A noncommutative extension of Mahler’s theorem on interpolation series
Jean-Eric Pin, Pedro Silva

To cite this version:
Jean-Eric Pin, Pedro Silva. A noncommutative extension of Mahler’s theorem on interpolation series. 2011. hal-00654030

HAL Id: hal-00654030
https://hal.science/hal-00654030
Preprint submitted on 20 Dec 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
A noncommutative extension of Mahler’s theorem on interpolation series

Jean-Éric Pin, Pedro V. Silva

Tuesday 20th December, 2011

Abstract

In this paper, we prove an extension of Mahler’s theorem on interpolation series, 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 free monoid \( A^* \) to \( \mathbb{Z} \), where \( d_p \) is replaced by the pro-\( p \) metric, the profinite metric on \( A^* \) defined by \( p \)-groups.

The aim of this paper is to give a noncommutative version of Mahler’s theorem on interpolation series [8], which applies to functions from a free monoid to the set of integers. This result was first announced in [15] and this article is the full version of this conference paper.

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 real 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 interpolation theorem [8]. It is usually stated for \( p \)-adic valued functions on the \( p \)-adic integers, but this full version can be easily recovered from the simpler version given in

---

1Work supported by the project ANR 2010 BLAN 0202 02 FREC, the European Regional Development Fund through the programme COMPETE and the Portuguese Government through FCT (Fundação para a Ciência e a Tecnologia) under the project PEst-C/MAT/UI0144/2011.

2LIAFA, Université Paris 7 and CNRS, Case 7014, 75205 Paris Cedex 13, France. Jean-Eric.Pin@liafa.jussieu.fr

3Centro de Matemática, Faculdade de Ciências, Universidade do Porto, R. Campo Alegre 687, 4169-007 Porto, Portugal. pvsilva@fc.up.pt
Theorem 1 below. First, for each function \( f : \mathbb{N} \to \mathbb{Z} \), there exists a unique family \( a_k \) of integers such that, for all \( n \in \mathbb{N} \),

\[
f(n) = \sum_{k=0}^{\infty} a_k \binom{n}{k}
\]
or, in short notation, \( f = \sum_{k=0}^{\infty} a_k \binom{\cdot}{k} \). The coefficients \( a_k \) are given by the formula

\[
a_k = (\Delta^k f)(0)
\]

where \( \Delta \) is the difference operator, defined by \( (\Delta f)(n) = f(n+1) - f(n) \). In other words, the binomial functions \( \binom{\cdot}{k} \) form a basis of the \( \mathbb{Z} \)-module of all functions from \( \mathbb{N} \) to \( \mathbb{Z} \). The series \( \sum_{k=0}^{\infty} a_k \binom{\cdot}{k} \) is called the Mahler’s expansion of \( f \).

Mahler’s theorem states that \( f \) is uniformly continuous for the \( p \)-adic metric if and only if the partial sums of its Mahler’s expansion converge uniformly to \( f \). More precisely:

**Theorem 1 (Mahler)** Let \( \sum_{k=0}^{\infty} a_k \binom{\cdot}{k} \) be the Mahler’s expansion of a function \( f : \mathbb{N} \to \mathbb{Z} \). The following conditions are equivalent:

1. \( f \) is uniformly continuous for the \( p \)-adic metric,
2. the sequence of partial sums \( \sum_{k=0}^{n} a_k \binom{\cdot}{k} \) converges uniformly to \( f \),
3. \( \lim_{k \to \infty} |a_k|_p = 0 \).

The most remarkable part of this 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. This nice feature makes Mahler’s theorem the dream of students in mathematics: a function is equal to the sum of its Newton series if and only if it is uniformly continuous.

Our goal is to give a noncommutative extension of this result, based on the following observation: since the additive monoid \( \mathbb{N} \) is isomorphic to the free monoid on a one-letter alphabet, the free monoid \( A^* \) on a finite alphabet \( A \) can be viewed as a noncommutative generalization of the natural numbers. Accordingly, our noncommutative extension concerns functions from \( A^* \) to \( \mathbb{Z} \). Nevertheless, generalizing Mahler’s theorem to these functions requires a few preliminary questions to be solved:

1. Extend the difference operators to word functions,
2. Extend the binomial coefficients to words,
3. Find a metric on \( A^* \) which generalizes \( d_p \),
4. Find a Mahler expansion for functions from \( A^* \) to \( \mathbb{Z} \).

Fortunately, (1) can be readily solved and (2) and (3) have already be addressed by Eilenberg [5]: there is a well-defined notion of binomial coefficients \( \binom{u}{v} \) for two words \( u \) and \( v \) and the pro-\( p \) metric on \( A^* \) is a natural
extension of the $p$-adic metric. Question (4) is the topic of Section 2, which culminates with Theorem 2.2 just like in the commutative case, every function $f : A^* \to \mathbb{Z}$ has a Mahler’s expansion of the form $\sum_{v \in A^*} \langle f, v \rangle \left( -v \right)$ where the coefficients $\langle f, v \rangle$ are integers related to the difference operators. There is also a striking similarity between Mahler’s theorem and our main result, which can be stated as follows:

**Theorem 2** Let $\sum_{v \in A^*} \langle f, v \rangle \left( -v \right)$ be the Mahler’s expansion of a function $f : A^* \to \mathbb{Z}$. The following conditions are equivalent:

1. $f$ is uniformly continuous for the pro-$p$ metric,
2. the sequence of partial sums $\sum_{0 \leq |v| \leq n} \langle f, v \rangle \left( -v \right)$ converges uniformly to $f$,
3. $\lim_{|v| \to \infty} |\langle f, v \rangle|_p = 0$.

Although this result can be considered as interesting on its own right, the reader may wonder about our original motivation for proving this theorem. It was actually inspired by automata theoretic questions related to the classification of regularity-preserving functions in language theory (see [14, 16] for more details). We discuss further this original motivation in Section 7.

In Section 3, we introduce the notion of Mahler polynomials and prove some of their properties. Section 4, devoted to uniform continuity, is a crucial step towards the proof of the main theorem, presented in Section 5. In Section 6, we come back to the commutative case by providing an alternative proof for Amice’s Theorem, a generalization of Mahler’s Theorem to functions from $\mathbb{N}^k$ to $\mathbb{Z}$.

## 1 Background and notations

In this paper, $p$ denotes a fixed prime number.

### 1.1 The $p$-adic valuation and the $p$-adic metric

The $p$-adic valuation of a nonzero integer $n$ is the natural number $\nu_p(n) = \max \left\{ k \in \mathbb{N} \mid p^k \text{ divides } n \right\}$ and its $p$-adic norm is the real number $|n|_p = p^{-\nu_p(n)}$.

By convention, we set $\nu_p(0) = +\infty$ and $|0|_p = 0$. Note that, for all $x, y \in \mathbb{Z}$,

\begin{align*}
|x + y|_p &\leq \max\{|x|_p, |y|_p\} \leq |x|_p + |y|_p \quad (1.1) \\
|xy|_p &\leq |x|_p|y|_p \leq \min\{|x|_p, |y|_p\} \quad (1.2)
\end{align*}
The \( p \)-adic metric \( d_p \) on \( \mathbb{N} \) and \( \mathbb{Z} \) is defined by setting, for each pair \((x, y)\) of integers,
\[
d_p(x, y) = |x - y|_p
\]  
(1.3)
These definitions can be extended to \( \mathbb{N}^k \) and to \( \mathbb{Z}^k \) as follows. Given \( x = (x_1, \ldots, x_k) \) and \( y = (y_1, \ldots, y_k) \) in \( \mathbb{Z}^k \) and a positive integer \( r \), we write \( x \equiv y \pmod{r} \) if, for \( 1 \leq i \leq k \), \( x_i \equiv y_i \pmod{r} \). The \( p \)-adic valuation of \( x \) is
\[
\nu_p(x) = \min \{ n \in \mathbb{N} \mid x \not\equiv 0 \pmod{p^{n+1}} \}
\]
and the \( p \)-adic norm \( |x|_p = p^{-\nu_p(x)} \). Finally, the \( p \)-adic metric on \( \mathbb{Z}^k \) is defined by the formula (1.3), for each pair \((x, y)\) of elements of \( \mathbb{Z}^k \):
\[
d_p(u, v) \leq \max\{d_p(u, w), d_p(w, v)\}
\]

1.2 Words and languages
Let \( A \) be a finite alphabet. Recall that the free monoid \( A^* \) on \( A \) is the set of all words on \( A \), endowed with the concatenation operation. The empty word, denoted \( 1 \), is the identity of \( A^* \). The length of a word \( u \) is denoted \( |u| \). If \( a \) is a letter of \( A \), we denote by \( |u|_a \) the number of occurrences of \( a \) in \( u \).

A language is a subset of \( A^* \). An important class of languages is the class of recognizable (also called regular) languages. Recognizable languages admit various characterizations, ranging from automata theory to topology or logic, but we will settle for the following algebraic definition: a language \( L \) of \( A^* \) is recognizable if and only if there exists a monoid morphism \( \varphi \) from \( A^* \) onto some finite monoid \( M \) such that \( L = \varphi^{-1}(\varphi(L)) \). It is then said that \( M \) recognizes the language \( L \).

1.3 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 \( v = v_0 a_1 v_1 \cdots a_n v_n \). For instance \( aaba \) is a subword of \( caacabca \). Following \([5,7]\), we define the binomial coefficient of \( u \) and \( v \) by setting
\[
\binom{v}{u} = |\{ (v_0, \ldots, v_n) \mid v = v_0 a_1 v_1 \cdots a_n v_n \}|.
\]
Observe that if \( a \) is a letter, then \( \binom{v}{u} \) is equal to \( |v|_a \). Further, if \( u = a^n \) 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 1.1 Let $u, v \in A^*$ and $a, b \in A$. Then

1. $\binom{u}{1} = 1$,
2. $\binom{u}{v} = 0$ if $|u| \leq |v|$ and $u \neq v$,
3. $\binom{ua}{vb} = \begin{cases} \binom{u}{v} & \text{if } a \neq b \\ \binom{u}{v} + \binom{u}{b} & \text{if } a = b \end{cases}$

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

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

which leads to the formula

$$\binom{u_1u_2}{x} = \sum_{x_1x_2=x} \binom{u_1}{x_1} \binom{u_2}{x_2}$$

1.4 The pro-$p$ metric

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 [5]. The pro-$p$ metric $d_p$ on $A^*$ is defined by setting

$$r_p(u, v) = \min \{ n \mid \text{there is a } p\text{-group of order } p^{n+1} \text{ separating } u \text{ and } v \}$$

$$d_p(u, v) = p^{-r_p(u, v)}$$

with the usual convention $\min \emptyset = +\infty$ and $p^{-\infty} = 0$. It turns out that $d_p$ is an ultrametric. 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. This monoid is in fact a compact group, called the free pro-$p$ group and denoted $\hat{F}_p(A)$. When $A$ is a one-letter alphabet, this group is of course $\mathbb{Z}_p$, the additive group of $p$-adic integers.

One can define in a similar way a pro-$p$ metric on $\mathbb{N}^k$ or $\mathbb{Z}^k$. The attentive reader will notice a potential conflict of notation between the $p$-adic metric and the pro-$p$ metric on $\mathbb{Z}^k$, which were both denoted by $d_p$. The next lemma gets rid of this ambiguity.
Lemma 1.2 The \( p \)-adic metric and the pro-\( p \) metric coincide on \( \mathbb{N}^k \) and on \( \mathbb{Z}^k \).

Proof. First, the pro-\( p \) metric on \( \mathbb{N}^k \) is the restriction to \( \mathbb{N}^k \) of the pro-\( p \) metric on \( \mathbb{Z}^k \) since every morphism from \( \mathbb{N}^k \) onto a group extends uniquely to \( \mathbb{Z}^k \). Therefore it suffices to prove that for all \( u, v \in \mathbb{Z}^k \), one has \( r_p(u, v) = \nu_p(u - v) \).

Since \( \mathbb{Z}^k \) is Abelian, a direct product of groups separates \( u \) and \( v \) if and only if one of its factors separates \( u \) and \( v \). Let \( C_r \) be the cyclic group of order \( r \). In view of the structure theorem of Abelian groups, one has

\[
r_p(u, v) = \min \{ n \mid C_{p^{n+1}} \text{ separates } u \text{ and } v \} \leq \nu_p(u - v)
\]

Further, for every homomorphism \( f : \mathbb{Z}^k \to \mathbb{Z} \), \( u \equiv v \pmod{p^{n+1}} \) implies \( f(u) \equiv f(v) \pmod{p^{n+1}} \). Therefore \( u \not\equiv v \pmod{p^{n+1}} \) is a necessary condition for \( C_{p^{n+1}} \) to separate \( u \) and \( v \). Thus \( r_p(u, v) = \nu_p(u - v) \) as claimed.

There is a nice connection \([11]\) between the pro-\( p \) metric and the binomial coefficients, which comes from the characterization of the languages recognized by a \( p \)-group given by Eilenberg and Schützenberger \([5, \text{Theorem 10.1, p. 239}]\). Let us call a \( p \)-group language a language recognized by a \( p \)-group. Note that such a language is recognizable by definition.

Proposition 1.3 A language of \( A^* \) is a \( p \)-group language if and only if it is a Boolean combination of the languages

\[
L(x, r, p) = \left\{ u \in A^* \mid \left( \begin{array}{c} u \\ x \end{array} \right) \equiv r \pmod{p} \right\},
\]

for \( 0 \leq r < p \) and \( x \in A^* \).

Let us set now

\[
r'_p(u, v) = \min \left\{ |x| \mid x \in A^* \text{ and } \left( \begin{array}{c} u \\ x \end{array} \right) \neq \left( \begin{array}{c} v \\ x \end{array} \right) \pmod{p} \right\}
\]

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

It is proved in \([11]\) Theorem 4.4] that \( d'_p \) is an ultrametric uniformly equivalent to \( d_p \). We shall use this fact in the proof of our main theorem.

The following result, which follows from \([15]\) Theorem 2.3] provides a characterization of uniform continuity in terms of \( p \)-group languages. Let \( A \) and \( B \) be two finite alphabets.

Theorem 1.4 A function \( f : A^* \to B^* \) is uniformly continuous for \( d_p \) if and only if, for every \( p \)-group language \( L \) of \( B^* \), the language \( f^{-1}(L) \) is a \( p \)-group language of \( A^* \).
Since $p$-group languages are closed under inverse of morphisms [5], one gets the following corollary.

**Corollary 1.5** Every monoid morphism from $A^*$ to $B^*$ is uniformly continuous for $d_p$.

As another consequence of Theorem 1.4, we obtain:

**Corollary 1.6** For every $v \in A^*$, the function from $A^*$ into $\mathbb{N}$ which maps every word $x$ onto the binomial coefficient $\binom{x}{v}$ is uniformly continuous for $d_p$.

**Proof.** Let $f_v : A^* \to \mathbb{N}$ be the function defined by $f_v(x) = \binom{x}{v}$. Since the only $p$-group quotients of $\mathbb{N}$ are the cyclic groups of order $p^n$, a language of $\mathbb{N}$ recognized by a $p$-group is a finite union of arithmetic progressions $r+p^n\mathbb{N}$ ($r \in \{0, \ldots, p^n-1\}$). In view of Theorem 1.4, and since the class of $p$-group languages is closed under union, we only need to show that $f_{v^{-1}}(r+p^n\mathbb{N})$ is a $p$-group language. Now, [10, Example 4, p. 450] shows that $f_{v^{-1}}(r+p^n\mathbb{N}) = \{ x \in A^* | \binom{x}{v} \equiv r \pmod{p^n} \}$ is indeed recognized by a $p$-group. \(\Box\)

## 2 Mahler’s expansions

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 $\Delta^a$ by

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

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

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

For instance,

$$\Delta^{aab} f(u) = -f(u) + 2f(ua) + f(ub) - f(uab) - 2f(uab) + f(uaab)$$

The next proposition gives a direct way to compute the difference operators.

**Proposition 2.1** For each word $w \in A^*$, the following relation holds:

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

(2.1)
Proof. We prove the result by induction on the length of $w$. If $w$ is the empty word, the result is trivial. Let $w$ be a word and let $a$ be a letter. Then $(\Delta^w f)(u) = (\Delta^w f)(ua) - (\Delta^w f)(u)$ and by the induction hypothesis applied to $w$,

$$(\Delta^w f)(u) = \sum_{0 \leq |x| \leq |w|} (-1)^{|w|+|x|} \binom{w}{x} (f(uax) - f(u{x}))$$

On the other hand, since

$$(awx) = \begin{cases} (wxaz) + (wz) & \text{if } x = az \\ (wx) & \text{otherwise} \end{cases}$$

we get

$$\sum_{0 \leq |x| \leq |aw|} (-1)^{|aw|+|x|} \binom{aw}{x} f(u{x})$$

$$= \sum_{0 \leq |x| \leq |w|+1} (-1)^{|w|+|az|+1} \binom{w}{az} + \binom{w}{x} f(uaz)$$

$$+ \sum_{0 \leq |x| \leq |w|+1} (-1)^{|w|+|x|+1} \binom{w}{x} f(u{x})$$

$$= \sum_{0 \leq |x| \leq |w|} (-1)^{|w|+|x|} \binom{w}{x} f(uaz)$$

$$+ \sum_{0 \leq |x| \leq |w|} (-1)^{|w|+|x|+1} \binom{w}{x} f(u{x})$$

$$= \sum_{0 \leq |x| \leq |w|} (-1)^{|w|+|x|} \binom{w}{x} (f(uax) - f(u{x}))$$

and thus

$$(\Delta^w f)(u) = \sum_{0 \leq |x| \leq |aw|} (-1)^{|aw|+|x|} \binom{aw}{x} f(u{x})$$

which gives the induction step for the proof of (2.1). □

By Corollary 1.6, for a fixed $v \in A^*$, the function $\binom{w}{x}$ from $A^*$ to $\mathbb{N}$ 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 $f$. 8
Theorem 2.2 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 \( u \in A^* \),

\[
f(u) = \sum_{v \in A^*} \langle f, v \rangle \binom{u}{v}
\]

This family is given by

\[
\langle f, v \rangle = (\Delta^v f)(1) = \sum_{0 \leq |x| \leq |v|} (-1)^{|v|+|x|} \binom{v}{x} f(x)
\]

Proof. First observe that, according to (2.1)

\[
(\Delta^v f)(1) = \sum_{0 \leq |x| \leq |v|} (-1)^{|v|+|x|} \binom{v}{x} f(x)
\]

Thus

\[
\sum_{v \in A^*} (\Delta^v f)(1) \binom{u}{v} = \sum_{v \in A^*} \sum_{|x| \leq |v|} (-1)^{|v|+|x|} \binom{v}{x} \binom{u}{v} f(x)
\]

\[
= \sum_{x \in A^*} (-1)^{|u|+|x|} \left( \sum_{0 \leq |v| \leq |u|} (-1)^{|v|+|u|} \binom{u}{v} \binom{v}{x} \right) f(x)
\]

\[
= f(u)
\]

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

\[
\sum_{0 \leq |u| \leq |v|} (-1)^{|u|+|v|} \binom{u}{v} \binom{v}{u} = \begin{cases} 1 & \text{if } u = w \\ 0 & \text{otherwise} \end{cases}
\]

(2.5)

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

\[
\langle f, u \rangle = f(u) - \sum_{0 \leq |v| < |u|} \langle f, v \rangle \binom{u}{v},
\]

a straightforward consequence of (2.2). \( \square \)

The series defined by (2.2) is called the Mahler’s expansion of \( f \).

Example 2.1 Let \( A \) be the binary alphabet \( \{0, 1\} \). The empty word will be exceptionally denoted by \( \varepsilon \) to avoid confusion. Let \( f : A^* \to \mathbb{N} \) be the function mapping a binary word onto its value as a binary number. For instance \( f(010111) = f(10111) = 23 \). Then one sees easily that

\[
(\Delta^v f) = \begin{cases} f + 1 & \text{if } v \in 1\{0, 1\}^* \\ f & \text{otherwise} \end{cases}
\]

\[
(\Delta^v f)(\varepsilon) = \begin{cases} 1 & \text{if } v \in 1\{0, 1\}^* \\ 0 & \text{otherwise} \end{cases}
\]
It follows that the Mahler expansion of $f$ is

$$f = \sum_{v \in \{0,1\}^*} \binom{-v}{v}$$

For instance, if $u = 01001$, one gets

$$f(u) = \left(\frac{u}{1}\right) + \left(\frac{u}{10}\right) + \left(\frac{u}{11}\right) + \left(\frac{u}{100}\right) + \left(\frac{u}{101}\right) + \left(\frac{u}{1001}\right) = 2 + 2 + 1 + 1 + 2 + 1 = 9.$$ 

The Mahler’s expansion of a function $f : A^* \to \mathbb{Z}$ can be used to compute its uniform norm $\|f\|_p$, defined by

$$\|f\|_p = \sup_{u \in A^*} |f(u)|_p$$

First observe that $\|f\|_p = \max_{u \in A^*} |f(u)|_p$ since $|f(u)|_p$ takes values in the set $\{p^{-n} \mid n \geq 0\} \cup \{0\}$.

**Theorem 2.3** Let $\sum_{v \in A^*} \langle f, v \rangle \binom{-v}{v}$ be the Mahler’s expansion of a function $f : A^* \to \mathbb{Z}$. Then $\|f\|_p = \max_{v \in A^*} |\langle f, v \rangle|_p$.

**Proof.** Let us choose a word $u$ of $A^*$ such that $\|f\|_p = |f(u)|_p$. Then $f(u) = \sum_{|v| \leq |u|} \langle f, v \rangle \binom{u}{v}$ and thus we get by (1.1) and (1.2)

$$\|f\|_p = |f(u)|_p \leq \max_{|v| \leq |u|} |\langle f, v \rangle|_p \leq \max_{v \in A^*} |\langle f, v \rangle|_p$$

Conversely, let $w$ be a word of minimal length such that $\max_{v \in A^*} |\langle f, v \rangle|_p$. The formula

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

yields, with the help of (1.1) and (1.2),

$$|\langle f, w \rangle|_p \leq \max(|f(w)|_p, \max_{0 \leq |v| < |w|} |\langle f, v \rangle|_p)$$

Now, for $0 \leq |v| < |w|$, one has $|\langle f, v \rangle|_p < |\langle f, w \rangle|_p$ by the minimality of $|w|$ and thus max$_{0 \leq |v| < |w|} |\langle f, v \rangle|_p < |\langle f, w \rangle|_p$. It follows that

$$|\langle f, w \rangle|_p \leq |f(w)|_p \leq \|f\|_p$$

and hence max$_{v \in A^*} |\langle f, v \rangle|_p = \|f\|_p$ as claimed. □
3 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 $\mathbb{Z}[\langle A \rangle]$ 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 $\mathbb{Z}[\langle A \rangle]$, denoted by $\uparrow$ and defined inductively as follows:

for all $u \in A^*$,

$$u \uparrow 1 = 1 \uparrow u = u,$$

(3.1)

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

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

(3.2)

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

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

(3.3)

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 $v_1, v_2 \in A^*$,

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

(3.4)

Formula (3.4) leads to an explicit computation of the Mahler’s expansion of the product of two functions.

**Proposition 3.1** 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:

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

11
Proof. Indeed, if \( f(u) = \sum_{v \in A^*} \langle f, v \rangle \left( \begin{array}{c} u \\ v \end{array} \right) \) and \( g(u) = \sum_{v \in A^*} \langle g, v \rangle \left( \begin{array}{c} u \\ v \end{array} \right) \), then
\[
fg(u) = \sum_{v_1, v_2 \in A^*} \langle f, v_1 \rangle \langle g, v_2 \rangle \left( \begin{array}{c} u \\ v_1 \\ v_2 \end{array} \right)
\]
and the result follows by Formula (3.4). □

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

**Proposition 3.2** 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 3.1 shows that Mahler polynomials are closed under product. □

### 4 Uniform continuity

In this section, we give a simple characterization of the uniformly continuous functions from \( A^* \) to \( \mathbb{Z} \). Uniform continuity always refers to the metric \( d_p \), but it can be replaced by \( d'_p \) whenever convenient in view of [11, Theorem 4.4].

**Theorem 4.1** Let \( \sum_{v \in A^*} \langle f, v \rangle \left( \begin{array}{c} -v \\ v \end{array} \right) \) be the Mahler’s expansion of a function \( f : A^* \to \mathbb{Z} \). Then \( f \) is uniformly continuous if and only if \( \lim_{|v| \to \infty} \langle f, v \rangle |_p = 0 \).

**Proof.** Suppose that \( \lim_{|v| \to \infty} \langle f, v \rangle |_p = 0 \) and let \( r \in \mathbb{N} \). Then there exists \( s \in \mathbb{N} \) such that, if \( |v| \geq s \), \( \nu_p((f, v)) \geq r \). Setting
\[
g = \sum_{|v| < s} \langle f, v \rangle \left( \begin{array}{c} -v \\ v \end{array} \right) \text{ and } h = \sum_{|v| \geq s} \langle f, v \rangle \left( \begin{array}{c} -v \\ v \end{array} \right)
\]
we get \( f = g + h \). Further \( p^r \) divides \( \langle f, v \rangle \) for \( |v| \geq s \). Since \( g \) is a Mahler polynomial, it is uniformly continuous by Corollary [11] and so there exists \( t \in \mathbb{N} \) such that \( d_p(u, u') \leq p^{-t} \) 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 4.2** 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 \( 0 \leq s < p \). Let \( \chi_{s,x} : A^* \rightarrow \mathbb{N} \) 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 4.3 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 \( P_{s,x}(u) = -\left[ \binom{u}{x} \right] \left[ \binom{u}{x} - 1 \right] \cdots \left[ \binom{u}{x} - (p-1) \right] \left( \binom{u}{x} - s \right) \)

Then \( P_{s,x} \) is a Mahler polynomial by Proposition 3.2. If \( \binom{v}{x} \not\equiv s \pmod{p} \), then \( P_{s,x}(u) \equiv 0 \pmod{p} \). If \( \binom{v}{x} \equiv s \pmod{p} \), then by Ibn Al-Haytham’s theorem (also known as Wilson’s theorem),

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

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

We now prove Theorem 4.2. Since \( f \) is uniformly continuous for the metric \( d'_p \), there exists a positive integer \( n \) such that if, for \( 0 \leq |x| \leq n \),

\[ \binom{u}{x} \equiv \binom{v}{x} \pmod{p} \]

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 \( \{ \binom{u}{x} \}_{0 \leq |x| \leq n} \).

Let \( R \) be the set of all families \( r = \{ r_x \}_{0 \leq |x| \leq n} \) such that \( 0 \leq r_x < p \). For \( 0 \leq i < p \), let \( R_i \) be the set of all families \( r \) of \( R \) satisfying the following condition:

if, for \( 0 \leq |x| \leq n \), \( \binom{u}{x} \equiv r_x \pmod{p} \), then \( f(u) \equiv i \pmod{p} \) \hspace{1cm} (4.1)

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

\[ f(u) \equiv \sum_{0 \leq i < p} iP_i(u) \pmod{p} \] \hspace{1cm} (4.2)
where $P_i$ is the Mahler polynomial

$$P_i = \sum_{r \in R_i} \prod_{0 \leq |x| \leq n} P_{r, x}$$

(4.3)

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

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

By construction, $\chi_r$ is defined by

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

and it follows from (4.1) and from the definition of $R_i$ that

$$f(u) \equiv \sum_{0 \leq i < p} \left( i \sum_{r \in R_i} \chi_r(u) \right) \pmod{p}$$

(4.4)

Now Lemma 4.3 gives immediately

$$\chi_r(u) \equiv \prod_{0 \leq |x| \leq n} P_{r, x}(u) \pmod{p}$$

(4.5)

and thus (4.2) follows now from (4.3), (4.4) and (4.5). The result follows, since

$$P = \sum_{0 \leq i < p} iP_i(u)$$

is a Mahler polynomial.

Theorem 4.2 can be extended as follows.

**Corollary 4.4** 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 4.2. 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 $\binom{v}{x} \equiv \binom{w}{x} \pmod{p^r}$ for $|x| \leq n$, 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 4.2 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 come back to the proof of Theorem 4.1. 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 (2.3) 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 $|v| \geq n_r$, $\langle P_r, v \rangle = 0$. It follows $|\langle f, v \rangle|_p \leq p^{-r}$ and thus $\lim_{|v| \to \infty} |\langle f, v \rangle|_p = 0$. □

5 An extension of Mahler’s theorem

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 $f(v) = c_v$. The answer to this question is part of our main result:

**Theorem 5.1** Let $\sum_{v \in A^*} \langle f, v \rangle (\underset{\text{-}}{v})$ be the Mahler’s expansion of a function $f : A^* \to \mathbb{Z}$. Then the following conditions are equivalent:

1. $f$ is uniformly continuous for $d_p$,
2. the sequence of partial sums $\sum_{0 \leq |v| \leq n} \langle f, v \rangle (\underset{\text{-}}{v})$ converges uniformly to $f$,
3. $\lim_{|v| \to \infty} |\langle f, v \rangle|_p = 0$,
4. $\lim_{|v| \to \infty} ||\Delta_v f||_p = 0$,
5. $f$ admits a continuous extension from $\hat{F}_p(A)$ to $\mathbb{Z}_p$.

**Proof.** (1) implies (3) follows from Theorem 4.1. (3) implies (2). Formulas (1.1) and (1.2) give for all $u \in A^*$ and for all $n \in \mathbb{N}$
\[ \left| \sum_{n<|v|} \langle f, v \rangle \binom{u}{v} \right|_p \leq \sup_{n<|v| \leq |u|} |\langle f, v \rangle|_p \leq \sup_{n<|v|} |(f, v)|_p \]
Let $\varepsilon > 0$. By (3), there is an integer $N$ such that if $|v| \geq N$, then $|\langle f, v \rangle|_p < \varepsilon$. It follows that for all $n \geq N$ and for all $u \in A^*$,
\[ \left| \sum_{n<|v|} \langle f, v \rangle \binom{u}{v} \right|_p \leq \sup_{n<|v|} |\langle f, v \rangle|_p < \varepsilon. \]
which proves (2), in view of (2.1).

(2) implies (1). Theorem 4.1 shows that every Mahler polynomial is uniformly continuous. Now, if a sequence of uniformly continuous functions converges uniformly, then its limit is uniformly continuous as well. Thus (2) implies that \( f \) is uniformly continuous.

(3) implies (4). A straightforward induction shows that \( \Delta^u(\Delta^v f) = \Delta^{uv} f \) holds for all \( u,v \in A^* \). Hence

\[
\langle \Delta^u f, u \rangle = \Delta^u(\Delta^v f)(1) = \Delta^{uv} f(1) = \langle f, uv \rangle
\]

and Theorem 2.3 yields

\[
\|\Delta^u f\|_p = \max_{u \in A^*} |\langle \Delta^v f, u \rangle|_p = \max_{u \in A^*} |\langle f, uv \rangle|_p
\]

Let \( \varepsilon > 0 \). By (3), there exists some \( N \in \mathbb{N} \) such that \( |\langle f, w \rangle|_p < \varepsilon \) whenever \( |w| \geq N \). Thus, whenever \( |v| \geq N \), we get \( |\langle f, uv \rangle|_p < \varepsilon \) and so \( \|\Delta^v f\|_p < \varepsilon \) by (5.1). Therefore (4) holds.

(4) implies (3) follows immediately from (5.1).

(1) implies (5) since \( \hat{F}_p(A) \) (respectively \( \hat{Z}_p \)) is the completion of \( A^* \) (respectively \( Z \)) for \( d_p \).

(5) implies (1). Since \( \hat{F}_p(A) \) is compact, the continuous extension of \( f \) to \( \hat{F}_p(A) \) is uniformly continuous and its restriction \( f \) to \( A^* \) is uniformly continuous as well. □

6 The commutative case

A Mahler expansion can also be obtained for functions from \( \mathbb{N}^k \) into \( Z \). First observe that the family of functions from \( \mathbb{N}^k \) into \( \mathbb{N} \)

\[
\left\{ \left( \begin{array}{c} r_1 \\ r_2 \\ \vdots \\ r_k \end{array} \right) \mid (r_1, \ldots, r_k) \in \mathbb{N}^k \right\}
\]

is locally finite. Thus, given a family \( (m_r)_{r \in \mathbb{N}^k} \), the formula

\[
f(n_1, \ldots, n_k) = \sum_{r \in \mathbb{N}^k} m_r \binom{n_1}{r_1} \cdots \binom{n_k}{r_k}
\]

defines a function \( f : \mathbb{N}^k \to \mathbb{Z} \). Conversely, each function from \( \mathbb{N}^k \) to \( \mathbb{Z} \) admits a unique Mahler expansion, a result proved in a more general setting in [2, 1].

**Proposition 6.1** Let \( f : \mathbb{N}^k \to \mathbb{Z} \) be an arbitrary function. Then there exists a unique family \( \langle f, r \rangle_{r \in \mathbb{N}^k} \) of integers such that, for all \( (n_1, \ldots, n_k) \in \mathbb{N}^k \),

\[
f(n_1, \ldots, n_k) = \sum_{r \in \mathbb{N}^k} \langle f, r \rangle \binom{n_1}{r_1} \cdots \binom{n_k}{r_k}
\]
Furthermore, the coefficients are given by

\[ \langle f, r \rangle = \sum_{i_1=0}^{r_1} \ldots \sum_{i_k=0}^{r_k} (-1)^{(r_1+\ldots+r_k)+(i_1+\ldots+i_k)} \prod_{l=1}^{r_l} \prod_{k=1}^{r_k} f(i_1, \ldots, i_k) \]

The next lemma relates commutative and noncommutative Mahler expansions. Given a finite alphabet \( A \), we denote by \( c \) the map from \( A^* \) to \( \mathbb{N}^{|A|} \) which maps every word \( u \) onto its commutative image \( c(u) = (|u|_a)_{a \in A} \).

**Lemma 6.2** Let \( \sum_{r \in \mathbb{N}^k} \langle f, r \rangle \left( -1 \right)^{r_1} \ldots \left( -1 \right)^{r_k} \) be the Mahler expansion of a function \( f \) from \( \mathbb{N}^k \) into \( \mathbb{Z} \). Let \( A = \{a_1, \ldots, a_k\} \) be a \( k \)-letter alphabet. Then \( \langle f, r \rangle = \langle f \circ c, a_1^{r_1} \cdots a_k^{r_k} \rangle \) for every \( r \in \mathbb{N}^k \).

**Proof.** One has

\[ f(n_1, \ldots, n_k) = f(c(a_1^{n_1} \cdots a_k^{n_k})) = \sum_{v \in A^*} \langle f \circ c, v \rangle \left( a_1^{n_1} \cdots a_k^{n_k} \right)_v = \sum_{r_1=0}^{n_1} \ldots \sum_{r_k=0}^{n_k} \langle f \circ c, a_1^{r_1} \cdots a_k^{r_k} \rangle \left( n_1 \right)^{r_1} \ldots \left( n_k \right)^{r_k} \]

The uniqueness of the Mahler expansion stated in Proposition 6.1 yields the equality \( \langle f, r \rangle = \langle f \circ c, a_1^{r_1} \cdots a_k^{r_k} \rangle \) for every \( r \in \mathbb{N}^k \). \( \square \)

We can now deduce a new proof for the hard implication of the generalization of Mahler’s Theorem to \( \mathbb{N}^k \) (see [2] or [18, Section 4.2.4]).

Given \( r \in \mathbb{N}^k \), let us write \( |r| \) for \( |r_1| + \ldots + |r_k| \).

**Corollary 6.3 (Amice [2])** Let \( \sum_{r \in \mathbb{N}^k} \langle f, r \rangle \left( -1 \right)^{r_1} \ldots \left( -1 \right)^{r_k} \) be the Mahler expansion of a function \( f \) from \( \mathbb{N}^k \) into \( \mathbb{Z} \). Then \( f \) is uniformly continuous for \( d_p \) if and only if \( \lim_{|r| \to \infty} |\langle f, r \rangle|_p = 0 \).

**Proof.** Assume that \( f \) is uniformly continuous. Let \( A = \{a_1, \ldots, a_k\} \). Since \( c : A^* \to \mathbb{N}^k \) is a morphism, it is uniformly continuous by Corollary 1.5 and so is the function \( f \circ c \). By Theorem 4.1, we obtain

\[ \lim_{|v| \to \infty} |\langle f \circ c, v \rangle|_p = 0. \]

In particular, \( \lim_{|r| \to \infty} |\langle f \circ c, a_1^{r_1} \cdots a_k^{r_k} \rangle|_p = 0 \) and so \( \lim_{|r| \to \infty} |\langle f, r \rangle|_p = 0 \) by Lemma 6.2.

The opposite implication is the easy one (see [2]) and can be deduced from the fact that if \( n_1 \equiv n_2 \pmod{p^{s+\nu_p(m)}} \) then \( \left( \frac{n_1}{m} \right) \equiv \left( \frac{n_2}{m} \right) \pmod{p^s} \). Details are omitted. \( \square \)
7 Back to language theory

Our original motivation was the study of regularity-preserving functions $f$ from $A^*$ to $B^*$, in the following sense: if $L$ is a regular language of $B^*$, then $f^{-1}(L)$ is a regular language of $A^*$. More generally, we are interested in functions preserving a given variety of languages $\mathcal{V}$: if $L$ is a language of $\mathcal{V}$, then $f^{-1}(L)$ is also a language of $\mathcal{V}$. There is an important literature on the regular case [20, 6, 13, 12, 13, 14, 3]. A remarkable contribution to this problem can be found in [17], where a characterization of sequential functions preserving star-free languages (respectively group-languages) is given. A similar problem was also recently considered for formal power series [4].

An open problem is to characterize the functions from $A^*$ to $B^*$ preserving the $p$-group languages. Theorems 1.4 and 5.1 give a solution when $B$ is a one-letter alphabet.

**Theorem 7.1** Let $f$ be a function from $A^*$ to $\mathbb{N}$. Then the following conditions are equivalent:

1. for each $p$-group language $L$ of $\mathbb{N}$, $f^{-1}(L)$ is a $p$-group language of $A^*$,
2. $f$ is uniformly continuous for $d_p$,
3. $\lim_{|v| \to \infty} \|\Delta^r f\|_p = 0$.

It would be interesting to find a suitable extension of this result for functions from $A^*$ to $B^*$, for any alphabet $B$. Another challenge would be to extend other results from $p$-adic analysis to the word case.

**Acknowledgement**

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

**References**


