MAHLER EXPANSIONS

KEITH CONRAD

1. Introduction

The set $C([0,1],\mathbf{R})$ of all continuous functions $f\colon [0,1]\to\mathbf{R}$ is a basic example of a function space in real analysis, and its p-adic analogue is the set $C(\mathbf{Z}_p,\mathbf{Q}_p)$ of all continuous functions $f\colon \mathbf{Z}_p\to\mathbf{Q}_p$. These are analogues both because $[0,1]$ and $\mathbf{Z}_p$ are compact and because $\mathbf{R}$ and $\mathbf{Q}_p$ are both complete. Examples of continuous functions are polynomials on $[0,1]$ with real coefficients and polynomials on $\mathbf{Z}_p$ with p-adic coefficients.

A metric can be put on $C([0,1],\mathbf{R})$ and $C(\mathbf{Z}_p,\mathbf{Q}_p)$ by measuring the distance between two functions as the largest distance between their values:

$$d(f,g)=\begin{cases}\max\limits_{x\in[0,1]}|f(x)-g(x)| & \text{if } f,g\colon[0,1]\to\mathbf{R},\\ \max\limits_{x\in\mathbf{Z}_p}|f(x)-g(x)|_p & \text{if } f,g\colon\mathbf{Z}_p\to\mathbf{Q}_p.\end{cases} \tag{1.1}$$

The maximum really occurs because $|f(x)-g(x)|$ in the real case and $|f(x)-g(x)|_p$ in the p-adic case are continuous functions $[0,1]\to\mathbf{R}$ and $\mathbf{Z}_p\to\mathbf{R}$, and any real-valued continuous function on a compact metric space has a maximum value (Extreme Value Theorem).

Theorem 1.1. *The function d in (1.1) is a metric and the spaces $C([0,1],\mathbf{R})$ and $C(\mathbf{Z}_p,\mathbf{Q}_p)$ are complete with respect to it.*

Proof. See the appendix. □

In both $C([0,1],\mathbf{R})$ and $C(\mathbf{Z}_p,\mathbf{Q}_p)$, the polynomial functions are a dense subset:

(1) if $f\colon[0,1]\to\mathbf{R}$ is continuous and $\varepsilon>0$ there is a polynomial $p(x)$ with real coefficients such that $|f(x)-p(x)|<\varepsilon$ for all $x\in[0,1]$.
(2) if $f\colon\mathbf{Z}_p\to\mathbf{Q}_p$ is continuous and $\varepsilon>0$ there is a polynomial $p(x)$ with coefficients in $\mathbf{Q}_p$ such that $|f(x)-p(x)|<\varepsilon$ for all $x\in\mathbf{Z}_p$.

In the real case, the denseness of polynomials in $C([0,1],\mathbf{R})$ is due to Weierstrass[1] (1885). The denseness of polynomials in $C(\mathbf{Z}_p,\mathbf{Q}_p)$ was first proved by Dieudonne (1944). While there is no nice description of all the functions in $C([0,1],\mathbf{R})$ (they are *not* describable as power series, since power series are infinitely differentiable and a continuous function need not be even once differentiable everywhere), in 1958 Mahler [3] gave a very nice description of all functions in $C(\mathbf{Z}_p,\mathbf{Q}_p)$ using infinite series of special polynomials.

Theorem 1.2 (Mahler)**.** *Every continuous function $f\colon\mathbf{Z}_p\to\mathbf{Q}_p$ can be written in the form*

$$f(x)=\sum_{n\geq 0}a_n\binom{x}{n}=a_0+a_1x+a_2\binom{x}{2}+a_3\binom{x}{3}+\cdots \tag{1.2}$$

[1]This was later generalized by Stone to a result now called the Stone–Weierstrass theorem, describing conditions under which a suitable collection of continuous real or complex-valued functions on a compact space is dense in the set of all continuous functions on that space.

for all $x \in \mathbf{Z}_p$, where $a_n \in \mathbf{Q}_p$ and $a_n \to 0$ as $n \to \infty$.

We'll see there is a formula for the coefficients a_n in terms of values of f that is reminiscent of Taylor's formula for the coefficients of a power series, but it must be stressed that (1.2) represents an arbitrary continuous function, not a function expressible by a power series. Just because each $\binom{x}{n}$ is differentiable (even infinitely differentiable) does not mean the series in (1.2) has that property too. This is similar to the "paradox" of Fourier series, where continuous functions $[0,1] \to \mathbf{R}$ have a representation as an infinite series of sines and cosines, but need not be differentiable even though $\sin(nx)$ and $\cos(nx)$ are differentiable.

The expansion (1.2) is called the *Mahler expansion* of f and the numbers a_n are called the *Mahler coefficients* of f. If $f_N(x)$ is the polynomial $\sum_{n=0}^{N} a_n\binom{x}{n}$ then $|f(x) - f_N(x)|_p = |\sum_{n\geq N+1} a_n\binom{x}{n}|_p \leq \max_{n\geq N+1} |a_n|_p$ since binomial coefficients at p-adic integers are p-adic integers. This maximum goes to 0 as $N \to \infty$ since $a_n \to 0$, so Mahler's theorem is an explicit (and useful!) way of approximating continuous functions $\mathbf{Z}_p \to \mathbf{Q}_p$ by polynomials.

2. Continuity and coefficient formula

Before we show every continuous function $\mathbf{Z}_p \to \mathbf{Q}_p$ has a Mahler expansion, let's see why an infinite series $\sum_{n\geq 0} a_n\binom{x}{n}$ with $a_n \to 0$ in $\mathbf{Q}_p$ is a continuous function on $\mathbf{Z}_p$.

Step 1: When $a_n \to 0$ in $\mathbf{Q}_p$, the infinite series $\sum_{n\geq 0} a_n\binom{x}{n}$ converges for all $x \in \mathbf{Z}_p$.

The key point is that even though $\binom{x}{n} = x(x-1)\cdots(x-(n-1))/n!$ has $n!$ in the denominator, which would ordinarily be bad p-adically (because $n!$ is very small p-adically as n grows), it is not bad when $x \in \mathbf{Z}_p$ because $\binom{x}{n} \in \mathbf{Z}_p$ by p-adic continuity of polynomials and its values on the dense subset $\mathbf{N}$. Thus $|\binom{x}{n}|_p \leq 1$ for $x \in \mathbf{Z}_p$, so $|a_n\binom{x}{n}|_p \leq |a_n|_p$, which proves $|a_n\binom{x}{n}|_p \to 0$ from $|a_n|_p \to 0$. Therefore $\sum_{n\geq 0} a_n\binom{x}{n}$ converges in $\mathbf{Q}_p$ for each $x \in \mathbf{Z}_p$.

Step 2: When $a_n \to 0$ in $\mathbf{Q}_p$, the function $f\colon \mathbf{Z}_p \to \mathbf{Q}_p$ defined by $f(x) = \sum_{n\geq 0} a_n\binom{x}{n}$ is continuous.

For $x_0 \in \mathbf{Z}_p$ we want to prove f is continuous at x_0. Pick $\varepsilon > 0$. Since $a_n \to 0$, there is an N such that $|a_n|_p < \varepsilon$ for $n \geq N$. Each of the finitely many functions $\binom{x}{n}$ for $0 \leq n \leq N-1$ is continuous at x_0, so by picking the minimal δ used for each of them in the ε-δ definition of continuity at x_0, there is a single $\delta > 0$ such that

$$|x - x_0|_p < \delta \Longrightarrow \left|\binom{x}{n} - \binom{x_0}{n}\right|_p < \varepsilon \tag{2.1}$$

for $n \in \{0, 1, \ldots, N-1\}$. If $|x - x_0|_p < \delta$,

$$|f(x) - f(x_0)|_p = \left|\sum_{n\geq 0} a_n\left(\binom{x}{n} - \binom{x_0}{n}\right)\right|_p \leq \max_{n\geq 0} |a_n|_p \left|\binom{x}{n} - \binom{x_0}{n}\right|_p.$$

Since $\binom{x}{n} - \binom{x_0}{n} \in \mathbf{Z}_p$ we have $|a_n|_p|\binom{x}{n} - \binom{x_0}{n}|_p \leq |a_n|_p < \varepsilon$ for $n \geq N$. For the terms preceding the Nth term, we instead say $|a_n|_p\left|\binom{x}{n} - \binom{x_0}{n}\right|_p \leq |a_n|_p\varepsilon$ by (2.1). Let $A = \max_{n\geq 0} |a_n|_p$ (this exists since the a_n's tend to 0), so $|a_n|_p \leq A$ for all n. Thus

$$|x - x_0|_p < \delta \Longrightarrow |f(x) - f(x_0)|_p \leq \max(\varepsilon, A\varepsilon) = \max(1, A)\varepsilon,$$

so f is continuous at x_0. As x_0 was arbitrary, f is continuous on $\mathbf{Z}_p$.

Next we will derive a formula for the coefficients in $f(x) = \sum_{n\geq 0} a_n\binom{x}{n}$ if $a_n \to 0$ in $\mathbf{Q}_p$. It is easy to get the value of the constant term: setting $x = 0$ in (1.2) we obtain $a_0 = f(0)$ since all terms vanish in the Mahler expansion except the one at $n = 0$. Similarly, if $x = 1$ then all terms in (1.2) vanish except for the first two: $f(1) = a_0 + a_1$, so $a_1 = f(1) - a_0 = f(1) - f(0)$. Setting $x = 2$, all terms in (1.2) vanish except the first three: $f(2) = a_0 + 2a_1 + a_2$, so

$$a_2 = f(2) - a_0 - 2a_1 = f(2) - f(0) - 2(f(1) - f(0)) = f(2) - 2f(1) + f(0).$$

These formulas suggest we can write a_n in terms of $f(0), f(1), \ldots, f(n)$, and we will see this is true.

The basic mechanism behind a formula for a_n is the discrete difference operator Δ: for any function $f\colon \mathbf{Z}_p \to \mathbf{Q}_p$, define $\Delta f\colon \mathbf{Z}_p \to \mathbf{Q}_p$ by

$$(\Delta f)(x) = f(x+1) - f(x).$$

This is also a function $\mathbf{Z}_p \to \mathbf{Q}_p$, and it can be iterated to have functions $\Delta^n f$ for $n \geq 2$: $\Delta^2 f = \Delta(\Delta f)$, and more generally $\Delta^n f = \Delta(\Delta^{n-1} f)$. Set $\Delta^0 f = f$, which is analogous to the zeroth derivative $f^{(0)}$ of a function f being the function itself.

The discrete difference operator Δ behaves nicely on the binomial coefficient polynomials because it shifts them down by one: $\Delta\binom{x}{n} = \binom{x}{n-1}$ for $n \geq 1$, and $\Delta\binom{x}{0} = \Delta(1)$ is the zero function. Indeed, by the Pascal's triangle recursion for binomial coefficients, if $n \geq 1$ then

$$\Delta\binom{x}{n} = \binom{x+1}{n} - \binom{x}{n} = \left(\binom{x}{n-1} + \binom{x}{n}\right) - \binom{x}{n} = \binom{x}{n-1}.$$

This resembles the effect of differentiation on x^n: $(x^n)' = nx^{n-1}$ for $n \geq 1$, and $(x^n)' = 0$ for $n = 0$. (A more accurate analogy would be with differentiation on the functions $x^n/n!$, whose derivative is $x^{n-1}/(n-1)!$ for $n \geq 1$.) Applying Δ to a function $f\colon \mathbf{Z}_p \to \mathbf{Q}_p$ having a Mahler expansion $\sum_{n\geq 0} a_n\binom{x}{n}$ where $a_n \to 0$,

$$(\Delta f)(x) = \sum_{n\geq 0} a_n\binom{x+1}{n} - \sum_{n\geq 0} a_n\binom{x}{n} = \sum_{n\geq 0} a_n\left(\binom{x+1}{n} - \binom{x}{n}\right) = \sum_{n\geq 1} a_n\binom{x}{n-1},$$

where a_0 drops out and other coefficients shift down one position. Reindexing the series to start at $n = 0$,

$$\Delta\sum_{n\geq 0} a_n\binom{x}{n} = \sum_{n\geq 0} a_{n+1}\binom{x}{n} = a_1 + a_2x + a_3\binom{x}{2} + a_4\binom{x}{3} + \cdots.$$

The effect of applying Δ to a Mahler expansion m times is to shift coefficients m positions:

$$\Delta^m\sum_{n\geq 0} a_n\binom{x}{n} = \sum_{n\geq 0} a_{n+m}\binom{x}{n} = a_m + a_{m+1}x + a_{m+2}\binom{x}{2} + a_{m+3}\binom{x}{3} + \cdots.$$

Setting $x = 0$ leaves only the constant term a_m, so we have proved the following theorem.

Theorem 2.1. *If $a_n \to 0$ in $\mathbf{Q}_p$ and $f(x) = \sum_{n\geq 0} a_n\binom{x}{n}$ for $x \in \mathbf{Z}_p$, then in terms of the function f we have $a_n = (\Delta^n f)(0)$.*

This is reminiscent of the formula $f^{(n)}(0)/n!$ for the coefficient of x^n in a power series. (A better analogy is that $f^{(n)}(0)$ is the coefficient of $x^n/n!$ in a power series.)

Now we will work out a formula for $(\Delta^n f)(0)$ to verify the earlier guess that a_n can be written in terms of $f(0), f(1), \ldots, f(n)$. Rather than focusing on the function $\Delta^n f$ just at $x = 0$, it is easier to see what's going on by getting a formula for $(\Delta^n f)(x)$ for general x.

Let's first work out formulas for $(\Delta^2 f)(x)$ and $(\Delta^3 f)(x)$:

$$\begin{aligned}(\Delta^2 f)(x) &= (\Delta(\Delta f))(x)\\ &= (\Delta f)(x+1) - (\Delta f)(x)\\ &= (f((x+1)+1) - f(x+1)) - (f(x+1) - f(x))\\ &= f(x+2) - 2f(x+1) + f(x)\end{aligned}$$

$$\begin{aligned}(\Delta^3 f)(x) &= (\Delta(\Delta^2 f))(x)\\ &= (\Delta^2 f)(x+1) - (\Delta^2 f)(x)\\ &= (f(x+3) - 2f(x+2) + f(x+1)) - (f(x+2) - 2f(x+1) + f(x))\\ &= f(x+3) - 3f(x+2) + 3f(x+1) - f(x).\end{aligned}$$

These hold for *all* functions $f\colon \mathbf{Z}_p \to \mathbf{Q}_p$.[2] Notice the binomial coefficients and alternating signs.

Theorem 2.2. *Let $f\colon \mathbf{Z}_p \to \mathbf{Q}_p$ be any function. For $n \geq 0$ and $x \in \mathbf{Z}_p$,*

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

Proof. We will give two proofs. One is a proof by induction that resembles a proof of the binomial theorem by induction (using the recursive formula for binomial coefficients) and the other is a slick proof using the binomial theorem on operators.

Proof by induction on n: When $n = 0$ the formula says $(\Delta^0 f)(x) = f(x)$, and when $n = 1$ the formula says $(\Delta^1 f)(x) = f(x+1) - f(x)$. These are true by definition (Δ^1 means Δ). If the formula is true for some n and all x in $\mathbf{Z}_p$, then

$$\begin{aligned}(\Delta^{n+1} f)(x) &= (\Delta(\Delta^n f))(x)\\ &= (\Delta^n f)(x+1) - (\Delta^n f)(x)\\ &= \sum_{k=0}^{n} (-1)^{n-k}\binom{n}{k} f((x+1)+k) - \sum_{k=0}^{n} (-1)^{n-k}\binom{n}{k} f(x+k)\\ &= \sum_{k=0}^{n} (-1)^{n-k}\binom{n}{k} f(x+(k+1)) + \sum_{k=0}^{n} (-1)^{n-k+1}\binom{n}{k} f(x+k)\\ &= \sum_{k=1}^{n+1} (-1)^{n-(k-1)}\binom{n}{k-1} f(x+k) + \sum_{k=0}^{n} (-1)^{n-(k-1)}\binom{n}{k} f(x+k).\end{aligned}$$

The term in the first sum at $k = n+1$ is $f(x+n+1)$. The term in the second sum at $k = 0$ is $(-1)^{n+1} f(x)$. The remaining terms in both sums run from $k = 1$ to $k = n$, and together equal

$$\sum_{k=1}^{n} (-1)^{n-(k-1)} \left(\binom{n}{k-1} + \binom{n}{k}\right) f(x+k) = \sum_{k=1}^{n} (-1)^{n-(k-1)} \binom{n+1}{k} f(x+k).$$

[2]The operator Δ on functions was used by Isaac Newton in his work on the calculus of finite differences, hundreds of years before its role in p-adic analysis.

The terms $f(x+n+1)$ and $(-1)^{n+1}f(x)$ fit into this sum at $k=n+1$ and $k=0$, so

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

Proof using binomial theorem: Define the shift operator S on functions by $(Sf)(x) = f(x+1)$. Then $(\Delta f)(x) = f(x+1) - f(x) = (Sf)(x) - f(x)$, so $\Delta = S - I$, where I is the identity operator on functions ($If = f$). The operators Δ, S, and I are all linear (*e.g.*, $S(af+bg) = aS(f) + bS(g)$ for functions f and g and scalars a and b). A product of operators A and B on functions is defined by composition, just like for matrices: $(AB)(f) = A(Bf)$. Thus powers of an operator mean repeated composition: $A^2f = A(Af)$, $A^3f = A(A^2f) = A(A(Af))$, and so on. For linear operators composition commutes with addition (like for matrices), and since S and I commute we can compute $\Delta^n = (S-I)^n$ by the binomial theorem:

$$\Delta^n = (S-I)^n = \sum_{k=0}^{n}\binom{n}{k}S^k(-I)^{n-k}.$$

Applying this formula for Δ^n to a function,

$$\Delta^n f = \sum_{k=0}^{n}\binom{n}{k}S^k((-I)^{n-k}f) = \sum_{k=0}^{n}\binom{n}{k}S^k((-1)^{n-k}f) = \sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}S^kf.$$

What does S^kf mean? Since $(Sf)(x) := f(x+1)$ shifts the variable in the function by 1, repeating this k times shifts the variable by k: $(S^kf)(x) = f(x+k)$ for all $k \geq 0$. Therefore

$$(\Delta^n f)(x) = \sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}(S^kf)(x) = \sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}f(x+k).$$

□

Corollary 2.3. *If* $f(x) = \sum_{n\geq 0} a_n\binom{x}{n}$ *for* $x \in \mathbf{Z}_p$*, where* $a_n \to 0$ *in* $\mathbf{Q}_p$*, then* $a_n = (\Delta^n f)(0) = \sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}f(k)$.

Proof. We have $a_n = (\Delta^n f)(0)$ by Theorem 2.1. Set $x = 0$ in Theorem 2.2. □

This corollary shows a continuous function $f\colon \mathbf{Z}_p \to \mathbf{Q}_p$ can be written as a Mahler expansion in at most one way, with the coefficients determined by the values of f on the nonnegative integers. That the knowledge of f on $\mathbf{N}$ should be enough to figure out the coefficients in a potential Mahler expansion for f is not a surprise: $\mathbf{N}$ is a dense subset of $\mathbf{Z}_p$, so any continuous function on $\mathbf{Z}_p$ is determined by its values on $\mathbf{N}$. But that is only a determination *in principle*. What is a surprise is how concretely the values in $f(\mathbf{N})$ are used in the coefficient formula.

3. Proof of Mahler's theorem

Up to this point we have checked Mahler expansions with coefficients tending to 0 are continuous functions and we have seen how the coefficients in such an expansion can be written in terms of values of the function. But we have not yet proved Mahler's theorem (Theorem 1.2): continuous functions $\mathbf{Z}_p \to \mathbf{Q}_p$ can be represented by Mahler expansions. This gap will be rectified now. Since the ideas used in the proof will not occur elsewhere, the reader can skip the proof to see what comes next and return to this section later.

Proof. For a continuous function $f\colon \mathbf{Z}_p \to \mathbf{Q}_p$, set $a_n := (\Delta^n f)(0) = \sum_{k=0}^n (-1)^{n-k}\binom{n}{k} f(k)$. Our goal is two-fold:

(1) show $a_n \to 0$ as $n \to \infty$, where $a_n = (\Delta^n f)(0)$,
(2) show $f(x) = \sum_{n\geq 0} a_n \binom{x}{n}$ for all $x \in \mathbf{Z}_p$.

Let's check first that (2) follows from (1). In (2), the series $\sum_{n\geq 0} a_n\binom{x}{n}$ is continuous on $\mathbf{Z}_p$ by (1), so to prove the series equals f it suffices to check these two continuous functions are equal on the dense subset of nonnegative integers. For $m \in \mathbf{N}$,

$$\sum_{n\geq 0} a_n \binom{m}{n} = \sum_{n=0}^m \left(\sum_{k=0}^n (-1)^{n-k}\binom{n}{k} f(k)\right)\binom{m}{n} = \sum_{k=0}^m \left(\sum_{n=k}^m (-1)^{n-k}\binom{n}{k}\binom{m}{n}\right) f(k)$$

Rewrite $\binom{n}{k}\binom{m}{n}$ as $\frac{n!}{k!(n-k)!}\frac{m!}{n!(m-n)!} = \frac{m!}{k!(n-k)!(m-n)!} = \frac{m!}{k!(m-k)!}\frac{(m-k)!}{(n-k)!(m-n)!} = \binom{m}{k}\binom{m-k}{n-k}$. Thus

$$\begin{aligned}\sum_{n\geq 0} a_n\binom{m}{n} &= \sum_{k=0}^m \left(\sum_{n=k}^m (-1)^{n-k}\binom{m}{k}\binom{m-k}{n-k}\right) f(k)\\ &= \sum_{k=0}^m \left(\sum_{n=k}^m (-1)^{n-k}\binom{m-k}{n-k}\right)\binom{m}{k} f(k)\\ &= \sum_{k=0}^m \left(\sum_{n=0}^{m-k} (-1)^{n}\binom{m-k}{n}\right)\binom{m}{k} f(k).\end{aligned}$$

The inner sum over n is the binomial expansion of $(1-1)^{m-k}$, which is 0 for $k < m$ and is 1 for $k = m$, so $\sum_{n\geq 0} a_n\binom{m}{n} = \binom{m}{m} f(m) = f(m)$. That proves $\sum_{n\geq 0} a_n\binom{x}{n}$ and $f(x)$ agree when $x \in \mathbf{N}$, so they agree everywhere by continuity of both functions on $\mathbf{Z}_p$.

It remains to show that if $f\colon \mathbf{Z}_p \to \mathbf{Q}_p$ is continuous then $(\Delta^n f)(0) \to 0$. Our argument comes from [4, p. 156] and is one of the nicer proofs of this fact that I have seen.

To prove what we want we'll prove something stronger: not just $|(\Delta^n f)(0)|_p \to 0$, but $|(\Delta^n f)(x)|_p \to 0$ uniformly in x. That is, we will show $||\Delta^n f|| \to 0$ as $n \to \infty$ where $||\Delta^n f|| := \max_{x\in\mathbf{Z}_p} |(\Delta^n f)(x)|_p$. For each n and x,

$$\begin{aligned}|(\Delta^{n+1} f)(x)|_p &= |(\Delta(\Delta^n f))(x)|_p\\ &= |(\Delta^n f)(x+1) - (\Delta^n f)(x)|_p\\ &\leq \max(|(\Delta^n f)(x+1)|_p, |(\Delta^n f)(x)|_p)\\ &\leq ||\Delta^n f||,\end{aligned}$$

and taking a maximum on the left over all x in $\mathbf{Z}_p$ gives us $||\Delta^{n+1} f|| \leq ||\Delta^n f||$. By induction $||\Delta^m f|| \leq ||\Delta^n f||$ for $m \geq n$, so it suffices to prove $||\Delta^{n_i} f|| \to 0$ along a sequence $n_1 < n_2 < n_3 < \cdots$ tending to ∞. We will use the powers of p, *i.e.*, prove $||\Delta^{p^r} f|| \to 0$ as $r \to \infty$. (This is a typical p-adic idea: things get small using high powers of p.)

For $n \geq 1$ and $x \in \mathbf{Z}_p$,

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

since $\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k} = (1-1)^n = 0$. The $k = 0$ term is $(-1)^n(f(x) - f(x)) = 0$, so drop it:

$$(\Delta^n f)(x) = \sum_{k=1}^{n}(-1)^{n-k}\binom{n}{k}(f(x+k) - f(x)).$$

Setting $n = p^r$ for $r \geq 0$, we will show that in the sum

$$(\Delta^{p^r} f)(x) = \sum_{k=1}^{p^r}(-1)^{p^r-k}\binom{p^r}{k}(f(x+k) - f(x)) \tag{3.1}$$

each term is small when r is large, independently of k and x.

Claim: For $1 \leq k \leq p^r$, $\mathrm{ord}_p \binom{p^r}{k} = r - \mathrm{ord}_p(k)$.

This is based on an important "factoring" formula for binomial coefficients: for $k \geq 1$,

$$\binom{x}{k} = \frac{x}{k}\binom{x-1}{k-1}. \tag{3.2}$$

The verification of this identity is an algebraic calculation left to the reader. Setting $x = p^r$,

$$\binom{p^r}{k} = \frac{p^r}{k}\binom{p^r-1}{k-1}.$$

We will show the integer $\binom{p^r-1}{k-1}$ is not divisible by p, so $\mathrm{ord}_p\binom{p^r}{k} = \mathrm{ord}_p(p^r) - \mathrm{ord}_p(k) = r - \mathrm{ord}_p(k)$, and that is the claim.

Using the formula for the highest power of p in a factorial,

$$\begin{aligned}
\mathrm{ord}_p\binom{p^r-1}{k-1} &= \mathrm{ord}_p\left(\frac{(p^r-1)!}{(k-1)!(p^r-k)!}\right) \\
&= \frac{p^r-1-s_p(p^r-1)}{p-1} - \frac{k-1-s_p(k-1)}{p-1} - \frac{p^r-k-s_p(p^r-k)}{p-1} \\
&= \frac{s_p(k-1)+s_p(p^r-k)-s_p(p^r-1)}{p-1}.
\end{aligned}$$

We will show the numerator is 0. It is clear if $k = p^r$, so assume $1 \leq k \leq p^r - 1$. Write k in base p as $c_ip^i + \cdots + c_{r-1}p^{r-1}$ with $c_i \neq 0$ (so $i = \mathrm{ord}_p(k)$). Then

$$\begin{aligned}
k-1 &= (p-1) + \cdots + (p-1)p^{i-1} + (c_i-1)p^i + c_{i+1}p^{i+1} + \cdots + c_{r-1}p^{r-1} \\
p^r-k &= (p-c_i)p^i + (p-1-c_{i+1})p^{i+1} + \cdots + (p-1-c_{r-1})p^{r-1} \\
p^r-1 &= (p-1) + (p-1)p + \cdots + (p-1)p^{r-1}.
\end{aligned}$$

Thus

$$\begin{aligned}
s_p(k-1) + s_p(p^r-k) &= ((p-1)i + s_p(k) - 1) + (1 + (p-1)(r-i) - s_p(k)) \\
&= (p-1)r \\
&= s_p(p^r-1),
\end{aligned}$$

which completes the proof of the claim.

For each $\varepsilon > 0$ we want to show $|(\Delta^{p^r} f)(x)|_p < \varepsilon$ for all large r and all x in $\mathbf{Z}_p$. By (3.1),

$$|(\Delta^{p^r} f)(x)|_p \leq \max_{1\leq k\leq p^r}\left|\binom{p^r}{k}\right|_p |f(x+k) - f(x)|_p.$$

Each term in the maximum is bounded above: $|\binom{p^r}{k}|_p \leq 1$ and $|f(x+k)-f(x)|_p \leq ||f||$. We'll show for large r that one factor in each term is small, so each term is (bounded)(small) = small.

For $1 \leq k \leq p^r$, by the claim $|\binom{p^r}{k}|_p = 1/p^{r-\mathrm{ord}_p(k)} = 1/(p^r|k|_p)$, so $|\binom{p^r}{k}|_p|k|_p = 1/p^r$. Therefore $|\binom{p^r}{k}|_p \leq 1/\sqrt{p^r}$ or $|k|_p \leq 1/\sqrt{p^r}$. Choosing $\varepsilon > 0$, there is $\delta > 0$ such that $|x-y|_p < \delta$ in $\mathbf{Z}_p$ implies $|f(x)-f(y)|_p < \varepsilon$. Pick R so large that $1/\sqrt{p^R} < \min(\delta, \varepsilon)$. For $r \geq R$ and each k between 1 and p^r we have two options:

(1) If $|\binom{p^r}{k}|_p \leq 1/\sqrt{p^r}$ then $|\binom{p^r}{k}|_p|f(x+k)-f(x)|_p \leq (1/\sqrt{p^r})||f|| \leq \varepsilon||f||$.
(2) If $|k|_p \leq 1/\sqrt{p^r}$ then $|k|_p < \delta$, so for all $x \in \mathbf{Z}_p$ we have $|(x+k)-x|_p < \delta$. Thus $|f(x+k)-f(x)|_p < \varepsilon$, so $|\binom{p^r}{k}|_p|f(x+k)-f(x)|_p \leq |f(x+k)-f(x)|_p < \varepsilon$.

Thus for each $\varepsilon > 0$ there is $R > 0$ such that $r \geq R \Longrightarrow ||\Delta^{p^r} f|| \leq \varepsilon \max(||f||, 1)$. □

Remark 3.1. We have shown continuous functions $f\colon \mathbf{Z}_p \to \mathbf{Q}_p$ admit a series representation $\sum_{n\geq 0} a_n\binom{x}{n}$ where $a_n \to 0$ in $\mathbf{Q}_p$. Could a *discontinuous* function $\mathbf{Z}_p \to \mathbf{Q}_p$ admit such a series representation with a_n not tending to 0? No. The reason is that $\binom{-1}{n} = (-1)^n$ (check!), so if $f(x) = \sum_{n\geq 0} a_n\binom{x}{n}$ for all x in $\mathbf{Z}_p$, where $a_n \in \mathbf{Q}_p$, then $f(-1) = \sum_{n\geq 0} a_n(-1)^n$ so $|a_n|_p \to 0$ and therefore f has to be continuous.

This trick of evaluating f at -1 is essential, because if we remove -1 from the domain then a Mahler-like expansion can converge everywhere else with coefficients not tending to 0. For example, if $x \in \mathbf{Z}_p - \{-1\}$ then set $f(x) = \frac{1}{x+1}\sum_{r\geq 0} a_r\binom{x+1}{p^r} = \sum_{r\geq 0}\frac{a_r}{p^r}\binom{x}{p^r-1}$ where $a_r \to 0$ in $\mathbf{Q}_p$ with $a_0 \neq 0$ and $a_r/p^r \not\to 0$ in $\mathbf{Q}_p$ (*e.g.*, $a_r = p^{[r/2]}$). The series $\sum_{r\geq 0} a_r\binom{x+1}{p^r}$ is continuous on $\mathbf{Z}_p$, so f is continuous on $\mathbf{Z}_p - \{-1\}$ and $(x+1)f(x) \to a_0 \neq 0$ as $x \to -1$. Therefore f can't be extended continuously to $x = -1$, but it has a Mahler-type expansion valid everywhere else on $\mathbf{Z}_p$ using the second formula that has unbounded coefficients.

Mahler's theorem extends to a description of all continuous functions on $\mathbf{Z}_p$ taking values in fields other than $\mathbf{Q}_p$.

Theorem 3.2. *Let $(K, |\cdot|)$ be a p-adic field, i.e., a complete valued extension of $\mathbf{Q}_p$. If $a_n \to 0$ in K then the series $\sum_{n\geq 0} a_n\binom{x}{n}$ converges for all $x \in \mathbf{Z}_p$ and is continuous. Conversely, every continuous function $f\colon \mathbf{Z}_p \to K$ has a unique representation in the form $\sum_{n\geq 0} a_n\binom{x}{n}$ where $a_n \in K$ and $a_n \to 0$. Explicitly, $a_n = (\Delta^n f)(0) = \sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}f(k)$.*

Proof. The proof is identical to the arguments used when $K = \mathbf{Q}_p$. □

Example 3.3. If $|a-1|_p < 1$ in $\mathbf{Q}_p$ then the series $f(x) = \sum_{n\geq 0}(a-1)^n\binom{x}{n}$ is continuous on $\mathbf{Z}_p$ since $|(a-1)^n|_p \to 0$. At a positive integer m we have $\binom{m}{n} = 0$ for $n > m$, so

$$f(m) = \sum_{n=0}^{m}(a-1)^n\binom{n}{m} = (1+(a-1))^m = a^m.$$

Therefore we have p-adically interpolated the power sequence $\{a^m\}$ from $m \in \mathbf{N}$ to a continuous function on $\mathbf{Z}_p$ that is denoted a^x:

$$a^x = \sum_{n\geq 0}(a-1)^n\binom{x}{n}. \tag{3.3}$$

For example, in $\mathbf{Q}_2$ with $a = -1$ we have

$$\sum_{n\geq 0}(-2)^n\binom{x}{n} = (-1)^x = \begin{cases} 1, & \text{if } x \in 2\mathbf{Z}_2, \\ -1, & \text{if } x \in 1+2\mathbf{Z}_2. \end{cases}$$

It's interesting how "nontrivial" the Mahler expansion of this locally constant function on $\mathbf{Z}_2$ looks.

4. Properties of Mahler expansions

Theorem 4.1. *For a continuous function $f\colon \mathbf{Z}_p \to \mathbf{Q}_p$ with Mahler expansion $\sum_{n\geq 0} a_n\binom{x}{n}$,*

$$\max_{x\in\mathbf{Z}_p}|f(x)|_p = \max_{n\geq 0}|a_n|_p.$$

Proof. For each $x \in \mathbf{Z}_p$,

$$|f(x)|_p = \left|\sum_{n\geq 0} a_n\binom{x}{n}\right|_p \leq \max_{n\geq 0}\left|a_n\binom{x}{n}\right|_p \leq \max_{n\geq 0}|a_n|_p.$$

Therefore $\max_{x\in\mathbf{Z}_p}|f(x)|_p \leq \max_{n\geq 0}|a_n|_p$.

To get the reverse inequality we use the formula for each a_n as $(\Delta^n f)(0)$:

$$|a_n|_p = \left|\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}f(k)\right|_p \leq \max_{0\leq k\leq n}\left|\binom{n}{k}f(k)\right|_p \leq \max_{0\leq k\leq n}|f(k)|_p \leq \max_{x\in\mathbf{Z}_p}|f(x)|_p.$$

Thus $\max_{n\geq 0}|a_n|_p \leq \max_{x\in\mathbf{Z}_p}|f(x)|_p$. □

Corollary 4.2. *For continuous functions $f, g\colon \mathbf{Z}_p \to \mathbf{Q}_p$ with respective Mahler expansions $\sum_{n\geq 0} a_n\binom{x}{n}$ and $\sum_{n\geq 0} b_n\binom{x}{n}$, $\max_{x\in\mathbf{Z}_p}|f(x)-g(x)|_p = \max_{n\geq 0}|a_n - b_n|_p$.*

Proof. The Mahler expansion of $f - g$ is $\sum_{n\geq 0}(a_n - b_n)\binom{x}{n}$. □

This theorem and corollary are true for continuous functions from $\mathbf{Z}_p$ to any p-adic field, not just to $\mathbf{Q}_p$.

Theorem 4.3. *In $\mathbf{Q}_p$, if $|a-1|_p \leq 1/p$ for $p \neq 2$ and $|a-1|_2 \leq 1/4$ for $p = 2$ then $|a^x - 1|_p = |a-1|_p|x|_p$ for all $x \in \mathbf{Z}_p$.*

Proof. This is obvious if $a = 1$ or if $x = 0$, so let's assume $a \neq 1$ and $x \neq 0$.

From the Mahler expansion and (3.2),

$$a^x = 1 + (a-1)x + \sum_{n\geq 2}(a-1)^n\binom{x}{n} = 1 + (a-1)x + \sum_{n\geq 2}(a-1)^n\frac{x}{n}\binom{x-1}{n-1}.$$

Subtract 1 and we get

$$|a^x - 1|_p = \left|(a-1)x + \sum_{n\geq 2}\frac{(a-1)^n}{n}x\binom{x-1}{n-1}\right|_p.$$

We will show when $a \neq 1$ that $|(a-1)^n/n|_p < |a-1|_p$ for all $n \geq 2$. This inequality is equivalent to $|a-1|_p < |n|_p^{1/(n-1)}$ and the reader can check $(1/p)^{1/(p-1)} \leq |n|_p^{1/(n-1)}$ for all $n \geq 2$ (there is equality only when $n = p$, but that doesn't matter), so from $|a-1|_p \leq$

$1/p < (1/p)^{1/(p-1)}$ when $p \neq 2$ or from $|a-1|_2 \leq 1/4 < 1/2$ when $p = 2$ we have verified $|(a-1)^n/n|_p < |a-1|_p$ for all $n \geq 2$. Then

$$\left|\sum_{n\geq 2} \frac{(a-1)^n}{n} x\binom{x-1}{n-1}\right|_p \leq \max_{n\geq 2}\left|\frac{(a-1)^n}{n}\right|_p |x|_p < |a-1|_p|x|_p \quad \text{(we use } a \neq 1, x \neq 0 \text{ here)},$$

so by the strong triangle inequality $|a^x - 1|_p = |a-1|_p|x|_p$. □

More generally, for any p-adic field $(K, |\cdot|)$ and $a \in K$ with $|a-1| < 1$, the continuous function a^x for $x \in \mathbf{Z}_p$ defined by the Mahler expansion (3.3) equals a^m when $x = m \in \mathbf{N}$, so Mahler expansions prove the power sequence $\{a^m\}$ has a continuous p-adic interpolation. Theorem 4.3 is valid when $|a-1| < (1/p)^{1/(p-1)}$: under that condition, $|a^x - 1| = |a-1||x|_p$.

Theorem 4.4. *For a p-adic field K, a continuous function $\sum_{n\geq 0} a_n\binom{x}{n}$ from $\mathbf{Z}_p$ to K is p-adic analytic (representable by a power series on $\mathbf{Z}_p$) if and only if $a_n/n! \to 0$ in K.*

Proof. Suppose $a_n/n! \to 0$. For $x \in \mathbf{Z}_p$,

$$f(x) = \sum_{n\geq 0} a_n\binom{x}{n} = \sum_{n\geq 0}\frac{a_n}{n!}x(x-1)\cdots(x-(n-1)).$$

Write $x(x-1)\cdots(x-(n-1)) = \sum_{m\geq 0} c_{mn}x^m$ where $c_{mn} \in \mathbf{Z}$ for $m \leq n$ and $c_{mn} = 0$ when $m > n$. (At $n = 1$ this is the constant polynomial 1.) Then

$$f(x) = \sum_{n\geq 0}\frac{a_n}{n!}x(x-1)\cdots(x-(n-1)) = \sum_{n\geq 0}\sum_{m\geq 0}\frac{a_n}{n!}c_{mn}x^m. \tag{4.1}$$

We want to switch the order of summation here. A condition that justifies this is having the (m,n)-th term tend to 0 as $\max(m,n) \to \infty$.

- If $m > n$ then the (m,n)-th term is 0 since $c_{mn} = 0$.
- If $n \geq m$, then from $c_{mn} \in \mathbf{Z}$ and $x \in \mathbf{Z}_p$ we can bound the (m,n)-th term:
$$\left|\frac{a_n}{n!}c_{mn}x^m\right| \leq \left|\frac{a_n}{n!}\right|,$$
and $|a_n/n!|$ can be made arbitrarily small for large enough n (independently of the choice of m when $m \leq n$).

Therefore the (m,n)-th term in the double sum in (4.1) tends to 0 as $\max(m,n) \to \infty$, so we can switch the order of summation and get

$$f(x) = \sum_{m\geq 0}\sum_{n\geq 0}\frac{a_n}{n!}c_{mn}x^m = \sum_{m\geq 0}\left(\sum_{n\geq m}\frac{a_n}{n!}c_{mn}\right)x^m,$$

which is a power series in x.

Conversely, assume $f(x) = \sum_{m\geq 0} b_m x^m$ on $\mathbf{Z}_p$ where $b_m \in K$. Since this power series converges at 1, $b_m \to 0$. The sequence of powers x^m and the sequence of "falling powers" $x^{\underline{m}} = x(x-1)\cdots(x-(m-1))$ are each $\mathbf{Z}$-linear combinations of each other (since each sequence has one term of each degree with integral coefficients and leading coefficient 1), so $x^m = \sum_{k=0}^m s_{mk}x^{\underline{k}} = \sum_{k=0}^m s_{mk}k!\binom{x}{k}$ where[3] $s_{mk} \in \mathbf{Z}$, so $x^m = \sum_{k\geq 0} d_{mk}\binom{x}{k}$, where

[3]The integers s_{mk} are called Stirling numbers of the second kind.

$d_{mk} = s_{mk}k!$ for $k \le m$ and $d_{mk} = 0$ for $k > m$. Note $k! \mid d_{mk}$. Then

$$f(x) = \sum_{m\ge 0} b_m \sum_{k \ge 0} d_{mk}\binom{x}{k} = \sum_{m \ge 0}\sum_{k\ge 0} b_m d_{mk}\binom{x}{k}.$$

Again we want to switch the order of summation. If $k > m$ then the (m, k)-th term is 0 since $d_{mk} = 0$. If $k \le m$ then $|b_m d_{mk}\binom{x}{k}| \le |b_m|$ since $d_{mk} \in \mathbf{Z}$, and this becomes arbitrarily small as m grows since $b_m \to 0$. Thus we can switch the order of summation and obtain

$$f(x) = \sum_{k\ge 0}\sum_{m\ge 0} b_m d_{mk}\binom{x}{k} = \sum_{k \ge 0}\left(\sum_{m\ge k} b_m d_{mk}\right)\binom{x}{k},$$

which is the Mahler expansion for f (by uniqueness). Since $k! \mid d_{mk}$,

$$\left|\frac{\sum_{m\ge k} b_m d_{mk}}{k!}\right| = \left|\sum_{m\ge k} b_m \frac{d_{mk}}{k!}\right| \le \left|\sum_{m \ge k} b_m\right| \le \max_{m\ge k}|b_m|,$$

and this maximum tends to 0 as $k \to \infty$, so the kth Mahler coefficient of f divided by $k!$ tends to 0. □

Corollary 4.5. *For a in a p-adic field K with $|a - 1| < 1$, the continuous function a^x for $x \in \mathbf{Z}_p$ is p-adic analytic on $\mathbf{Z}_p$ if and only if $|a - 1| < (1/p)^{1/([p-1)}$.*

Proof. Since $a^x = \sum_{n\ge 0}(a - 1)^n\binom{x}{n}$, this function is representable by a power series on $\mathbf{Z}_p$ if and only if $|(a - 1)^n/n!| \to 0$. We want to show this condition holds if and only if $|a - 1| < (1/p)^{1/(p-1)}$. To start,

$$\left|\frac{(a-1)^n}{n!}\right| = \frac{|a-1|^n}{|n!|_p} = |a-1|^n p^{(n-s_p(n))/(p-1)} = \frac{(|a-1|p^{1/(p-1)})^n}{p^{s_p(n)/(p-1)}} \le (|a-1|p^{1/(p-1)})^n$$

since $s_p(n) \ge 0$. Therefore if $|a - 1| < (1/p)^{1/(p-1)}$ we get $(|a - 1|p^{1/(p-1)})^n \to 0$ as $n \to \infty$, so $|(a - 1)^n/n!| \to 0$. And if $|a - 1| \ge (1/p)^{1/(p-1)}$ then $|a - 1|p^{1/(p-1)} \ge 1$, so

$$\left|\frac{(a-1)^n}{n!}\right| = \frac{(|a-1|p^{1/(p-1)})^n}{p^{s_p(n)/(p-1)}} \ge \frac{1}{p^{s_p(n)/(p-1)}},$$

and when n is a power of p we have $s_p(n) = 1$, so the lower bound $|(a-1)^n/n!| \ge 1/p^{1/(p-1)}$ occurs infinitely often, which shows $(a - 1)^n/n!$ does not tend to 0 in K. □

If $|a - 1| < (1/p)^{1/(p-1)}$ then there is an explicit power series for a^x using the p-adic exponential and logarithm: $a^x = e^{x \log a} = \sum_{n\ge 0}((\log a)^n/n!)x^n$. This particular power series formula breaks down if $(1/p)^{1/(p-1)} \le |a - 1| < 1$, and Corollary 4.5 says there is no power series for a^x anyway in this case: the function is continuous but not p-adic analytic.

We have described two properties of a continuous function in terms of its Mahler coefficients: its maximal size (Theorem 4.1) and whether or not it is p-adic analytic (Theorem 4.4). There is a criterion for pointwise differentiability also: $f(x) = \sum_{n\ge 0} a_n\binom{x}{n}$ is differentiable at a p-adic integer y if and only if $(\Delta^n f)(y)/n \to 0$ as $n \to \infty$, in which case $f'(y) = \sum_{n\ge 1}(-1)^{n-1}(\Delta^n f)(y)/n$. For example, if $f(x) = \sum_{r\ge 0} p^r\binom{x}{p^r}$ then $(\Delta^{p^k} f)(y) = p^k + \sum_{r\ge k+1} p^r\binom{y}{p^r - p^k}$, so $|(\Delta^{p^k} f)(y)/p^k|_p = 1$ and thus differentiability at each y fails: the series $\sum_{r\ge 0} p^r\binom{x}{p^r}$ is a continuous nowhere differentiable function $\mathbf{Z}_p \to \mathbf{Q}_p$.

The equality of maxima in Theorem 4.1 can be interpreted as an isometry between two metric spaces, as follows. Let $c_0(\mathbf{Q}_p) = \{(a_0, a_1, a_2, \dots) : a_n \in \mathbf{Q}_p, a_n \to 0\}$ be the set of all sequences in $\mathbf{Q}_p$ that tend to 0. For example, $c_0(\mathbf{Q}_p)$ contains $(1, p, p^2, \dots, p^n, \dots)$ and all sequences whose terms eventually equal 0. Under componentwise addition and the natural scaling rule $t(a_0, a_1, \dots, a_n, \dots) = (ta_0, ta_1, \dots, ta_n, \dots)$, $c_0(\mathbf{Q}_p)$ is a vector space over $\mathbf{Q}_p$. Make $c_0(\mathbf{Q}_p)$ a metric space by defining the distance between two sequences in $c_0(\mathbf{Q}_p)$ to be the maximum distance between terms in corresponding positions:

$$d(\mathbf{a}, \mathbf{b}) = \max_{n \geq 0} |a_n - b_n|_p.$$

Sending each $f \in C(\mathbf{Z}_p, \mathbf{Q}_p)$ to its sequence $(a_0, a_1, \dots)$ of Mahler coefficients is a mapping $C(\mathbf{Z}_p, \mathbf{Q}_p) \to c_0(\mathbf{Q}_p)$ that is a bijection: it is surjective by Section 2 and it is injective since a continuous function $\mathbf{Z}_p \to \mathbf{Q}_p$ is completely determined by its Mahler coefficients. Theorem 4.1 tells us that if $f(x) = \sum_{n \geq 0} a_n \binom{x}{n}$ and $g(x) = \sum_{n \geq 0} b_n \binom{x}{n}$ in $C(\mathbf{Z}_p, \mathbf{Q}_p)$ then $\max_{x \in \mathbf{Z}_p} |f(x) - g(x)|_p = \max_{n \geq 0} |a_n - b_n|_p$, which says the distance between two functions $C(\mathbf{Z}_p, \mathbf{Q}_p)$ and between their Mahler coefficient sequences in $c_0(\mathbf{Q}_p)$ are equal.

5. Integration of functions in $C(\mathbf{Z}_p, \mathbf{Q}_p)$

One of the most important applications of the Mahler expansion is in the development of p-adic integration. What is a p-adic integral? We will abstract properties of the definite integral of continuous functions $f\colon [0, 1] \to \mathbf{R}$ to define integration of functions in $C(\mathbf{Z}_p, \mathbf{Q}_p)$.

For $f \in C([0, 1], \mathbf{R})$, the definite integral $\int_0^1 f(x)\,dx$ is a real number and integration is a mapping $C([0, 1], \mathbf{R}) \to \mathbf{R}$ that has a few properties:

- Linearity: $\int_0^1 (af(x) + bg(x))\,dx = a\int_0^1 f(x)\,dx + b\int_0^1 g(x)\,dx$ for $a, b \in \mathbf{R}$ and $f, g \in C([0, 1], \mathbf{R})$,
- Continuity: functions that are close in $C([0, 1], \mathbf{R})$ have close integrals. This is made explicit with the bound $|\int_0^1 (f(x) - g(x))\,dx| \leq \max_{x \in [0,1]} |f(x) - g(x)|$ for all $f, g \in C([0, 1], \mathbf{R})$.

Integration also preserves positivity (if $f \geq 0$ on $[0, 1]$ then $\int_0^1 f(x)\,dx \geq 0$), but we don't pay attention to this since it doesn't carry over to the p-adics.

There are other mappings $I\colon C([0, 1], \mathbf{R}) \to \mathbf{R}$ that are linear and continuous besides $I(f) = \int_0^1 f(x)\,dx$. Examples include $I(f) = f(0)$ and $I(f) = \int_0^1 f(x)(2x^2 + x)\,dx$. They can be regarded as generalized integrals on $C([0, 1], \mathbf{R})$.

Definition 5.1. An *integral* on $C(\mathbf{Z}_p, \mathbf{Q}_p)$ is a continuous $\mathbf{Q}_p$-linear map $C(\mathbf{Z}_p, \mathbf{Q}_p) \to \mathbf{Q}_p$.

Example 5.2. For a *bounded* sequence $\mathbf{b} = \{b_n\}$ in $\mathbf{Q}_p$, define $I_{\mathbf{b}}\colon C(\mathbf{Z}_p, \mathbf{Q}_p) \to \mathbf{Q}_p$ by

$$I_{\mathbf{b}}(f) = \sum_{n \geq 0} a_n b_n, \tag{5.1}$$

where the a_n's are the Mahler coefficients of f. This series converges: letting $|b_n|_p \leq B$ for all $n \geq 0$, we have $|a_n b_n| \leq |a_n| B \to 0$. It is easy to see $I_{\mathbf{b}}$ is $\mathbf{Q}_p$-linear, and it is continuous since

$$|I_{\mathbf{b}}(f)|_p = \left|\sum_{n \geq 0} a_n b_n\right|_p \leq \max_{n \geq 0} |a_n b_n|_p \leq \max_{n \geq 0} |a_n|_p B = \max_{x \in \mathbf{Z}_p} |f(x)|_p B,$$

so small continuous functions on $\mathbf{Z}_p$ have small integrals.

In effect we are defining $I_{\mathbf{b}}(\binom{x}{n}) = b_n$ for all n and extending $I_{\mathbf{b}}$ to the rest of $C(\mathbf{Z}_p, \mathbf{Q}_p)$ by linearity and continuity. The continuity meant here is for the metric d on $C(\mathbf{Z}_p, \mathbf{Q}_p)$ that makes it complete: if $f(x) = \sum_{n\geq 0} a_n\binom{x}{n}$ then $\sum_{n=0}^N a_n\binom{x}{n} \to f$ in $C(\mathbf{Z}_p, \mathbf{Q}_p)$ as $N \to \infty$, so

$$I_{\mathbf{b}}\left(\sum_{n\geq 0} a_n\binom{x}{n}\right) = I_{\mathbf{b}}\left(\lim_{N\to\infty}\sum_{n\geq N} a_n\binom{x}{n}\right) = \lim_{N\to\infty} I_{\mathbf{b}}\left(\sum_{n=0}^{N} a_n\binom{x}{n}\right),$$

and by linearity $I_{\mathbf{b}}(\sum_{n=0}^N a_n\binom{x}{n}) = \sum_{n=0}^N a_n I_{\mathbf{b}}\left(\binom{x}{n}\right) = \sum_{n=0}^N a_n b_n$. Passing to the limit as $N \to \infty$, $I_{\mathbf{b}}\left(\sum_{n\geq 0} a_n\binom{x}{n}\right) = \sum_{n\geq 0} a_n b_n$.

It can be shown that this example is in fact completely general: every integral on $C(\mathbf{Z}_p, \mathbf{Q}_p)$ is an infinite dot product of Mahler coefficients against a bounded sequence, as in (5.1).[4] Specific choices of bounded sequences in $\mathbf{Q}_p$ (or in other p-adic fields) define integrals that lead to p-adic zeta-functions and L-functions [2, Chap. 4]. These are p-adic analogues of the zeta and L-functions on $\mathbf{C}$ from analytic number theory. Integration on $C(\mathbf{Z}_p, \mathbf{Q}_p)$ can also be developed without using Mahler expansions [1, Chap. 2].

Appendix A. Completeness

We want to prove that $C([0,1], \mathbf{R})$ and $C(\mathbf{Z}_p, \mathbf{Q}_p)$ are complete with respect to the metric described in (1.1). All that matters about $[0,1]$ and $\mathbf{Z}_p$ is that they are compact in order for the maximum in (1.1) to occur. We will prove completeness of these spaces as a special case of completeness of the space of continuous functions from any compact metric space to a complete valued field.

Theorem A.1. *Let (X, d_X) be a compact metric space and $(K, |\cdot|)$ be a complete valued field. On the set $C(X, K)$ of all continuous functions from X to K, define*

$$d(f, g) = \max_{x\in X} |f(x) - g(x)|.$$

This is a metric and $C(X, K)$ is complete for this metric.

Proof. It is easy to check that d fits the axioms for a metric:

(1) Easily $d(f, g) \geq 0$ with equality if and only if $f(x) = g(x)$ for all $x \in X$ (so $f = g$ in $C(X, K)$)
(2) Obviously $d(f, g) = d(g, f)$.
(3) (Triangle inequality) For three continuous functions $f, g, h \in C(X, K)$, let $d(f, g) = |f(x_0) - g(x_0)|$. Using the triangle inequality for the absolute value on K,

$$d(f, g) = |f(x_0) - g(x_0)| \leq |f(x_0) - h(x_0)| + |h(x_0) - g(x_0)| \leq d(f, h) + d(h, g).$$

Actually, the only delicate issue about the metric d is that it *makes sense*: the maximum really exists. We discussed this in Section 1 for $K = \mathbf{R}$ and $\mathbf{Q}_p$, but do it again. Showing there is a maximum relies on compactness of X: for continuous functions $f, g\colon X \to K$, the function $|f - g|\colon X \to \mathbf{R}$ is continuous (the difference $f - g$ is continuous $X \to K$ and the absolute value $|\cdot|\colon K \to \mathbf{R}$ is continuous, and the composite of continuous functions is continuous) and a continuous real-valued function on any compact metric space has a maximum (as well as a minimum) value.

[4] See Theorems 3.2 and 3.4 of https://kconrad.math.uconn.edu/blurbs/analysis/sequencespaceto0.pdf.

To show every Cauchy sequence $\{f_n\}$ in $C(X,K)$ has a limit in $C(X,K)$ will be broken up into three steps:

Step 1: Create a candidate limit function f.

For each $a \in X$ the sequence of numbers $\{f_n(a)\}$ in K is Cauchy since

$$|f_m(a) - f_n(a)| \leq \max_{x \in X} |f_m(x) - f_n(x)| = d(f_m, f_n)$$

and the value on the right is arbitrarily small for all large enough m and n. Therefore $\lim_{n\to\infty} f_n(a)$ exists in K by completeness. Call the limit value $f(a)$. We have defined a function $f\colon X \to K$.

Step 2: Show f is continuous.

This will be proved with an $\varepsilon/3$ argument.

Pick $a \in X$ and $\varepsilon > 0$. We need to find $\delta > 0$ such that $d_X(a,X) < \delta \Longrightarrow |f(x) - f(a)| < \varepsilon$.

Since $\{f_n\}$ is Cauchy, there is N such that $m, n \geq N \Longrightarrow d(f_m, f_n) < \varepsilon/3$, so for all $x \in X$ and $m \geq N$ we have $|f_m(x) - f_N(x)| < \varepsilon/3$. Letting $m \to \infty$ in this inequality, $|f(x) - f_N(x)| \leq \varepsilon/3$ for all x in X. Thus for each x in X,

$$\begin{aligned} |f(x) - f(a)| &\leq |f(x) - f_n(x)| + |f_N(x) - f_N(a)| + |f_N(a) - f(a)| \\ &\leq \frac{\varepsilon}{3} + |f_N(x) - f_N(a)| + \frac{\varepsilon}{3} \\ &= \frac{2}{3}\varepsilon + |f_N(x) - f_N(a)|. \end{aligned}$$

Since f_N is continuous at a, there is $\delta > 0$ such that $|x-a| < \delta \Longrightarrow |f_N(x) - f_N(a)| < \varepsilon/3$. Therefore

$$|x - a| < \delta \Longrightarrow |f(x) - f(a)| \leq \frac{2}{3}\varepsilon + |f_N(x) - f_N(a)| < \varepsilon,$$

which proves f is continuous at each $a \in X$.

Step 3: Show $d(f_n, f) \to 0$.

We essentially repeat the beginning of Step 2. From $\{f_n\}$ being Cauchy there is an N such that $m, n \geq N \Longrightarrow d(f_m, f_n) < \varepsilon/3$, so for all $x \in X$ and $m, n \geq N$ we have $|f_m(x) - f_n(x)| < \varepsilon/3$. Letting $m \to \infty$ here, we get $|f(x) - f_n(x)| \leq \varepsilon/3$ for all $x \in X$. Therefore $d_\infty(f, f_n) \leq \varepsilon/3 < \varepsilon$.

□

A similar theorem can be formulated where the domain X is not assumed to be compact, but then we have to work with the continuous functions that are *bounded* (which is automatic for continuous functions $X \to K$ when X is compact). If (X, d_X) is a metric space and $C_b(X,K)$ is the set of bounded continuous functions $X \to K$, then a metric on $C_b(X,K)$ is given by

$$d(f,g) = \sup_{x \in X} |f(x) - g(x)|,$$

where we use a supremum instead of a maximum because there is no guarantee (without X being compact) that a maximum value of $|f(x) - g(x)|$ over all x exists; when f and g have bounded values in K then the real numbers $|f(x) - g(x)|$ are bounded so the supremum of this set in $\mathbf{R}$ exists. Since it is not guaranteed that the supremum defining $d(f,g)$ is achieved, we can't assume $d(f,g) = |f(x_0) - g(x_0)|$ for some $x_0 \in X$, so proving d is a metric on $C_b(X,K)$ requires a little ε-fiddling with the definition of the supremum compared to the case when X is compact. This is left to the reader.

Theorem A.2. *If* (X, d_X) *is any metric space and* $(K, |\cdot|)$ *is a complete valued field then* $C_b(X, K)$ *is complete for the metric* d *above.*

The proof of this is left to the reader. It is similar to the proof of Theorem A.1, but there is an extra step: proving the limit function is bounded, not just continuous.

There is a theorem of this type for functions having values not just in a complete valued field, but in any complete metric space.

Theorem A.3. *For metric spaces* (X, d_X) *and* (Y, d_Y)*, let* $C_b(X, Y)$ *denote the set of bounded continuous functions from* X *to* Y*. On* $C_b(X, Y)$ *the function*

$$d(f, g) = \sup_{x \in X} d_Y(f(x), g(x))$$

is a metric, and if Y *is complete for* d_Y *then* $C_b(X, Y)$ *is complete for* d*.*

The proof is very much like that of Theorem A.2 except for added notation: expressions like $|f(x) - f(a)|$ have to be replaced by $d_Y(f(x), f(a))$.

References

[1] N. Koblitz, *p-Adic Numbers, p-Adic Analysis, and Zeta-functions*, 2nd ed., Springer-Verlag, 1984.

[2] S. Lang, *Cyclotomic Fields I and II*, combined 2nd ed., Springer-Verlag, 1990.

[3] K. Mahler, "An interpolation series for continuous functions of a p-adic variable", Journal für die reine und angewandte Mathematik, **199** (1958), 23–34.

[4] W. Schikhof, *Ultrametric Calculus: An Introduction to p-Adic Analysis*, Cambridge Univ. Press, 1984.