1\}, \delta, q_0, \{0, 1\}, \tau),$$

where

$$\delta(q_0, 0) = \delta(q_1, 1) = q_0, \quad \delta(q_0, 1) = \delta(q_1, 0) = q_1,$$

and  $\tau(q_0) = 0$ ,  $\tau(q_1) = 1$ .

We recall the classical version of Cobham's theorem. Given a finite set  $\Delta$ , a function  $h : \mathbb{N} \rightarrow \Delta$  is said to be  $k$ -automatic if  $(h(n))_{n \geq 0}$  is a  $k$ -automatic sequence.

**Theorem 3.1** (Cobham). *Let  $k$  and  $\ell$  be multiplicatively independent positive integers and  $\Delta$  be a finite set. A function  $h : \mathbb{N} \rightarrow \Delta$  is both  $k$ - and  $\ell$ -automatic if and only if it is eventually periodic.*

### 4. GENERALIZED POWER SERIES

As mentioned above, Kedlaya gives a complete generalization of Christol's theorem by extending it to algebraic elements in an algebraically closed overring of  $\mathbb{F}_q(t)$ . To describe his result, we must first look at Hahn's work [7] on *generalized power series*.

We recall that a subset  $S$  of a totally ordered group is said to be well-ordered if every nonempty subset of  $S$  has a minimal element or, equivalently, if there is no infinite decreasing sequence within  $S$ . Given a commutative

ring  $R$  and a totally ordered abelian group  $G$  we construct a commutative ring, denoted by  $R((t^G))$ , which is defined to be the collection of all elements of the form

$$f(t) := \sum_{\alpha \in G} r_\alpha t^\alpha$$

which satisfy:

- $r_\alpha \in R$  for all  $\alpha \in G$ ;
- The support of  $f(t)$  is well ordered; that is, the subset  $\{\alpha \mid r_\alpha \neq 0_G\}$  is a well-ordered set.

Addition and multiplication are defined via the rules

$$\sum_{\alpha \in G} r_\alpha t^\alpha + \sum_{\alpha \in G} s_\alpha t^\alpha = \sum_{\alpha \in G} (r_\alpha + s_\alpha) t^\alpha$$

and

$$\left( \sum_{\alpha \in G} r_\alpha t^\alpha \right) \left( \sum_{\alpha \in G} s_\alpha t^\alpha \right) = \sum_{\alpha \in G} \sum_{\beta \in G} (r_\beta s_{\alpha-\beta}) t^\alpha.$$

We note that the fact that the support of valid series expansion is well-ordered means that no problems with possible infinite sums appearing in the expression for the coefficients in a product of two generalized power series will occur. We call the ring  $R((t^G))$  the *ring of generalized power series over  $R$  with exponent in  $G$* . We recall that a group is *divisible* if for every  $g \in G$  and  $n \geq 1$ , there exists some  $h \in G$  such that  $h^n = g$ .

For an algebraically closed field  $\mathbb{K}$  and a divisible group  $G$ , the field  $\mathbb{K}((t^G))$  is known to be algebraically closed [8] (see also [9, 12]). In what follows, we will only consider the particular case of the divisible group  $\mathbb{Q}$  and of a finite field  $\mathbb{F}_q$  ( $q$  being a power of a prime  $p$ ). We then have the series of containments

$$\mathbb{F}_q(t) \subset \mathbb{F}_q((t)) \subset \mathbb{F}_q((t^{\mathbb{Q}})).$$

Though  $\mathbb{F}_q((t^{\mathbb{Q}}))$  is not algebraically closed, it is sufficient for our purpose to consider such fields. Indeed, taking  $\bigcup_{n \geq 1} \mathbb{F}_{p^n}$  as an algebraic closure of  $\mathbb{F}_p$ , it follows from the remark above that the field  $(\bigcup_{n \geq 1} \mathbb{F}_{p^n})((t^{\mathbb{Q}}))$  is algebraically closed.

Kedlaya asks [9] whether one can, as in the classical Christol theorem, give an automata-theoretic characterization of the elements of  $\mathbb{F}_q((t^{\mathbb{Q}}))$  that are algebraic over  $\mathbb{F}_q(t)$ . In the next section we describe Kedlaya's later work in answering this question [10].

## 5. KEDLAYA'S THEOREM

In this section, we describe the work of Kedlaya and his generalization of Christol's theorem.

Let  $k > 1$  be a positive integer. We set

$$\Sigma'_k = \{0, 1, \dots, k-1, \cdot\}$$

and we denote by  $\mathcal{L}(k)$  the language on the alphabet  $\Sigma'_k$  consisting of all words on  $\Sigma'_k$  with exactly one occurrence of the letter ' $\cdot$ ' (the radix point) and whose first and last letters are not equal to 0. This is a regular language

[10, Lemma 2.3.3]. We let  $S_k$  denote the set of nonnegative  $k$ -adic rationals; i.e.,

$$S_k = \{a/k^b \mid a, b \in \mathbb{Z}, a \geq 0\}.$$

We note that there is a bijection  $[\cdot]_k : \mathcal{L}(k) \rightarrow S_k$  given by

$$s_1 \cdots s_{i-1} s_{i+1} \cdots s_n \in \mathcal{L}(k) \mapsto \sum_{j=1}^{i-1} s_j k^{i-1-j} + \sum_{j=i+1}^n s_j k^{i-j},$$

where  $s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n \in \{0, 1, \dots, k-1\}$ . So, for example, we have  $[110.32]_4 = [20.875]_{10} = 167/8$ . We also note that the fact that we exclude strings whose initial and terminal letters are 0 means that we have the awkward looking expression  $[\cdot]_k = 0$ .

Kedlaya works with a different definition of a  $k$ -automatic sequence than the usual one, focusing instead on maps from  $S_k$  into a finite set.

**Definition 5.1.** We say that a map  $h : S_k \rightarrow \Delta$  is  $K_k$ -automatic if the support of  $h$  is well-ordered and there is a finite state machine which takes words on  $\Sigma'_k$  as input such that for each  $W \in \mathcal{L}_k$ ,  $h([W]_k)$  is generated by the machine using the word  $W$  as input.

In the sequel, we will need a more general notion of automatic functions defined over the set of rationals. For this purpose, we always implicitly consider sets  $\Delta$  containing a special element called zero and denoted by 0 (of course, when  $\Delta$  is a subset of  $\mathbb{R}$  or  $\mathbb{N}$ , or if it denotes a finite field, zero will preserve its usual meaning). Then, we will talk about functions  $h : \mathbb{Q} \rightarrow \Delta$  as being  $K_k$ -automatic if their support is contained in  $S_k$  and the restriction of  $h$  to  $S_k$  is  $K_k$ -automatic (the support of such a function being defined as the set  $S = \{x \in \mathbb{Q} \mid h(x) \neq 0\}$ ).

**Example 5.2.** For  $W \in \mathcal{L}(2)$ , define

$$h([W]_2) = \begin{cases} 0 & \text{if there are an even number of 1's in } W \\ 1 & \text{otherwise.} \end{cases}$$

Then  $h : S_2 \rightarrow \{0, 1\}$  is  $K_2$ -automatic.

*Proof.* Note the finite state machine in Figure 1 has the property that the output of a word  $W$  is just the number, reduced modulo 2, of ones in the expansion of  $W$ .  $\square$

**Definition 5.3.** Let  $k$  be a positive integer and let  $h : \mathbb{Q} \rightarrow \Delta$  be a function with well-ordered support  $S$ . We say that  $h$  is  $k$ -quasi-automatic if there exist integers  $a$  and  $b$  with  $a > 0$  such that:

- the set  $aS + b$  consists of nonnegative  $k$ -adic rationals.
- the map  $h((x - b)/a)$  is  $K_k$ -automatic.

We will find it convenient to talk about  $K_k$ -automatic and  $k$ -quasi-automatic subsets of  $\mathbb{Q}$ . We say that a set is  $K_k$ -automatic (resp.  $k$ -quasi-automatic) if its characteristic function is  $K_k$ -automatic (resp.  $k$ -quasi-automatic).

We note that if  $f$  is  $k$ -quasi-automatic with support  $S$ , then the map  $f((x - b)/a)$  will be  $K_k$ -automatic whenever  $a > 0$  and  $b$  are integers which satisfy  $aS + b \subseteq S_k$ .

Kedlaya's main theorem is the following.

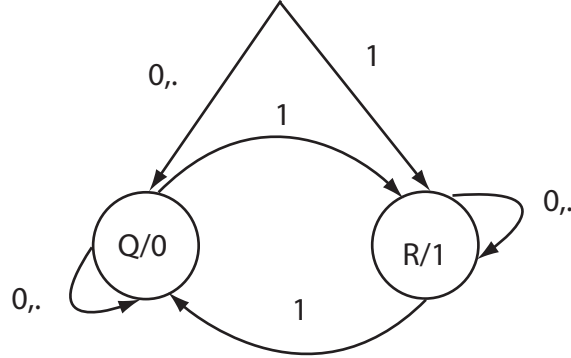


FIGURE 1. A finite state machine with input alphabet  $\{0, 1, .\}$  and output alphabet  $\{0, 1\}$ .

**Theorem 5.4** (Kedlaya). *Let  $p$  be a prime, let  $q$  be a power of  $p$ , and let  $f : \mathbb{Q} \rightarrow \mathbb{F}_q$ . Then  $\sum_{\alpha \in \mathbb{Q}} f(\alpha)t^\alpha$  is algebraic over  $\mathbb{F}_q(t)$  if and only if the function  $f : \mathbb{Q} \rightarrow \mathbb{F}_q$  is  $p$ -quasi-automatic.*

### 6. SAGUARO SETS

In this section, motivated by the work of Kedlaya, we introduce the notion of a *Saguaro set* and obtain a number-theoretic description of such sets. The behaviour of quasi-automatic functions has two main components: the pre-radix point behaviour, which is much like ordinary automatic sequences; and the post-radix point behaviour, which has restrictions. To describe the post-radix point behaviour of a quasi-automatic function, we define Saguaro sets.

**Definition 6.1.** We say that a set  $S$  is  $k$ -Saguaro if  $S$  is the support of a  $K_k$ -automatic function and  $S \subseteq [0, 1]$ .

**Example 6.2.** Let  $S$  be the set of all numbers of the form  $[\cdot 2^a 1^b]_3$ . Then  $S$  is 3-Saguaro.

*Proof.* Note that the automaton in Figure 2 shows that  $S$  is  $K_3$ -automatic. Observe that  $S$  is well-ordered since  $[\cdot 2^a 1^b]_3 < [\cdot 2^c 1^d]_3$  if and only if  $c > a$  or  $c = a$  and  $d > b$ . Thus given any subset of  $S$  we can choose the least element in the subset simply by picking the element of the form  $[\cdot 2^a 1^b]_3$  with  $a$  minimal, and among all such elements with this minimal value of  $a$ , pick the one with the minimum value of  $b$ . Hence  $S$  is well-ordered and so it is 3-quasi-automatic.  $\square$

The motivation for the name Saguaro comes from graph theory. A connected, undirected graph is called a *cactus* if each vertex in the graph lies

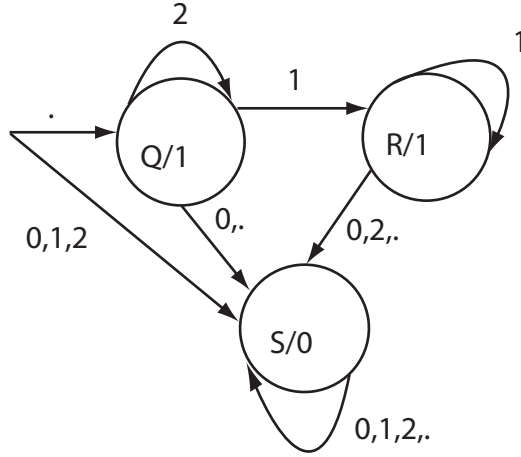


FIGURE 2. A finite state machine which produces a Saguaro set.

on at most one minimal cycle. Using this as his motivation, Kedlaya [10] defines a class of digraphs, which he calls *Saguaro digraphs*.

**Definition 6.3.** A rooted directed graph with vertex set  $V$ , edge set  $E$ , and distinguished vertex  $v_0 \in V$  is called a *rooted Saguaro* if each  $v \in V$  lies on at most one minimal cycle, and there is a directed path from  $v_0$  to each vertex  $v \in V$ .

**Example 6.4.** Let  $G$  be the transition graph of the automaton in Figure 2 and let  $v_0$  be the vertex corresponding to state  $Q$ . Then  $G$  is a rooted Saguaro.

In fact, Kedlaya shows there is a strong connection between Saguaro digraphs and well-ordered  $k$ -quasi-automatic sets. To do this, he introduces the idea of a *proper  $k$ -labelling* of a Saguaro digraph.

**Definition 6.5.** Let  $G$  be a rooted Saguaro with edge set  $E$  and vertex set  $V$ . Then a *proper  $k$ -labelling* of  $G$  is a map  $\ell : E \rightarrow \{0, 1, \dots, k-1\}$  satisfying:

- if  $v, w, x \in V$  and  $vw, vx \in E$ , then  $\ell(vw) \neq \ell(vx)$ ;
- if  $v, w, x \in V$ , and  $vw \in E$  lies on a minimal cycle and  $vx \in E$  does not, then  $\ell(vw) > \ell(vx)$ .

**Theorem 6.6.** (Kedlaya [10, Theorem 7.1.6]). *Let  $M$  be a minimal, well-formed deterministic finite automaton with input alphabet  $\Sigma_k^+$ . Then  $M$  is well-ordered if and only if for each relevant post-radix state  $q$ , the subgraph  $G_q$  of the transition graph consisting of relevant states that can be reached from  $q$  is a rooted Saguaro with a proper  $k$ -labelling and root  $q$ .*

**Lemma 6.7.** *Let  $G$  be a Saguaro digraph with a proper  $k$ -labelling, and let  $v, w$  be two vertices in  $G$ . Then the collection of labelled paths from  $v$  to  $w$*

is a finite union of sets of the form

$$(1) \quad \{V_0 W_1^{i_1} V_1 W_2^{i_2} \cdots W_d^{i_d} V_d \mid i_1, \dots, i_d \geq 0\},$$

where  $W_i$  is a word corresponding to a labelled minimal cycle for  $1 \leq i \leq d$  and  $V_j$  contains no cycles for  $0 \leq j \leq d$ .

*Proof.* Let  $\mathcal{C}$  denote the (finite) set of minimal cycles in  $G$ . We create a digraph  $\widehat{G}$  whose vertices are the elements of  $\mathcal{C}$ , in which we declare that there is an edge between two cycles  $C_1$  and  $C_2$  if there is a directed path in  $G$  from a vertex in  $C_1$  to a vertex in  $C_2$ . We note that  $\widehat{G}$  is an acyclic digraph since  $G$  is a Saguario.

Observe that any labelled path in  $G$  from  $v$  to  $w$  can be written in the form

$$V_0 W_1^{i_1} V_1 W_2^{i_2} \cdots W_d^{i_d} V_d$$

for some  $d \geq 0$ , where  $W_1, \dots, W_d$  are words corresponding to minimal cycles. Since  $\widehat{G}$  is acyclic, and there are only finitely many acyclic paths connecting two cycles, we immediately see that the collection of labelled paths from  $v$  to  $w$  is a finite union of sets of the form given in item (1).  $\square$

**Proposition 6.8.** *Let  $S$  be a  $k$ -Saguario set. Then  $S$  is contained in a finite union of sets of the form*

$$\left\{ \sum_{i=0}^d \alpha_i k^{-m_i} \mid m_0, \dots, m_d \in \mathbb{Z}_{\geq 0} \right\},$$

where  $d$  is a nonnegative integer and  $\alpha_0, \dots, \alpha_d$  are rational numbers with  $\alpha_0 > 0$  and  $\alpha_1, \dots, \alpha_d < 0$ .

*Proof.* By Kedlaya's theorem, there is a Saguario digraph  $G$  with a proper  $k$ -labelling and vertices  $v$  and  $w$  such that elements in  $S$  correspond to the labelled paths from  $v$  to  $w$  via the correspondence  $W \mapsto [.W]_k$ .

By Lemma 6.7, it is sufficient to prove that a set of the form given in item (1) satisfies the conclusion of the statement of the proposition.

Let  $x \in [0, 1]$  be a real number with base  $k$  expansion of the form

$$.V_0 W_1^{i_1} V_1 W_2^{i_2} V_2 \cdots W_d^{i_d} V_d.$$

For convenience, define  $W_0$  and  $W_{d+1}$  to be empty words and define

$$a_i := \text{length}(W_i) \quad \text{for } 0 \leq i \leq d+1,$$

$$b_i := \text{length}(V_i) \quad \text{for } 0 \leq i \leq d,$$

and

$$m_j = \text{length}(V_0 W_1^{i_1} \cdots V_{j-1} W_j^{i_j}) \quad \text{for } 0 \leq j \leq d.$$

Then

$$\begin{aligned}
x &= [.V_0 W_1^{i_1} V_1 W_2^{i_2} \cdots W_d^{i_d} V_d]_k \\
&= \sum_{j=0}^{d-1} k^{-m_j} [.V_j W_{j+1}^{i_{j+1}}]_k + k^{-m_d} [.V_d]_k \\
&= \sum_{j=0}^{d-1} k^{-m_j} \left( [.V_j]_k + k^{-b_j} [.W_{j+1}^{i_{j+1}}]_k \right) + k^{-m_d} [.V_d]_k \\
&= \sum_{j=0}^{d-1} k^{-m_j} \left( [.V_j]_k + k^{-b_j} [.W_{j+1}]_k \frac{(1 - k^{-a_{j+1} i_{j+1}})}{(1 - k^{-a_{j+1}})} \right) + k^{-m_d} [.V_d]_k \\
&= \sum_{j=0}^{d-1} k^{-m_j} \left( [.V_j]_k + k^{-b_j} (1 - k^{-a_{j+1}})^{-1} [.W_{j+1}]_k \right) \\
&\quad - \sum_{j=0}^{d-1} k^{-m_j - b_j - a_{j+1} i_{j+1}} (1 - k^{a_{j+1}})^{-1} [.W_{j+1}]_k + k^{-m_d} [.V_d]_k \\
&= \sum_{j=0}^{d-1} k^{-m_j} \left( [.V_j]_k + k^{-b_j} (1 - k^{-a_{j+1}})^{-1} [.W_{j+1}]_k \right) - \\
&\quad \sum_{j=0}^{d-1} k^{-m_{j+1}} (1 - k^{a_{j+1}})^{-1} [.W_{j+1}]_k + k^{-m_d} [.V_d]_k \\
&= \sum_{j=0}^d k^{-m_j} \left( [.V_j]_k + k^{-b_j} (1 - k^{-a_{j+1}})^{-1} [.W_{j+1}]_k - (1 - k^{a_j})^{-1} [.W_j]_k \right).
\end{aligned}$$

Take

$$\alpha_i = [.V_j]_k + k^{-b_j} (1 - k^{-a_{j+1}})^{-1} [.W_{j+1}]_k - (1 - k^{a_j})^{-1} [.W_j]_k$$

for  $0 \leq i \leq d$ .

Then we see that every element in the set

$$\left\{ [.V_0 W_1^{i_1} V_1 \cdots W_d^{i_d} V_d]_k \mid i_1, \dots, i_d \geq 0 \right\}$$

is of the form  $\sum_{i=0}^d \alpha_i k^{-m_i}$  for some nonnegative integers  $m_0, \dots, m_d$ . To finish the proof, note that  $\alpha_0 = [.V_0]_k + k^{-b_0} (1 - k^{-a_1})^{-1} [.W_1]_k > 0$  and for  $1 \leq j \leq d$  we have

$$\begin{aligned}
\alpha_j &= \left( [.V_j]_k + k^{-b_j} (1 - k^{-a_{j+1}})^{-1} [.W_{j+1}]_k - (1 - k^{a_j})^{-1} [.W_j]_k \right) \\
&= [.V_j W_{j+1} W_{j+1} \cdots]_k - [.W_j W_j \cdots]_k.
\end{aligned}$$

Since the digraph  $G$  has a proper  $k$ -labelling, the first letter of  $V_j$  is smaller than the first letter of  $W_j$ . Hence  $\alpha_j \leq 0$  for  $1 \leq j \leq d$ . Of course, it is no loss of generality to assume that  $\alpha_j < 0$  for  $1 \leq j \leq d$ , since zero terms can be ignored. This completes the proof.  $\square$

## 7. COBHAM'S THEOREM AND SAGUARO SETS

In this section we obtain an analogue of Cobham's theorem for Saguaro sets. More specifically, we prove the following result.

**Theorem 7.1.** *Let  $k$  and  $\ell$  be multiplicatively independent positive integers. If  $S$  is a set that is both  $k$ -Saguaro and  $\ell$ -Saguaro, then  $S$  is finite.*

To prove this result, we need to use the theory of  $S$ -unit equations [6, §1.5].

**Lemma 7.2.** *Let  $d$  and  $e$  be two nonnegative integers and suppose that  $\alpha_0, \dots, \alpha_d, \beta_0, \dots, \beta_e$  are rational numbers. Then there are only finitely many integer solutions  $(m_0, \dots, m_d, n_0, \dots, n_e) \in \mathbb{Z}^{d+e+2}$  which satisfy the following three conditions:*

- (1)  $\sum_j \alpha_j k^{m_j} = \sum_j \beta_j \ell^{n_j}$ .
- (2)  $\sum_{j \in I} \alpha_j k^{m_j}$  is nonzero for every nonempty subset  $I$  of  $\{0, 1, \dots, d\}$ .
- (3)  $\sum_{j \in I} \beta_j \ell^{n_j}$  is nonzero for every nonempty subset  $I$  of  $\{0, 1, \dots, e\}$ .

*Proof.* Suppose this is not true. We well-order  $\mathbb{Z}^2$  by declaring that  $(d, e) < (d', e')$  if and only if  $d < d'$  or  $d = d'$  and  $e < e'$ . Pick  $(d, e) \in \mathbb{Z}_{\geq 0}^2$  minimal with respect to the property that there exist  $\alpha_0, \dots, \alpha_d, \beta_0, \dots, \beta_e$  such that there exist infinitely many integer solutions satisfying conditions (1), (2), and (3).

Then the minimality of  $d$  and  $e$  and conditions (2) and (3) show that there are infinitely many solutions  $(m_0, \dots, m_d, n_0, \dots, n_e)$  such that no non-trivial proper sub-sum of the expression

$$\alpha_0 k^{m_0} + \dots + \alpha_d k^{m_d} - \beta_0 \ell^{n_0} - \dots - \beta_e \ell^{n_e}$$

is equal to zero.

Let  $A$  be the finitely generated multiplicative subgroup of  $\mathbb{Q}^\times$  generated by  $\alpha_0, \dots, \alpha_d, \beta_0, \dots, \beta_e$ , and  $k, \ell$ , and  $-1$ . Then Everest *et al.* [6, Theorem 1.19] show that the equation

$$X_1 + \dots + X_{d+e+2} = 0$$

has only finitely many solutions  $(a_1, \dots, a_{d+e+2}) \in A^{d+e+2}$  up to multiplication by  $A$  if we assume that no proper sub-sum vanishes. Equation (1) thus admits two distinct solutions,  $(\alpha_0 k^{m_0}, \dots, \alpha_d k^{m_d}, -\beta_0 \ell^{n_0}, \dots, -\beta_e \ell^{n_e})$  and  $(\alpha_0 k^{m'_0}, \dots, \alpha_d k^{m'_d}, -\beta_0 \ell^{n'_0}, \dots, -\beta_e \ell^{n'_e})$ , that differ by a multiple  $a \in A$ . Then, for  $0 \leq i \leq d$  and  $0 \leq j \leq e$ , we have

$$k^{m_i - m'_i} = \ell^{n_j - n'_j} = a.$$

But  $k$  and  $\ell$  are multiplicatively independent, and hence the only way this can occur is if  $m_i - m'_i = 0 = n_j - n'_j$  for  $0 \leq i \leq d$  and  $0 \leq j \leq e$ . This implies that our two solutions are equal, and so we obtain a contradiction. This ends the proof.  $\square$

*Proof of Theorem 7.1.* By Proposition 6.8 it is sufficient to show that the intersection of two sets

$$T_1 := \left\{ \sum_{i=0}^d \alpha_i k^{-m_i} \mid m_0, \dots, m_d \in \mathbb{Z}_{\geq 0} \right\},$$

and

$$T_2 := \left\{ \sum_{i=0}^e \beta_i \ell^{-n_i} \mid n_0, \dots, n_e \in \mathbb{Z}_{\geq 0} \right\}$$

with  $\alpha_0, \beta_0 > 0$  and  $\alpha_i, \beta_j < 0$  for  $1 \leq i \leq d$ ,  $1 \leq j \leq e$  can have at most finitely many elements in  $[0, 1]$ .

Any point  $x$  lying in the set  $T_1 \cap T_2$  gives rise to an integer solution  $(m_0, \dots, m_d, n_0, \dots, n_e)$  to the equation

$$\sum_{j=0}^d \alpha_j k^{m_j} = \sum_{j=0}^e \beta_j \ell^{n_j}.$$

If, in addition to this,  $x \in (0, 1]$  then we have that

$$\sum_{j=0}^d \alpha_j k^{m_j} > 0.$$

Since  $\alpha_0 k^{m_0}$  is the only positive term which appears in this sum, we see that

$$\sum_{j \in I} \alpha_j k^{m_j}$$

is nonzero for every nonempty subset  $I$  of  $\{0, 1, \dots, d\}$ . Similarly,

$$\sum_{j \in I} \beta_j \ell^{n_j}$$

is nonzero for every nonempty subset  $I$  of  $\{0, 1, \dots, e\}$ . Hence conditions 1, 2, and 3 in the statement of Lemma 7.2 are satisfied. It follows that there are only finitely many  $x \in T_1 \cap T_2 \cap (0, 1]$ . Of course, possibly adding 0 into  $T_1 \cap T_2$  does not affect the fact that there are only finitely many points in the intersection of  $T_1$  and  $T_2$  that are also in  $[0, 1]$ . The result follows.  $\square$

*Remark 7.3.* If  $k$  and  $\ell$  are relatively prime, then the Proof of Theorem 7.1 is trivial. The reason for this is that if  $k$  and  $\ell$  are relatively prime, then  $S_k \cap S_\ell = \mathbb{Z}_{\geq 0}$  and so any set which is both  $k$ - and  $\ell$ -Saguaro will consist of at most two points; namely 0 and 1.

## 8. PROOFS OF THEOREMS 2.1 AND 2.2

In this section we give a proof of our main results.

We first need a simple lemma.

**Lemma 8.1.** *Let  $S$  be the support of a  $K_k$ -automatic function. Then the set*

$$S_1 := \{x \in [0, 1] \mid x + m \in S \text{ for some } m \in \mathbb{Z}\}$$

*is a  $k$ -Saguaro set.*

*Proof.* By assumption, there is some finite state machine which accepts words in  $\Sigma'_k$  such that for  $W \in \mathcal{L}(k)$ ,  $[W]_k \in S$  if and only if the machine outputs 1 when  $W$  is the input. There are only finitely many possible transitions that can occur when the radix point is entered and thus  $S_1$  is easily seen to be a finite union of  $k$ -Saguaro sets. The fact that a finite union of  $k$ -Saguaro sets is itself a  $k$ -Saguaro set can be obtained by mimicking the

proof of assertion (b) in Theorem 5.6.3 of [1]. The latter result is that a finite union of  $k$ -automatic sets is itself a  $k$ -automatic set.  $\square$

We are now ready to prove Theorems 2.1 and 2.2.

*Proof of Theorem 2.1.* Since  $f$  is  $k$ -quasi-automatic and  $\ell$ -quasi-automatic, there exist integers  $a_1, a_2, b_1, b_2$  with  $a_1, a_2 > 0$  such that  $f((x - b_1)/a_1)$  has support in  $S_k$  and is  $K_k$ -automatic, and such that  $f((x - b_2)/a_2)$  has support in  $S_\ell$  and is  $K_\ell$ -automatic. Let  $a = a_1 a_2$  and let  $b = |a_1 b_2| + |a_2 b_1|$ . Then  $f((x - b)/a)$  has support in both  $S_k$  and  $S_\ell$  and is both  $K_k$ - and  $K_\ell$ -automatic.

Let  $S$  denote the support of  $f(ax + b)$  and let  $S_1 = \{x \in [0, 1] \mid x + m \in S \text{ for some } m \in \mathbb{Z}\}$ . By Lemma 8.1,  $S_1$  is both a  $k$ -Saguaro and a  $\ell$ -Saguaro and is thus a finite set by Theorem 7.1. It follows that there is a positive integer  $n$  such that  $nS \subseteq \mathbb{N}$ .

Replacing  $a$  by  $na$  and  $b$  by  $nb$  we may assume that the support of  $h(x) := f((x - b)/a)$  is the set of nonnegative integers; that is  $h(x) = 0$  for  $x \in \mathbb{Q} \setminus \mathbb{N}$ . Notice that  $h|_{\mathbb{N}}$  is an ordinary  $k$ - and  $\ell$ -automatic function and hence is eventually periodic in virtue of Cobham's theorem. The result follows.  $\square$

*Proof of Theorem 2.2.* We keep the notation as in Theorem 2.2. We first remark that, given a positive integer  $n$ , the field  $\mathbb{F}_q(t^{1/n})$  is an algebraic extension (of degree  $n$ ) of the field of rational functions  $\mathbb{F}_q(t)$ . Since  $f(t^n)$  is a rational function (with respect to the indeterminate  $t$ ) if and only if  $f$  lies in  $\mathbb{F}_q(t^{1/n})$ , we obtain that such a function  $f$  is algebraic over  $\mathbb{F}_q(t)$  (of degree at most  $n$ ). This proves the first part of the Theorem.

Now, let us assume that  $f(t)$  and  $g(t)$  are both algebraic. Theorem 5.4 implies that the function  $h_1 : \mathbb{Q} \mapsto \mathbb{F}_{q_1}$  defined by  $h_1(\alpha) = i_1(r_\alpha)$ , is a  $p_1$ -quasi-automatic function. Just as, the function  $h_2 : \mathbb{Q} \mapsto \mathbb{F}_{q_2}$  defined by  $h_2(\alpha) = i_2(r_\alpha)$ , is a  $p_2$ -quasi-automatic function. Since  $i_1$  and  $i_2$  are two injections, this straightforwardly implies that the function  $h : \mathbb{Q} \mapsto \mathcal{A}$  that maps  $\alpha$  to  $r_\alpha$  is both  $p_1$ - and  $p_2$ -quasi-automatic. By Theorem 2.1, there exist integers  $a$  and  $b$  with  $a > 0$  such that

$$f(t) = \sum_{\alpha \in \mathbb{Q}} i_1(r_\alpha) t^\alpha = \sum_{k \geq 0} i_1(r_{(k-b)/a}) t^{(k-b)/a} = (t^{1/a})^{-b} \sum_{k \geq 0} a_k (t^{1/a})^k,$$

where  $(a_k)_{k \geq 0} := (i_1(r_{(k-b)/a}))_{k \geq 0}$ . Moreover, Theorem 2.1 also implies under our assumptions that the sequence  $(a_k)_{k \geq 0}$  is eventually periodic. Consequently,  $f(t)$  is a rational function in  $t^{1/a}$  and thus  $f(t^a)$  lies in the field  $\mathbb{F}_{q_1}(t)$ . The same reasoning applies to  $g(t)$ . This completes the proof.  $\square$

## 9. ACKNOWLEDGMENTS

We thank the referee for his careful reading of the manuscript.

## REFERENCES

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

- [2] C. Chevalley, Introduction to the Theory of Algebraic Functions of One Variable, Mathematical Surveys No. VI, Amer. Math. Soc., 1951.
- [3] G. Christol, Ensembles presque périodiques  $k$ -reconnaissables, *Theor. Comput. Sci.* **9** (1979), 141–145.
- [4] G. Christol, T. Kamae, M. Mendès France & G. Rauzy, Suites algébriques, automates et substitutions, *Bull. Soc. Math. France* 108 (1980), 401–419.
- [5] A. Cobham, On the base-dependence of sets of numbers recognizable by finite automata, *Math. Systems Theory* **3** (1969), 186–192.
- [6] G. Everest, A. van der Poorten, I. Shparlinski & T. Ward, Recurrence Sequences, Mathematical Surveys and Monographs **104**, American Mathematical Society, Providence, RI, 2003.
- [7] H. Hahn, Über die nichtarchimedische Größensysteme, *Wien. Ber.* **116** (1907), 601–655.
- [8] I. Kaplansky, Maximal fields with valuations, *Duke Math. J.* **9** (1942), 303–321.
- [9] K. Kedlaya, The algebraic closure of the power series field in positive characteristic, *Proc. Amer. Math. Soc.* **129** (2001), 3461–3470.
- [10] K. Kedlaya, Finite automata and algebraic extensions of function fields, *J. Théor. Nombres Bordeaux*, to appear.
- [11] M. Mendès France, Sur les décimales des nombres algébriques réels, in *Sémin. Théor. Nombres, Bordeaux, 1979–1980*, Exp. No. 28, 7 pp., Univ. Bordeaux I, Talence, 1980.
- [12] J.-P. Serre, Local Fields, Springer–Verlag, 1979.

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

`Boris.Adamczewski@math.univ-lyon1.fr`

Jason Bell  
Department of Mathematics  
Simon Fraser University  
8888, University Drive  
Burnaby BC V5A 1S6  
(CANADA)

`jpb@sfu.ca`