A Mahler's theorem for functions from words to integers *

Jean-Éric Pin† and Pedro V. Silva‡

December 16, 2007

Abstract

In this paper, we prove an extension of Mahler's theorem, a celebrated result of p-adic analysis. Mahler's original result states that a function from $\mathbb{N}$ to $\mathbb{Z}$ is uniformly continuous for the p-adic metric d_p if and only if it can be uniformly approximated by polynomial functions. We prove the same result for functions from A^* to $\mathbb{Z}$, where d_p is now the profinite metric defined by p-groups (pro-p metric).

This paper was originally motivated by two main research lines of automata theory, but resulted into an approximation theorem that goes far beyond our original project. We first present our original motivations and then describe our main result. Recall that a *variety of languages* is a class of regular languages closed under Boolean operations, left and right quotients and inverse morphisms.

1 Motivations

Our first motivation was the study of *regularity-preserving* functions f from A^* to B^*, in the following sense: if X is a regular language of B^*, then $f^{-1}(X)$ is a regular language of A^*. More generally, we were interested in functions *preserving a given variety of languages* $\mathcal{V}$: if X is a language of $\mathcal{V}$, then $f^{-1}(X)$ is also a language of $\mathcal{V}$. There is an important literature on the regular case [20, 6, 18, 12, 13, 2], including the authors recent paper [14]. A similar problem was also recently considered for formal power series [3]. A remarkable contribution to the second problem can be found in [16], where a characterization of sequential functions preserving aperiodic languages (respectively group-languages) is given.

Our second motivation was the study of certain reductions. A fundamental idea of descriptive set theory is to use *continuous reductions* to classify topological spaces: given two sets X and Y, Y reduces to X if there exists a continuous function f such that $X = f^{-1}(Y)$. For Polish spaces, this gives rise to the Wadge hierarchy [21], nowadays entirely described and very well understood. Wagner [22] was the first to consider the restriction of the Wadge hierarchy to ω-rational languages. He proved in particular that

*The authors acknowledge support from the AutoMathA programme of the European Science Foundation. The second author acknowledges support from Project ASA (PTDC/MAT/65481/2006) and C.M.U.P., financed by F.C.T. (Portugal) through the programmes POCTI and POSI, with national and European Community structural funds.

†LIAFA, Université Paris-Diderot and CNRS, Case 7014, 75205 Paris Cedex 13, France.

‡Centro de Matemática, Faculdade de Ciências, Universidade do Porto, R. Campo Alegre 687, 4169-007 Porto, Portugal.

the hierarchy doesn't change if continuous functions are replaced by sequential functions, a much more restricted class of functions (see [10, Chapter 5] for more details).

Our idea was to consider similar reductions for regular languages (of finite words this time). The first obstacle was to find an appropriate topology, but there is a natural candidate: the profinite topologies, a notion first introduced in [5]. Indeed, by Eilenberg's theorem, to each variety of languages $\mathcal{V}$ corresponds a unique variety of finite monoids $\mathbf{V}$, which in turn defines the pro-$\mathbf{V}$ topology [15]. We shall not give here the precise definition, but it suffices to know that, in the most interesting cases, this topology can be defined by a metric $d_{\mathbf{V}}$. Let us call $\mathbf{V}$*-reduction* a uniformly continuous function between the metric spaces $(A^*, d_{\mathbf{V}})$ and $(B^*, d_{\mathbf{V}})$. These $\mathbf{V}$-reductions define a hierarchy similar to the Wadge hierarchy among regular languages, which we would like to explore. Note that a different notion of reduction for regular languages was recently considered in [19]. The first author had very instructive discussions with Selivanov and Kunc in June 2006 on the comparison between these two reductions and this paper is partly motivated by this conversation.

Regularity-preserving functions and $\mathbf{V}$-reductions are actually strongly related. Indeed, one can show that a function from $(A^*, d_{\mathbf{V}})$ to $(B^*, d_{\mathbf{V}})$ is uniformly continuous if and only if, for every language L in $\mathcal{V}(B^*)$, $f^{-1}(L)$ belongs to $\mathcal{V}(A^*)$. This encouraging fact lead us to search for a more precise description of $\mathbf{V}$-reductions. However, apart from general results, not so much is known on pro-$\mathbf{V}$ topologies, except when $\mathbf{V}$ is a variety of finite groups. Among groups, the variety $\mathbf{G}_p$ of p-groups, where p is a given prime, is of special interest for two reasons. First, Eilenberg and Schützenberger gave a very nice description of the languages recognized by a p-group (see Proposition 2.2 below). Second, a special case of the metric d_p has been widely studied in mathematics: indeed, the free monoid over a one-letter alphabet is isomorphic to $\mathbb{N}$, and the metric d_p is known as the *p-adic metric*. The completion of the metric space $(\mathbb{N}, d_p)$ is the space of *p-adic numbers* and *p-adic analysis* is the branch of number theory that deals with functions of p-adic numbers [1, 17, 9].

Our main result takes advantage of this powerful mathematical framework to provide a characterization of the $\mathbf{G}_p$-reductions from A^* to $\mathbb{N}$, that is, the uniformly continuous functions from (A^*, d_p) to $(\mathbb{N}, d_p)$. It turns out that this characterization extends a celebrated result of number theory, Mahler's theorem (see `http://en.wikipedia.org/wiki/Mahler's_theorem`), giving our result a mathematical interest on its own. Our result states that a function from A^* to $\mathbb{N}$ is uniformly continuous for d_p if and only if it can be uniformly approximated by a sequence of polynomial functions. Before stating this result in a precise form, we need a few formal definitions.

2 The p-adic and pro-p topologies

In the sequel, A denotes a finite alphabet, A^* is the free monoid on A and 1 denotes the empty word.

Let p be a prime number. Recall that a *p-group* is a finite group whose order is a power of p. Let u and v be two words of A^*. A p-group G *separates* u and v if there is a monoid morphism from A^* onto G such that $\varphi(u) \neq \varphi(v)$. One can show that any pair of distinct words can be separated by a p-group. We now set

$$r_p(u, v) = \min \{ \ n \mid \text{there is a } p\text{-group of order } p^n \text{ separating } u \text{ and } v \ \}$$
$$d_p(u, v) = 2^{-r_p(u,v)}$$

with the usual convention $\min \emptyset = -\infty$ and $2^{-\infty} = 0$. One can show that d_p is an *ultrametric*, that is, satisfies the following properties, for all $u, v, w \in A^*$:

(1) $d_p(u, v) = 0$ if and only if $u = v$,

(2) $d_p(u, v) = d_p(v, u)$,

(3) $d_p(u, v) \leqslant \max(d_p(u, w), d_p(w, v))$

One can also show that the concatenation product on A^* is uniformly continuous for this metric. It follows that the completion of the metric space (A^*, d_p) is naturally equipped with a structure of monoid, which is in fact a compact group, called the *free pro-p group*. The topology defined by the metric d_p is usually called the *pro-p topology* in the literature.

There is a nice connection [11] between this topology and a generalization of the *binomial coefficients*. Let u and v be two words of A^*. Let $u = a_1 \cdots a_n$, with $a_1, \ldots, a_n \in A$. Then u is a *subword* of v if there exist $v_0, \ldots, v_n \in A^*$ such that $v = v_0 a_1 v_1 \ldots a_n v_n$. Following [4, 7], we define the binomial coefficient of u and v by setting

$$\binom{v}{u} = |\{(v_0, \ldots, v_n) \mid v = v_0 a_1 v_1 \ldots a_n v_n\}|\,.$$

Observe that if a is a letter, then $\binom{v}{a}$ is simply the number of occurrences of a in v, also denoted by $|v|_a$. Also note that if $A = \{a\}$, $u = a^n$ and $v = a^m$, then

$$\binom{v}{u} = \binom{m}{n}$$

and hence these numbers constitute a generalization of the classical binomial coefficients. The next proposition, whose proof can be found in [7, Chapter 6], summarizes the basic properties of the generalized binomial coefficients and can serve as an alternative definition.

Lemma 2.1 *Let $u, v \in A^*$ and $a, b \in A$. Then*

(1) $\binom{u}{1} = 1$,

(2) $\binom{u}{v} = 0$ *if* $|u| \leqslant |v|$ *and* $u \neq v$,

(3) $\binom{ua}{vb} = \begin{cases} \binom{u}{vb} & \text{if } a \neq b \\ \binom{u}{vb} + \binom{u}{v} & \text{if } a = b \end{cases}$

A third way to define the binomial coefficients is to use the *Magnus automorphism* of the algebra $\mathbb{Z}\langle A\rangle$ of polynomials in noncommutative indeterminates in A defined by $\mu(a) = 1 + a$ for all $a \in A$. One can show that, for all $u \in A^*$,

$$\mu(u) = \sum_{x \in A^*} \binom{u}{x} x \tag{2.1}$$

which leads to the formula

$$\binom{u_1 u_2}{x} = \sum_{x_1 x_2 = x} \binom{u_1}{x_1}\binom{u_2}{x_2} \tag{2.2}$$

The connection between the pro-p topology and the binomial coefficients comes from the characterization of the languages recognized by a p-group given by Eilenberg and Schützenberger (see [4, Theorem 10.1, p. 239]). Let us call a p-group language a language recognized by a p-group. Note that such a language is necessarily regular.

Proposition 2.2 *A language of A^* is a p-group language if and only if it is a Boolean combination of the languages*

$$L(x, r, p) = \{u \in A^* \mid \binom{u}{x} \equiv r \bmod p\},$$

for $0 \leqslant r < p$ and $x \in A^*$.

Let us set now

$$r'_p(u,v) = \min\left\{ |x| \mid x \in A^* \text{ and } \binom{u}{x} \not\equiv \binom{v}{x} \pmod p \right\}$$

$$d'_p(u,v) = p^{-r'_p(u,v)}.$$

It is proved in [11, Theorem 4.4] that d'_p is an ultrametric uniformly equivalent to d_p. We shall use this result under a slightly different form.

Theorem 2.3 *A function $f : A^* \to B^*$ is uniformly continuous for d_p if and only if, for every regular language L of A^* recognized by a p-group, the language $f^{-1}(L)$ is also recognized by a p-group.*

Proof. Let L be a language recognized by a p-group G of order p^k. Then there exists a monoid morphism $\varphi : A^* \to G$ such that $L = \varphi^{-1}(\varphi(L))$. If f is uniformly continuous for d_p, there exists $n > 0$ such that if $r_p(u,v) \geqslant n$, then $r_p(f(u), f(v)) \geqslant k$. It follows in particular that $f(u)$ and $f(v)$ cannot be separated by G and hence $\varphi(f(u)) = \varphi(f(v))$.

Let $(\psi_i)_{i\in I}$ be the family of all monoid morphisms from A^* onto a p-group H_i of order $\leqslant p^n$. Let $\psi : A^* \to \prod_{i\in I} H_i$ be the morphism defined by $\psi(x) = (\psi_i(x))_{i\in I}$ and let H be the range of ψ. Then H is a p-group and if $\psi(u) = \psi(v)$, then $r_p(u,v) \geqslant n$ and thus $\varphi(f(u)) = \varphi(f(v))$. We claim that

$$\psi^{-1}(\psi(f^{-1}(L)) = f^{-1}(L)$$

First, $f^{-1}(L)$ is clearly a subset of $\psi^{-1}(\psi(f^{-1}(L))$. To prove the opposite inclusion, let $u \in \psi^{-1}(\psi(f^{-1}(L))$. Then $\psi(u) \in \psi(f^{-1}(L))$, that is, $\psi(u) = \psi(v)$ for some $v \in f^{-1}(L)$. It follows that $\varphi(f(u)) = \varphi(f(v))$ and since $f(v) \in L$, $f(u) \in \varphi^{-1}(\varphi(L))$ and finally $f(u) \in L$ since $L = \varphi^{-1}(\varphi(L))$. This proves the claim and shows that $f^{-1}(L)$ is a p-group language.

Suppose now that if L is a p-group language, then $f^{-1}(L)$ is also a p-group language. Let φ be a morphism from A^* onto a p-group G. For each $g \in G$, $\varphi^{-1}(g)$ is a p-group language and hence $f^{-1}(\varphi^{-1}(g))$ is recognized by a morphism $\psi_g : A^* \to H_g$ onto a p-group. Let $\psi : A^* \to \prod_{g\in G} H_g$ be the mapping defined by $\psi(x) = (\psi_g(x))_{g\in G}$ and let $H = \psi(A^*)$. Then H is also a p-group and if $\psi(u) = \psi(v)$, then $\psi_g(u) = \psi_g(v)$ for all $g \in G$. Since ψ_g recognizes $f^{-1}(\varphi^{-1}(g))$, it follows that $u \in f^{-1}(\varphi^{-1}(g))$ if and only if $v \in f^{-1}(\varphi^{-1}(g))$ and hence $\varphi(f(u)) = \varphi(f(v))$.

Now let $k \in \mathbb{N}$. If we consider all the morphisms φ from A^* onto a p-group of order $\leqslant p^k$, and take $n \in \mathbb{N}$ large enough so that every group H corresponding to φ has order $\leqslant p^n$, it follows that

$$r_p(u,v) > n \Rightarrow r_p(f(u), f(v)) > k$$

holds for all $u, v \in A^*$. This shows that f is uniformly continuous for d_p. $\square$

In the case of a one-letter alphabet, A^* is isomorphic to the additive monoid $\mathbb{N}$ and the definition of d_p can be further simplified. If n is a non-zero integer, recall that the *p-adic valuation* of n is the integer

$$\nu_p(n) = \max\left\{k \in \mathbb{N} \mid p^k \text{ divides } n\right\}$$

By convention, $\nu_p(0) = +\infty$. The *p-adic norm* of n is the real number

$$|n|_p = p^{-\nu_p(n)}.$$

The p-adic norm satisfies the following axioms, for all $n, m \in \mathbb{N}$:

(1) $|n|_p \geqslant 0$,

(2) $|n|_p = 0$ if and only if $n = 0$,
(3) $|mn|_p = |m|_p|n|_p$,
(4) $|m + n|_p \leqslant \max\{|m|_p, |n|_p\}$.
Finally, the metric d_p can be defined by

$$d_p(u, v) = |u - v|_p.$$

and is known as the p-adic metric. Since $\mathbb{Z}$ is naturally embedded in the p-adic completion of $\mathbb{N}$, the definitions above can be readily extended to $\mathbb{Z}$. In the sequel, it will be more convenient to use the metric space $(\mathbb{Z}, d_p)$ in place of $(\mathbb{N}, d_p)$.

3 Mahler's expansions

The classical Stone-Weierstrass approximation theorem states that a continuous function defined on a closed interval can be uniformly approximated by a polynomial function. In particular, if a function f is infinitely differentiable in the neighbourhood of 0, it can be approximated, under some convergence conditions, by its Taylor polynomials

$$\sum_{n=0}^{k} \frac{f^{(n)}(0)}{n!}\, x^n$$

The p-adic analogue of these results is Mahler's theorem [8]. For a fixed $k \in \mathbb{N}$, the binomial polynomial function

$$u \to \binom{u}{k}$$

defines a uniformly continuous function from $(\mathbb{N}, d_p)$ to $(\mathbb{Z}, d_p)$. The *Mahler's expansion* of a function f from $\mathbb{N}$ to $\mathbb{Z}$ is defined as the series

$$\sum_{k=0}^{\infty} (\Delta^k f)(0) \binom{u}{k}$$

where Δ is the *difference operator*, defined by

$$(\Delta f)(u) = f(u + 1) - f(u)$$

Mahler's theorem states that f is uniformly continuous for d_p if and only if its Mahler's expansion converges uniformly to f. Of course, the most remarkable part of the theorem is the fact that any uniformly continuous function can be approximated by polynomial functions, in contrast to Stone-Weierstrass approximation theorem, which requires much stronger conditions.

For instance, if f is the Fibonacci sequence defined by $f(0) = f(1) = 1$ and $f(n) = f(n - 1) + f(n - 2)$ for $n \geqslant 2$, then

$$f(n) = \sum_{k=0}^{\infty} (-1)^{k+1} f(k) \binom{n}{k}$$

This function is not uniformly continuous for d_p for any choice of p. If $f(n) = r^n$, then

$$f(n) = \sum_{k=0}^{\infty} (r - 1)^k \binom{n}{k}$$

and f is uniformly continuous for d_p if and only if p divides $r - 1$.

The first step to extend Mahler's theorem to functions from words to integers is to define a suitable notion of Mahler's expansion for these functions.

Let $f : A^* \to \mathbb{Z}$ be a function. For each letter a, we define the difference operator Δ^a by

$$(\Delta^a f)(u) = f(ua) - f(u)$$

One can now define inductively an operator Δ^w for each word $w \in A^*$ by setting $(\Delta^1 f)(u) = f(u)$, and for each letter $a \in A$,

$$(\Delta^{aw} f)(u) = (\Delta^a(\Delta^w f))(u).$$

It is easy to see that these operators can also be defined directly by setting

$$\Delta^w f(u) = \sum_{0 \leqslant |x| \leqslant |w|} (-1)^{|w|+|x|} \binom{w}{x} f(ux) \tag{3.3}$$

For instance, $\Delta^{aab} f(u) = -f(u) + 2f(ua) + f(ub) - f(uaa) - 2f(uab) + f(uaab)$.
For a fixed $v \in A^*$, the function

$$u \to \binom{u}{v}$$

from A^* to $\mathbb{Z}$ which maps a word u to the binomial coefficient $\binom{u}{v}$ is uniformly continuous for d_p. This family of functions, for v ranging over A^*, is *locally finite* in the sense that, for each $u \in A^*$, the binomial coefficient $\binom{u}{v}$ is null for all but finitely many words v. In particular, if $(m_v)_{v \in A^*}$ is a family of integers, there is a well-defined function from A^* to $\mathbb{Z}$ defined by the formula

$$f(u) = \sum_{v \in A^*} m_v \binom{u}{v}$$

We can now state our first result, which doesn't require any assumption on f.

Theorem 3.1 *Let $f : A^* \to \mathbb{Z}$ be an arbitrary function. Then there exists a unique family $\langle f, v \rangle_{v \in A^*}$ of integers such that, for all $u \in A^*$,*

$$f(u) = \sum_{v \in A^*} \langle f, v \rangle \binom{u}{v} \tag{3.4}$$

This family is given by

$$\langle f, v \rangle = (\Delta^v f)(1) = \sum_{0 \leqslant |x| \leqslant |v|} (-1)^{|v|+|x|} \binom{v}{x} f(x) \tag{3.5}$$

Proof. First observe that, according to (3.3)

$$(\Delta^v f)(1) = \sum_{0 \leqslant |x| \leqslant |v|} (-1)^{|v|+|x|} \binom{v}{x} f(x) \tag{3.6}$$

Thus

$$\begin{aligned}
\sum_{v \in A^*} (\Delta^v f)(1) \binom{u}{v} &= \sum_{v \in A^*} \sum_{|x| \leqslant |v|} (-1)^{|v|+|x|} \binom{v}{x} \binom{u}{v} f(x) \\
&= \sum_{x \in A^*} (-1)^{|u|+|x|} \left(\sum_{0 \leqslant |v| \leqslant |u|} (-1)^{|v|+|u|} \binom{u}{v} \binom{v}{x} \right) f(x) \\
&= f(u)
\end{aligned}$$

in view of the following relation from [7, Corollary 6.3.8]:

$$\sum_{0 \leqslant |v| \leqslant |u|} (-1)^{|u|+|v|} \binom{u}{v} \binom{v}{w} = \begin{cases} 1 & \text{if } u = w \\ 0 & \text{otherwise} \end{cases} \tag{3.7}$$

Uniqueness of the coefficients $\langle f, v\rangle$ follows inductively from the formula

$$\langle f, u\rangle = f(u) - \sum_{0 \leqslant |v| \, < \, |u|} \langle f, v\rangle \binom{u}{v},$$

a straightforward consequence of (3.4). $\square$

The series defined by (3.4) is called the *Mahler's expansion* of f.

For instance, let $f : \{0,1\}^* \to \mathbb{N}$ be the function mapping a binary word onto its value as a binary number. Thus $f(010111) = f(10111) = 23$. Then one has

$$(\Delta^v f) = \begin{cases} f+1 & \text{if } |v|_1 > 0 \\ f & \text{otherwise} \end{cases}$$

$$(\Delta^v f)(\varepsilon) = \begin{cases} 1 & \text{if } |v|_1 > 0 \\ 0 & \text{otherwise} \end{cases}$$

Thus, if $u = 01001$, one gets

$$f(u) = \binom{u}{1} + \binom{u}{10} + \binom{u}{11} + \binom{u}{100} + \binom{u}{101} + \binom{u}{1001} = 2+2+1+1+2+1 = 9$$

4 Mahler polynomials

A function $f : A^* \to \mathbb{Z}$ is a *Mahler polynomial* if its Mahler's expansion has *finite support*, that is, if the number of nonzero coefficients $\langle f, v\rangle$ is finite. In this section, we prove in particular that Mahler polynomials are closed under addition and product. We first introduce a convenient combinatorial operation, the infiltration product. We follow the presentation of [7].

Let $\mathbb{Z}\langle\langle A\rangle\rangle$ be the ring of formal power series in noncommutative indeterminates in A. Any series s is written as a formal sum $s = \sum_{u \in A^*} \langle s, u\rangle u$, a notation not to be confused with our notation for the Mahler's expansion. The *infiltration product* is the binary operation on $\mathbb{Z}\langle\langle A\rangle\rangle$, denoted by $\uparrow$ and defined inductively as follows:

for all $u \in A^*$,

$$u \uparrow 1 = 1 \uparrow u = u, \tag{4.8}$$

for all $u, v \in A^*$, for all $a, b \in A$

$$ua \uparrow bv = \begin{cases} (u \uparrow vb)a + (ua \uparrow v)b + (u \uparrow v)a & \text{if } a = b \\ (u \uparrow vb)a + (ua \uparrow v)b & \text{otherwise} \end{cases} \tag{4.9}$$

for all $s, t \in \mathbb{Z}\langle\langle A\rangle\rangle$,

$$s \uparrow t = \sum_{u,v \in A^*} \langle s, u\rangle \langle t, v\rangle (u \uparrow v) \tag{4.10}$$

Intuitively, the coefficient $\langle u \uparrow v, x\rangle$ is the number of pairs of subsequences of x which are respectively equal to u and v and whose union gives the whole sequence x. For instance,

$$ab \uparrow ab = ab + 2aab + 2abb + 4aabb + 2abab$$
$$ab \uparrow ba = aba + bab + abab + 2abba + 2baab + baba$$

Also note that $\langle u \uparrow v, u\rangle = \binom{u}{v}$. We shall need the following relation (see [7, p.131]). For all $v_1, v_2 \in A^*$,

$$\binom{u}{v_1}\binom{u}{v_2} = \sum_{x \in A^*} \langle v_1 \uparrow v_2, x\rangle \binom{u}{x} \tag{4.11}$$

Formula 4.11 leads to an explicit computation of the Mahler's expansion of the product of two functions.

Proposition 4.1 *Let f and g be two functions from A^* to $\mathbb{N}$. Then the coefficients of the Mahler's expansion of fg are given by the formula:*

$$\langle fg, x\rangle = \sum_{v_1, v_2 \in A^*} \langle f, v_1\rangle\langle g, v_2\rangle\langle v_1 \uparrow v_2, x\rangle$$

Proof. Indeed, if $f(u) = \sum_{v\in A^*} \langle f, v\rangle\binom{u}{v}$ and $g(u) = \sum_{v\in A^*} \langle g, v\rangle\binom{u}{v}$, then

$$fg(u) = \sum_{v_1, v_2 \in A^*} \langle f, v_1\rangle\langle g, v_2\rangle\binom{u}{v_1}\binom{u}{v_2}$$

and the result follows by Formula (4.11). □

It is now easy to prove the result announced at the beginning of this section.

Proposition 4.2 *Mahler polynomials form a subring of the ring of all functions from A^* to $\mathbb{Z}$ for addition and multiplication.*

Proof. It is clear that the difference of two Mahler polynomials is a Mahler polynomial. Further Proposition 4.1 shows that Mahler polynomials are closed under product. □

5 Mahler's theorem

We are now ready to state our main result, which extends Mahler's theorem. In this section, uniform continuity always refers to the metric d_p.

Theorem 5.1 *Let $f(u) = \sum_{v\in A^*} \langle f, v\rangle\binom{u}{v}$ be the Mahler's expansion of a function from A^* to $\mathbb{Z}$. Then f is uniformly continuous if and only if $\lim_{|v|\to\infty} |\langle f, v\rangle|_p = 0$.*

Proof. Suppose that $\lim_{|v|\to\infty} |\langle f, v\rangle|_p = 0$. Then there exists $s \in \mathbb{N}$ such that, if $|v| \geqslant s$, $\nu_p(\langle f, v\rangle) \geqslant r$. Setting

$$g(u) = \sum_{|v|<s} \langle f, v\rangle\binom{u}{v} \text{ and } h(u) = \sum_{|v|\geqslant s} \langle f, v\rangle\binom{u}{v}$$

we get $f = g + h$. Further p^r divides $\langle f, v\rangle$ for $|v| \geqslant s$. Since g is a Mahler polynomial, it is uniformly continuous and there exists $t \in \mathbb{N}$ such that $d_p(u, u') \leqslant p^{-t}$ implies $g(u) \equiv g(u') \pmod{p^r}$ and hence $f(u) \equiv f(u') \pmod{p^r}$. Thus f is uniformly continuous.

This proves the easy direction of the theorem. The key argument for the opposite direction is the following approximation result.

Theorem 5.2 *Let $f : A^* \to \mathbb{N}$ be a uniformly continuous function. Then there exists a Mahler polynomial P such that, for all $u \in A^*$, $f(u) \equiv P(u) \pmod p$.*

Proof. We first prove the theorem for some characteristic functions related to the binomial coefficients. The precise role of these functions will appear in the course of the main proof.

Let $x \in A^*$ and let s be an integer such that $0 \leqslant s < p$. Let $\chi_{s,x} : A^* \to \mathbb{N}$ be the function defined by

$$\chi_{s,x}(u) = \begin{cases} 1 & \text{if } \binom{u}{x} \equiv s \pmod p \\ 0 & \text{otherwise} \end{cases}$$

Lemma 5.3 *There is a Mahler polynomial $P_{s,x}$ such that, for all $u \in A^*$, $\chi_{s,x}(u) \equiv P_{s,x}(u) \pmod p$.*

Proof. Let

$$P_{s,x}(u) = -\frac{\left[\binom{u}{x}\right]\left[\binom{u}{x} - 1\right] \cdots \left[\binom{u}{x} - (p-1)\right]}{\binom{u}{x} - s}$$

Then $P_{s,x}$ is a Mahler polynomial by Proposition 4.2. If $\binom{u}{x} \not\equiv s \pmod p$, then $P_{s,x}(u) \equiv 0 \pmod p$. If $\binom{u}{x} \equiv s \pmod p$, then by Al-Haytham's theorem,

$$P_{s,x}(u) \equiv -(p-1)! \equiv 1 \pmod p$$

It follows that $P_{s,x}(u) \equiv \chi_{s,x}(u) \pmod p$ in all cases. $\square$

We now prove Theorem 5.2. Since f is uniformly continuous, there exists a positive integer n such that if, for $0 \leqslant |x| \leqslant n$,

$$\binom{u}{x} \equiv \binom{v}{x} \pmod p$$

then

$$f(u) \equiv f(v) \pmod p$$

It follows that the value of $f(u)$ modulo p depends only on the residues modulo p of the family $\left\{\binom{u}{x}\right\}_{0 \leqslant |x| \leqslant n}$.

Let C be the set of all families $r = \{r_x\}_{0 \leqslant |x| \leqslant n}$ such that $0 \leqslant r_x < p$. For $0 \leqslant i < p$, let C_i be the set of all families r of C satisfying the following condition:

$$\text{if, for } 0 \leqslant |x| \leqslant n, \binom{u}{x} \equiv r_x \pmod p, \text{ then } f(u) \equiv i \pmod p \tag{5.12}$$

The sets $(C_i)_{0 \leqslant i < p}$ are pairwise disjoint and their union is C. We claim that, for all $u \in A^*$,

$$f(u) \equiv \sum_{0 \leqslant i < p} i P_i(u) \pmod p \tag{5.13}$$

where P_i is the Mahler polynomial

$$P_i = \sum_{r \in C_i} \prod_{0 \leqslant |x| \leqslant n} P_{r_x,x} \tag{5.14}$$

First consider, for $r \in C$, the characteristic function

$$\chi_r(u) = \prod_{0 \leqslant |x| \leqslant n} \chi_{r_x,x}(u)$$

By construction, χ_r is defined by

$$\chi_r(u) = \begin{cases} 1 & \text{if, for } 0 \leqslant |x| \leqslant n,\ \binom{u}{x} \equiv r_x \pmod p \\ 0 & \text{otherwise} \end{cases}$$

and it follows from (5.12) and from the definition of C_i that

$$f(u) \equiv \sum_{0 \leqslant i < p} \Bigl(i \sum_{r \in C_i} \chi_r(u) \Bigr) \pmod p \tag{5.15}$$

Now Lemma 5.3 gives immediately

$$\chi_r(u) \equiv \prod_{0 \leqslant |x| \leqslant n} P_{r_x, x}(u) \pmod p \tag{5.16}$$

and thus (5.13) follows now from (5.14), (5.15) and (5.16). The result follows, since

$$P = \sum_{0 \leqslant i < p} i P_i(u)$$

is a Mahler polynomial. □

Theorem 5.2 can be extended as follows.

Corollary 5.4 *Let $f : A^* \to \mathbb{N}$ be a uniformly continuous function. Then, for each positive integer r, there exists a Mahler polynomial P_r such that, for all $u \in A^*$, $f(u) \equiv P_r(u) \pmod{p^r}$.*

Proof. We prove the result by induction on r. For $r = 1$, the result follows from Theorem 5.2. If the result holds for r, there exists a Mahler polynomial P_r such that, for all $u \in A^*$, $f(u) - P_r(u) \equiv 0 \pmod{p^r}$. Let $g = f - P_r$. Since g is uniformly continuous, there exists a positive integer n such that if $\binom{u}{x} \equiv \binom{v}{x} \pmod p$ for $|x| \leqslant n$, then $g(u) \equiv g(v) \pmod{p^{2r}}$. It follows that $\frac{1}{p^r} g(u) \equiv \frac{1}{p^r} g(v) \pmod{p^r}$, and thus $\frac{1}{p^r} g$ is uniformly continuous.

Applying Theorem 5.2 to $\frac{1}{p^r} g$, we get a Mahler polynomial P such that, for all $u \in A^*$,

$$\frac{1}{p^r} g(u) \equiv P(u) \pmod p$$

Setting $P_{r+1} = P_r + p^r P$, we obtain finally

$$f(u) \equiv P_{r+1}(u) \pmod{p^{r+1}}$$

which concludes the proof. □

We now conclude the proof of Theorem 5.1. For each positive integer r, there exists a Mahler polynomial P_r such that, for all $u \in A^*$, $f(u) \equiv P_r(u) \pmod{p^r}$. Using (3.5) to compute explicitly the coefficients $\langle f - P_r, v\rangle$, we obtain

$$\langle f - P_r, v\rangle \equiv 0 \pmod{p^r}$$

Since P_r is a polynomial, there exists an integer n_r such that for all $v \in A^*$ such that $|v| \geqslant n$, $\langle P_r, v\rangle = 0$. It follows $|\langle f, v\rangle|_p < p^{-r}$ and thus $\lim_{|v| \to \infty} |\langle f, v\rangle|_p = 0$. □

Mahler's theorem is often presented as an interpolation result (see for instance [9, p. 57]). This can also be extended to functions from words to integers. Given a family of integers $(c_v)_{v \in A^*}$, one can ask whether there is a (uniformly) continuous function f from the free pro-p group to $\mathbb{Z}$ such that $f(v) = c_v$. Then answer is yes if and only if $\lim_{|v| \to \infty} |m_v|_p = 0$, where $m_v = \sum_{0 \leqslant |x| \leqslant |v|} (-1)^{|v|+|x|} \binom{v}{x} c_x$.

6 Conclusion

We proved an extension of Mahler's theorem for functions from words to integers. It would be interesting to find a suitable extension for functions from words to words. It would also be interesting to see whether other results from p-adic analysis can be extended to the word case.

Acknowledgement

The authors would like to thank Daniel Barsky and Gilles Christol for pointing out several references.

References

[1] Y. Amice, Interpolation p-adique, *Bull. Soc. Math. France* **92** (1964), 117–180.

[2] J. Berstel, L. Boasson, O. Carton, B. Petazzoni and J.-E. Pin, Operations preserving recognizable languages, *Theoret. Comput. Sci.* **354** (2006), 405–420.

[3] M. Droste and G.-Q. Zhang, On transformations of formal power series, *Inform. and Comput.* **184**,2 (2003), 369–383.

[4] S. Eilenberg, *Automata, Languages and Machines*, vol. B, Academic Press, New York, 1976.

[5] M. Hall, Jr., A topology for free groups and related groups, *Ann. of Math. (2)* **52** (1950), 127–139.

[6] S. R. Kosaraju, Regularity preserving functions, *SIGACT News* **6 (2)** (1974), 16–17. Correction **6 (3)** (1974), 22.

[7] M. Lothaire, *Combinatorics on words*, *Cambridge Mathematical Library*, Cambridge University Press, Cambridge, 1997.

[8] K. Mahler, An interpolation series for continuous functions of a p-adic variable., *J. Reine Angew. Math.* **199** (1958), 23–34. Correction **208** (1961), 70–72.

[9] M. R. Murty, *Introduction to p-adic analytic number theory*, *Studies in Advanced Mathematics*, American Mathematical Society, Cambridge, 2002.

[10] D. Perrin and J.-E. Pin, *Infinite Words*, *Pure and Applied Mathematics* vol. 141, Elsevier, 2004. ISBN 0-12-532111-2.

[11] J.-E. Pin, Topologie p-adique sur les mots, *Journal de théorie des nombres de Bordeaux* **5** (1993), 263–281.

[12] J.-E. Pin and J. Sakarovitch, Operations and transductions that preserve rationality, in *6th GI Conference*, Berlin, 1983, pp. 617–628, *Lect. Notes Comp. Sci.* n° 145, Lect. Notes Comp. Sci.

[13] J.-E. Pin and J. Sakarovitch, Une application de la représentation matricielle des transductions, *Theoret. Comput. Sci.* **35** (1985), 271–293.

[14] J.-E. Pin and P. V. Silva, A topological approach to transductions, *Theoret. Comput. Sci.* **340** (2005), 443–456.

[15] J.-E. Pin and P. Weil, Uniformities on free semigroups, *International Journal of Algebra and Computation* **9** (1999), 431–453.

[16] C. Reutenauer and M.-P. Schützenberger, Variétés et fonctions rationnelles, *Theoret. Comput. Sci.* **145**,1-2 (1995), 229–240.

[17] A. M. Robert, *A course in p-adic analysis*, *Graduate Texts in Mathematics* vol. 198, Springer-Verlag, New York, 2000.

[18] J. I. Seiferas and R. McNaughton, Regularity-preserving relations, *Theoret. Comp. Sci.* **2** (1976), 147–154.

[19] V. L. Selivanov and K. W. Wagner, A reducibility for the dot-depth hierarchy, *Theoret. Comput. Sci.* **345**,2-3 (2005), 448–472.

[20] R. E. Stearns and J. Hartmanis, Regularity preserving modifications of regular expressions, *Information and Control* **6** (1963), 55–69.

[21] W. Wadge, *Reducibility and determinateness in the Baire space*, PhD thesis, University of California, Berkeley, 1983.

[22] K. Wagner, On ω-regular sets, *Inform. and Control* **43**,2 (1979), 123–177.