INFINITE SERIES IN p-ADIC FIELDS

KEITH CONRAD

1. Introduction

One of the major topics in a course on real analysis is the representation of functions as power series

$$\sum_{n\geq 0} a_n x^n,$$

where the coefficients a_n are real numbers and the variable x belongs to $\mathbf{R}$. In these notes we will develop the theory of power series over complete nonarchimedean fields.

Let K be a field that is complete with respect to a nontrivial nonarchimedean absolute value $|\cdot|$, such as $\mathbf{Q}_p$ with its absolute value $|\cdot|_p$. We will look at power series over K (that is, with coefficients in K) and a variable x coming from K. Such series share some properties in common with real power series:

(1) There is a radius of convergence R (a nonnegative real number, or ∞), for which there is convergence at $x \in K$ when $|x| < R$ and not when $|x| > R$.
(2) A power series is uniformly continuous on each closed disc (of positive radius) where it converges.
(3) A power series can be differentiated termwise in its disc of convergence.

There are also some contrasts:

(1) The convergence of a power series at its radius of convergence R does not exhibit mixed behavior: it is either convergent at all x with $|x| = R$ or not convergent at all x with $|x| = R$, unlike $\sum_{n\geq 1} x^n/n$ on $\mathbf{R}$, where $R = 1$ and there is convergence at $x = -1$ but not at $x = 1$.
(2) A power series can be expanded around a new point in its disc of convergence and the new series has exactly the same disc of convergence as the original series. Over $\mathbf{R}$, a recentered power series often converges at some numbers outside the interval of convergence of the original series.

2. Generalities on infinite series

Before discussing power series, we set out a few general properties of infinite series over K. Since our eventual goal is power series, we will write infinite series as $\sum_{n\geq 0} a_n$, with the indexing starting at $n = 0$ instead of $n = 1$.

The most basic property of an infinite series $\sum_{n\geq 0} a_n$ over K is that it converges if and only if $|a_n| \to 0$: convergence implies the general term tends to 0 by the same reasoning as with infinite series over $\mathbf{R}$, and conversely the general term tending to 0 makes the sequence of partial sums a Cauchy sequence in K by the strong triangle inequality, so the series converges in K by completeness.

Theorem 2.1. *If $a_n \to 0$ in K then $|\sum_{n\geq 0} a_n| \leq \max_{n\geq 0} |a_n|$.*

Proof. Let $s_n = a_0 + \cdots + a_n$ be the nth partial sum and $s = \sum_{n\geq 0} a_n = \lim_{n\to\infty} s_n$. If $s = 0$ then $|\sum_{n\geq 0} a_n| = 0 \leq \max |a_n|$. If $s \neq 0$ then $|s| = |s_N|$ for $N \gg 0$ by the strong triangle inequality: for large N we we have $|s - s_N| < |s|$, so $|s| = |s_N|$. Since s_N has only finitely many terms,

$$|s_N| = |a_0 + \cdots + a_N| \leq \max(|a_0|, \ldots, |a_N|) \leq \max_{n\geq 0} |a_n|. \qquad \square$$

Theorem 2.2. *If $\sum_{n\geq 0} a_n$ and $\sum_{n\geq 0} b_n$ converge then $\sum_{n\geq 0}(a_n+b_n) = \sum_{n\geq 0} a_n + \sum_{n\geq 0} b_n$ and, for each $c \in K$, $\sum_{n\geq 0} ca_n = c\sum_{n\geq 0} a_n$.*

Proof. Let $A = \sum_{n\geq 0} a_n$ and $B = \sum_{n\geq 0} b_n$. Then $\sum_{n=0}^N a_n \to A$ and $\sum_{n=0}^N b_n \to B$ as $N \to \infty$, so by continuity of addition and multiplication with respect to $|\cdot|$ we have

$$\sum_{n=0}^N (a_n + b_n) = \sum_{n=0}^N a_n + \sum_{n=0}^N b_n \to A + B$$

and

$$\sum_{n=0}^N ca_n = c\sum_{n=0}^N a_n \to cA$$

as $N \to \infty$. $\square$

Theorem 2.3 (Comparison Test)**.** *If $\{a_n\}$ is a sequence in K and there is a sequence $\{b_n\}$ in $\mathbf{R}$ such that $|a_n| \leq b_n$ for $n \gg 0$, then convergence of $\sum_{n\geq 0} b_n$ in $\mathbf{R}$ implies convergence of $\sum_{n\geq 0} a_n$ in K.*

Proof. Exercise. This is true for all complete valued fields K using just the triangle inequality, not the strong triangle inequality. $\square$

One of the subtle aspects of infinite series in the real numbers is that a convergent series can have its terms rearranged to produce an infinite series with a different value.

Example 2.4. Set $L = 1-1/2+1/3-1/4+1/5-\cdots = \sum_{n\geq 1}(-1)^{n-1}/n$ be the alternating harmonic series in $\mathbf{R}$. Its value is between 1 and $1 - 1/2 = 1/2$. (By calculus $L = \ln(2) \approx .693$, but we don't need this.) Let's rearrange the terms of the series so each positive term is followed by two negative terms:

$$L' = 1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \frac{1}{12} + \cdots$$

To compute L', add each positive term to the negative term immediately after it, which will not affect the value of L' since the terms are kept in the same order:

$$L' = \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \frac{1}{10} - \frac{1}{12} + \cdots$$

This is precisely half the alternating harmonic series, so $L' = L/2$. By rearranging the terms of a series we obtained a new series with half the value of the original series!

This example is due to Dirichlet (in an 1837 paper on number theory), and it illustrates why care is needed when using series in $\mathbf{R}$. Dirichlet proved (in that same paper) a sufficient condition for a convergent infinite series $\sum_{n\geq 0} a_n$ to have all of its rearrangements converge to the same value: the series is absolutely convergent, meaning $\sum_{n\geq 0} |a_n|$ converges. This is the primary reason for the significance of absolute convergence in real analysis. Riemann later proved that absolute convergence is not just sufficient for all rearrangements of a

series in $\mathbf{R}$ to have the same value, but it is necessary as well: a series in $\mathbf{R}$ that is not absolutely convergent (like the alternating harmonic series in Example 2.4) can have its terms rearranged to converge to an arbitrary real number or to be ∞ or $-\infty$.

While the concept of absolute convergence makes sense for an infinite series over a complete nonarchimedean field K, there is no need for it because there are no difficulties with rearrangments:

Theorem 2.5. *If $a_n \to 0$ in K and $\{a'_n\}$ is a rearrangement of $\{a_n\}$ then $a'_n \to 0$ and $\sum_{n\geq 0} a'_n = \sum_{n\geq 0} a_n$.*

Proof. Pick $\varepsilon > 0$. To prove $a'_n \to 0$ we seek an N such that $n > N \Longrightarrow |a'_n| \leq \varepsilon$.

Since $a_n \to 0$ there is an $\widetilde{N}$ such that $n > \widetilde{N} \Longrightarrow |a_n| \leq \varepsilon$. Since $\{a_n\} = \{a'_n\}$, the list $\{a_1, a_2, \ldots, a_{\widetilde{N}}\}$ lies inside the list $\{a'_1, a'_2, \ldots, a'_N\}$ for some N. Therefore $n > N \Longrightarrow a'_n \notin \{a_1, a_2, \ldots, a_{\widetilde{N}}\}$, so $|a'_n| \leq \varepsilon$.

Set $s = \sum_{n\geq 0} a_n$. The series $\sum_{n\geq 0} a'_n$ converges in K since the general term tends to 0. We want to show $\sum_{n\geq 0} a'_n = s$:

$$a'_1 + \cdots + a'_n \to s \text{ as } n \to \infty.$$

Pick $\varepsilon > 0$. Since $a_n \to 0$, the index set $A = \{n : |a_n| > \varepsilon\}$ is *finite*. Let m be the largest integer in A, so $\{n : |a_n| > \varepsilon\} \subset \{1, 2, \ldots, m\}$. Then

$$\begin{aligned} s - \sum_{n\in A} a_n &= s - \sum_{n=0}^{m} a_n + \sum_{n=0}^{m} a_n - \sum_{n\in A} a_n \\ &= \sum_{n>m} a_n + \left(\sum_{n=0}^{m} a_n - \sum_{n\in A} a_n \right). \end{aligned}$$

In the first series each term has absolute value at most ε, so the series has absolute value at most ε (Theorem 2.1). In the parentheses we have two finite sums, and the last sum removes from the first sum all terms that have absolute value greater than ε. Therefore the number in parentheses is a sum of finitely many terms that are each at most ε in absolute value, so by the strong triangle inequality

$$\left| s - \sum_{n\in A} a_n \right| \leq \varepsilon. \tag{2.1}$$

Next let $A' = \{n : |a'_n| > \varepsilon\}$, another finite set of indices. Set $m' = \max\{n : |a'_n| > \varepsilon\}$, so reasoning as above,

$$N \geq m' \Longrightarrow \left| \sum_{n=0}^{N} a'_n - \sum_{n\in A'} a'_n \right| \leq \varepsilon, \tag{2.2}$$

The finite lists of numbers $\{a_n : n \in A\}$ and $\{a'_n : n \in A'\}$ are identical, since they're both just the terms from the same list (in original and rearranged form) having absolute value exceeding ε. Thus

$$\sum_{n\in A'} a'_n = \sum_{n\in A} a_n, \tag{2.3}$$

so for $N \geq m'$

$$\begin{aligned}\left|s - \sum_{n=0}^{N} a'_n\right| &= \left|s - \sum_{n\in A} a_n + \sum_{n\in A} a_n - \sum_{n=0}^{N} a'_n\right| \\ &= \left|s - \sum_{n\in A} a_n + \sum_{n\in A'} a'_n - \sum_{n=0}^{N} a'_n\right| \\ &\leq \max\left(\left|s - \sum_{n\in A} a_n\right|, \left|\sum_{n\in A'} a'_n - \sum_{n=0}^{N} a'_n\right|\right).\end{aligned}$$

The first term in the maximum is at most ε by (2.1) and the second term is at most ε by (2.2). □

Another important operation on series besides rearranging terms is swapping the order of a double series:

$$\sum_{m\geq 0}\sum_{n\geq 0} a_{mn} \stackrel{?}{=} \sum_{n\geq 0}\sum_{m\geq 0} a_{mn}. \tag{2.4}$$

To be clear, the meaning of a double series $\sum_{m\geq 0}\sum_{n\geq 0} a_{mn}$ is $\sum_{m\geq 0}(\sum_{n\geq 0} a_{mn})$. Putting the doubly-indexed terms a_{mn} in an infinite matrix

$$\begin{pmatrix} a_{00} & a_{01} & a_{02} & \dots \\ a_{10} & a_{11} & a_{12} & \dots \\ a_{20} & a_{21} & a_{22} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{2.5}$$

the left side of (2.4) is the sum of the row series and the right side is the sum of the column series.

Double series arise when justifying operations on power series like multiplication and termwise differentiation (we will see this later). Therefore it is important to know conditions under which the two sides of (2.4) are equal.

When a double series on one side of (2.4) converges, it is not generally true that the other double series converges to the same value, or even converges at all. Here are examples in $\mathbf{R}$ and in each $\mathbf{Q}_p$.

Example 2.6. Let $a_{m,n}$ be given by the matrix below: $a_{m,0} = 1/(m+1)$, $a_{m,m} = -1/(m+1)$ for $m > 0$, and all other $a_{m,n}$ are 0.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \dots \\ 1/2 & -1/2 & 0 & 0 & \dots \\ 1/3 & 0 & -1/3 & 0 & \dots \\ 1/4 & 0 & 0 & -1/4 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

The row sums are $\sum_{n\geq 0} a_{0,n} = 1$ and $\sum_{n\geq 0} a_{m,n} = 0$ for $m \geq 1$, so $\sum_{m\geq 0}\sum_{n\geq 0} a_{m,n} = 1$. However, the first column sum $\sum_{m\geq 0} a_{m,0}$ is the harmonic series, so $\sum_{n\geq 0}\sum_{m\geq 0} a_{m,n}$ does not converge in $\mathbf{R}$ even though the later column sums $\sum_{m\geq 0} a_{m,n}$ each converge, consisting of just a single nonzero term each.

Example 2.7. In $\mathbf{Q}_p$ consider $\sum_{m\geq 0} p^m = 1/(1-p)$. Writing p^m as a sum of p^m 1's:

$$\frac{1}{1-p} = \sum_{m\geq 0} p^m = \sum_{m\geq 0}\sum_{n=0}^{p^m-1} 1 = \sum_{m\geq 0}\sum_{n\geq 0} a_{mn},$$

where

$$a_{mn} = \begin{cases} 1, & \text{if } 0 \leq n < p^m, \\ 0, & \text{if } n \geq p^m, \end{cases}$$

so $\sum_{m\geq 0}\sum_{n\geq 0} a_{mn} = 1/(1-p)$ in $\mathbf{Q}_p$. In $\sum_{n\geq 0}\sum_{m\geq 0} a_{mn} = \sum_{n\geq 0}\left(\sum_{\{m:p^m>n\}} 1\right)$, the nth inner sum (for each n) does not converge since it is a sum of infinitely many 1's.

Example 2.8. Define a_{mn} to be 1, -1 or 0 by the rules $a_{nn} = 1$, $a_{n+1,n} = -1$, and $a_{mn} = 0$ if $m \neq n$ or $n+1$:

$$(a_{mn}) = \begin{pmatrix} 1 & 0 & 0 & \dots \\ -1 & 1 & 0 & \dots \\ 0 & -1 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

In $\mathbf{R}$ and in $\mathbf{Q}_p$, $\sum_{m\geq 0}\sum_{n\geq 0} a_{mn}$ is the sum of row sums, which is 1, while $\sum_{n\geq 0}\sum_{m\geq 0} a_{mn}$ is the sum of column sums, which is 0. The two double sums converge with different values.

In real analysis, we can guarantee the convergence of both sides of (2.4) and their equality by conditions related to absolute convergence (see [6, pp. 143-144] for a precise statement). In a complete nonarchimedean field K we don't use absolute convergence, but instead the following theorem with weaker hypotheses (see [2, Lemma 4.1.3, Prop. 4.1.4] and [5, Theorem 3.8]).

Theorem 2.9. *If $a_{mn} \to 0$ in K as $\max(m,n) \to \infty$ then both sides of (2.4) converge and are equal, and moreover both double series equal $\lim_{N\to\infty}\sum_{m=0}^{N}\sum_{n=0}^{N} a_{mn}$.*

Let's be clear about what the condition "$a_{mn} \to 0$ as $\max(m,n) \to \infty$" means. Here are three ways to think about it.

(1) As long as one of the indices in a_{mn} is big, a_{mn} is small. (Don't confuse this with requiring both indices to be big for a_{mn} to be small. That would mean $a_{mn} \to 0$ as $\min(m,n) \to \infty$, often written as $\lim_{m,n\to\infty} a_{mn} = 0$.)
(2) For each $\varepsilon > 0$ there is an N such that if $\max(m,n) \geq N$ then $|a_{mn}| \leq \varepsilon$.
(3) Putting the doubly-indexed terms a_{mn} in an infinite matrix (a_{mn}) as in (2.5), the terms with $\max(m,n) \geq N$ are those outside the upper left $N \times N$ square, so what is happening when $a_{mn} \to 0$ as $\max(m,n) \to \infty$ is that all the terms of the matrix lying *outside* a sufficiently large upper left square are as small as desired. This implies the terms tend to 0 along each row and along each column, but it is stronger than that.

The equality (2.4) does not hold for the double sums in Examples 2.7 and 2.8, and the hypothesis of Theorem 2.9 is not satisfied in those examples: $a_{m,p^m-1} = 1$ in Example 2.7 and $a_{mm} = 1$ in Example 2.8, so $a_{mn} \not\to 0$ as $\max(m,n) \to \infty$. Now we prove Theorem 2.9.

Proof. First we show $\sum_{m\geq 0}\sum_{n\geq 0} a_{mn}$ makes sense. The inner sum $\sum_{n\geq 0} a_{mn}$, for each m, makes sense since it is a row sum in the matrix and the terms along each row tend to 0.

Set $b_m = \sum_{n\geq 0} a_{mn}$, so we want to show $\sum_{m\geq 0}\sum_{n\geq 0} a_{mn} = \sum_{m\geq 0} b_m$ makes sense. That is, why does $b_m \to 0$?

By Theorem 2.1, $|b_m| \leq \max_{n\geq 0}|a_{mn}|$. The hypothesis that $\lim_{\max(m,n)\to\infty} a_{mn} = 0$ implies for $\varepsilon > 0$ that there's an N such that $\max(m,n) \geq N \Rightarrow |a_{mn}| \leq \varepsilon$, so in particular $m \geq N \Rightarrow |a_{mn}| \leq \varepsilon$ for all n (if $m \geq N$ then $\max(m,n) \geq m \geq N$). Thus $m \geq N \Rightarrow |b_m| \leq \max_{n\geq 0}|a_{mn}| \leq \varepsilon$, so we have proved $b_m \to 0$ as $m \to \infty$. Thus we have convergence of $\sum_{m\geq 0} b_m = \sum_{m\geq 0}\sum_{n\geq 0} a_{mn}$ and

$$\left|\sum_{m\geq 0}\sum_{n\geq 0} a_{mn}\right| = \left|\sum_{m\geq 0} b_m\right| \leq \max_{m\geq 0}|b_m| \leq \max_{m\geq 0}\max_{n\geq 0}|a_{mn}|. \tag{2.6}$$

Swapping the roles of m and n shows the double series $\sum_{n\geq 0}\sum_{m\geq 0} a_{mn}$ also converges.

Now that we know both double series in (2.4) make sense, we show they are equal by comparing each one to the finite sum $\sum_{m=0}^{N}\sum_{n=0}^{N} a_{mn}$ as N grows (this is the sum of terms in the upper left $(N+1)\times(N+1)$ square of the matrix). Let's consider for each $N \geq 0$ the difference

$$\sum_{m\geq 0}\sum_{n\geq 0} a_{mn} - \sum_{m=0}^{N}\sum_{n=0}^{N} a_{mn} = \sum_{m=0}^{N}\left(\sum_{n\geq 0} a_{mn}\right) + \sum_{m\geq N+1}\left(\sum_{n\geq 0} a_{mn}\right) - \sum_{m=0}^{N}\sum_{n=0}^{N} a_{mn}.$$

Combining the first and third sums,

$$\begin{aligned}\sum_{m\geq 0}\sum_{n\geq 0} a_{mn} - \sum_{m=0}^{N}\sum_{n=0}^{N} a_{mn} &= \sum_{m=0}^{N}\left(\sum_{n\geq 0} a_{mn} - \sum_{n=0}^{N} a_{mn}\right) + \sum_{m\geq N+1}\sum_{n\geq 0} a_{mn} \\ &= \sum_{m=0}^{N}\sum_{n\geq N+1} a_{mn} + \sum_{m\geq N+1}\sum_{n\geq 0} a_{mn}. \end{aligned} \tag{2.7}$$

Pick $\varepsilon > 0$. There is an $N \geq 0$ such that $\max(m,n) \geq N+1 \Longrightarrow |a_{mn}| \leq \varepsilon$. Then $n \geq N+1 \Longrightarrow |a_{mn}| \leq \varepsilon$ for each m and $m \geq N+1 \Longrightarrow |a_{mn}| \leq \varepsilon$ for each n, so in (2.7) both double series have absolute value at most ε (argue as in (2.6)). Hence for each $\varepsilon > 0$ there is an $N \geq 0$ such that

$$\left|\sum_{m\geq 0}\sum_{n\geq 0} a_{mn} - \sum_{m=0}^{N}\sum_{n=0}^{N} a_{mn}\right| \leq \varepsilon,$$

which proves $\lim_{N\to\infty}\sum_{m=0}^{N}\sum_{n=0}^{N} a_{mn} = \sum_{m\geq 0}\sum_{n\geq 0} a_{mn}$. By swapping the roles of m and n, $\lim_{N\to\infty}\sum_{n=0}^{N}\sum_{m=0}^{N} a_{mn} = \sum_{n\geq 0}\sum_{m\geq 0} a_{mn}$ too, and since $\sum_{m=0}^{N}\sum_{n=0}^{N} a_{mn} = \sum_{n=0}^{N}\sum_{m=0}^{N} a_{mn}$ we get

$$\sum_{m\geq 0}\sum_{n\geq 0} a_{mn} = \lim_{N\to\infty}\sum_{m=0}^{N}\sum_{n=0}^{N} a_{mn} = \sum_{n\geq 0}\sum_{m\geq 0} a_{mn}. \qquad \square$$

Corollary 2.10. *If $a_{mn} \to 0$ in K as $\max(m,n) \to \infty$ then*

$$\sum_{m\geq 0}\sum_{n\geq 0} a_{mn} = \sum_{\ell\geq 0}\sum_{m+n=\ell} a_{mn},$$

where the inner sum on the right runs over the finitely many pairs of nonnegative integers (m, n) adding up to ℓ.

Proof. By Theorem 2.9, $\displaystyle\sum_{m\geq 0}\sum_{n\geq 0} a_{mn} = \lim_{L\to\infty}\sum_{m=0}^{L}\sum_{n=0}^{L} a_{mn}$. We will show the limit as $L\to\infty$ equals $\sum_{\ell\geq 0}\sum_{m+n=\ell} a_{mn}$. For $m, n \geq 0$, if $m+n=\ell$ then $m\leq \ell$ and $n\leq \ell$, so for $L\geq 0$

$$\sum_{m=0}^{L}\sum_{n=0}^{L} a_{mn} - \sum_{\ell=0}^{L}\sum_{m+n=\ell} a_{mn} = \sum_{\substack{0\leq m,n\leq L\\ m+n\geq L}} a_{mn}. \tag{2.8}$$

In the finite sum on the right, $m+n\geq L \Longrightarrow m\geq L/2$ or $n\geq L/2$.

Pick $\varepsilon > 0$. By hypothesis there is an N such that if $\max(m,n)\geq N$ then $|a_{mn}|\leq\varepsilon$, so if $m+n\geq 2N$ then m or n is $\geq N$, so $|a_{mn}|\leq\varepsilon$. Thus if $L\geq 2N$ then (2.8) has absolute value at most ε. That proves (2.8) tends to 0 as $L\to\infty$, which is what we wanted. □

If we want to write the product of two convergent series $\sum_{m\geq 0} a_m$ and $\sum_{n\geq 0} b_n$ as a single convergent series of two-term products $a_m b_n$, it is natural to list the terms according to increasing value of $m+n$, getting

$$a_0b_0 + (a_0b_1 + a_1b_0) + (a_0b_2 + a_1b_1 + a_2b_0) + \cdots = \sum_{\ell\geq 0}\sum_{m+n=\ell} a_m b_n. \tag{2.9}$$

In $\mathbf{R}$ this series need not converge: if $a_n = b_n = (-1)^n/\sqrt{n+1}$, then $\sum_{n\geq 0} a_n$ and $\sum_{n\geq 0} b_n$ converge by the alternating series test but (2.9) does not converge since $|\sum_{m+n=\ell} a_m b_n| \geq 1$ for all ℓ. To have (2.9) converge in $\mathbf{R}$ and equal $(\sum_{m\geq 0} a_m)(\sum_{n\geq 0} b_n)$, absolute convergence of both of the original series $\sum_{m\geq 0} a_m$ and $\sum_{n\geq 0} b_n$ is sufficient [6, pp. 146-147]. (In fact absolute convergence of just one of the original series is sufficient [6, p. 321].) Over a nonarchimedean complete field, let's show no problems occur.

Corollary 2.11. *If $\sum_{m\geq 0} a_m$ and $\sum_{n\geq 0} b_n$ converge in K then*

$$\left(\sum_{m\geq 0} a_m\right)\left(\sum_{n\geq 0} b_n\right) = \sum_{\ell\geq 0}\sum_{m+n=\ell} a_m b_n.$$

Proof. Set $s = \sum_{n\geq 0} b_n$. Then

$$\begin{aligned}\left(\sum_{m\geq 0} a_m\right)\left(\sum_{n\geq 0} b_n\right) &= \left(\sum_{m\geq 0} a_m\right) s\\ &= \sum_{m\geq 0} a_m s\\ &= \sum_{m\geq 0} a_m\left(\sum_{n\geq 0} b_n\right)\\ &= \sum_{m\geq 0}\sum_{n\geq 0} a_m b_n\\ &= \sum_{m\geq 0}\sum_{n\geq 0} c_{mn}\end{aligned}$$

where $c_{mn} = a_m b_n$. Let's check $c_{mn} \to 0$ as $\max(m,n) \to \infty$.

Since $a_m \to 0$ as $m \to \infty$ and $b_n \to 0$ as $n \to \infty$, both sequences are bounded, say $|a_m| \leq A$ for all m and $|b_n| \leq B$ for all n. Then for $\varepsilon > 0$ there are M and N such that $m \geq M \Rightarrow |a_m| \leq \varepsilon$ and $n \geq N \Rightarrow |b_n| \leq \varepsilon$. In the first case, $|c_{mn}| = |a_m||b_n| \leq |a_m|B \leq \varepsilon B$, and in the second case $|c_{mn}| = |a_m||b_n| \leq A|b_n| \leq A\varepsilon$, so together

$$\max(m,n) \geq \max(M,N) \Longrightarrow |c_{mn}| = |a_m||b_n| \leq \varepsilon \max(A,B).$$

Thus $c_{mn} \to 0$ as $\max(m,n) \to \infty$. We can now apply Corollary 2.10:

$$\left(\sum_{m\geq 0} a_m\right)\left(\sum_{n\geq 0} b_n\right) = \sum_{m\geq 0}\sum_{n\geq 0} c_{mn} = \sum_{\ell \geq 0}\sum_{m+n=\ell} c_{mn} = \sum_{\ell\geq 0}\sum_{m+n=\ell} a_m b_n. \qquad \square$$

3. Convergence, algebra, and continuity with power series

A *power series* over K is an infinite series of the form

$$f(x) = \sum_{n\geq 0} a_n x^n = \lim_{n\to\infty} (a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n)$$

with coefficients $a_n \in K$ and variable x in K. We will be concerned initially with describing the set of x where the power series converges and then performing operations on power series (*e.g.*, addition, multiplication, and differentiation). The series obviously converges if $x = 0$, with $f(0) = a_0$.

The simplest example of a power series (aside from polynomials, which are power series with all but finitely many coefficients equal to 0) is a geometric series $\sum_{n\geq 0} x^n$, whose convergence is easy to describe: the series converges if and only if $|x| < 1$, in which case

$$\sum_{n\geq 0} x^n = \frac{1}{1-x}.$$

This is just like the behavior of geometric series in $\mathbf{R}$ and does not depend on whether or not the absolute value is archimedean or nonarchimedean.

Theorem 3.1. *If $\sum_{n\geq 0} a_n x^n$ and $\sum_{n\geq 0} b_n x^n$ both converge for an $x \in K$, then*

$$\sum_{n\geq 0} a_n x^n + \sum_{n\geq 0} b_n x^n = \sum_{n\geq 0} (a_n + b_n) x^n,$$

$$\left(\sum_{m\geq 0} a_m x^m\right)\left(\sum_{n\geq 0} b_n x^n\right) = \sum_{\ell\geq 0}\left(\sum_{m=0}^{\ell} a_m b_{\ell-m}\right) x^\ell.$$

Proof. The formula for the sum follows from Theorem 2.2 with a_n and b_n there replaced by $a_n x^n$ and $b_n x^n$. The formula for the product follows from Corollary 2.11 with a_m and b_n there replaced by $a_m x^m$ and $b_n x^n$: if $m+n = \ell$ then $a_m x^m b_n x^n = a_m b_n x^\ell = a_m b_{\ell-m} x^\ell$. $\square$

The geometric series converges on the open unit disc. On what kind of set in K does a general power series converge? In calculus you learn that a real power series has a radius of convergence, usually found with the ratio test. But the ratio test is not always applicable (it just seems so in calculus courses because no other technique is available). There is a formula

for the radius of convergence R of every real power series $\sum_{n\geq 0} a_n x^n$, due to Cauchy [1, pp. 132, 143, 151] (1821) and rediscovered by Hadamard [3] (1888):

$$\frac{1}{R} = \varlimsup_{n\to\infty} \sqrt[n]{|a_n|},$$

where $1/\infty = 0$, $1/\infty = 0$, and $\varlimsup$ is defined next.

Definition 3.2. For a sequence $\{x_n\} \subset \mathbf{R}$, its *limit supremum* $\varlimsup_{n\to\infty} x_n$ is the largest limit point of the sequence $\{x_n\}$ in $\mathbf{R} \cup \{\pm\infty\}$.

Don't confuse this with the supremum $\sup x_n$, which is the least upper bound of the x_n's.

Example 3.3. If $x_n = 1 + 1/n$ for $n \geq 1$ then $\sup x_n = 2$ and $x_n \to 1$, so the sequence has only one limit point: $\varlimsup_{n\to\infty} x_n = \lim_{n\to\infty} x_n = 1$.

Example 3.4. If $x_n = (-1)^n/n$ for $n \geq 1$ then $\sup x_n = 1/2$ and $x_n \to 0$, so $\varlimsup_{n\to\infty} x_n = 0$.

Example 3.5. If $x_n = (-1)^n n/(n+1)$ for $n \geq 1$, then the sequence has two limit points: 1 and -1. So $\lim_{n\to\infty} x_n$ does not exist but $\varlimsup_{n\to\infty} x_n = 1$.

When $L := \varlimsup_{n\to\infty} x_n$ is finite, it is characterized by properties from above and below:

- for all $\varepsilon > 0$, $x_n < L + \varepsilon$ for $n \gg 0$ and
- for all $\varepsilon > 0$, $L - \varepsilon < x_n$ for *infinitely many* n (not for $n \gg 0$; see Example 3.5)

We will need two more important properties of $\varlimsup_{n\to\infty} x_n$:

- $\varlimsup_{n\to\infty} x_n = \infty$ if and only if a *subsequence* of $\{x_n\}$ tends to ∞ (it's not necessary to have $x_n \to \infty$).
- If $x_n \geq 0$ for all n, $\varlimsup_{n\to\infty} x_n = 0$ if and only if $x_n \to 0$ as $n \to \infty$ (it's not enough for a subsequence to tend to 0 – try $x_n = 1/n$ for even n and $x_n = 1$ for odd n).

Remark 3.6. We call $\varlimsup_{n\to\infty} x_n$ a limit supremum from its connection to suprema of the sequences $\{x_k, x_{k+1}, x_{k+2}, \ldots\}$ where more and more initial terms are omitted:

$$\sup_{n\geq 1} x_n \geq \sup_{n\geq 2} x_n \geq \sup_{n\geq 3} x_n \geq \cdots$$

and the sequence $s_k = \sup_{n\geq k} x_n$ in $\mathbf{R}$ has a limit (possibly being $\pm\infty$) since all monotonic sequences in $\mathbf{R}$ have a limit in $\mathbf{R} \cup \{\pm\infty\}$. The limit of the s_k's is precisely $\varlimsup_{n\to\infty} x_n$:

$$\varlimsup_{n\to\infty} x_n = \lim_{k\to\infty} \left(\sup_{n\geq k} x_n \right).$$

This is why $\varlimsup_{n\to\infty} x_n$ is also written $\limsup_{n\to\infty} x_n$ and is pronounced "lim sup x_n".[1] For a sequence $\{x_n\}$ in $\mathbf{R}$, its limit in $\mathbf{R} \cup \{\pm\infty\}$ may not exist but $\varlimsup_{n\to\infty} x_n$ always does.

The point of introducing $\varlimsup$ here is the Cauchy–Hadamard radius of convergence formula. It works not only over $\mathbf{R}$, but also over K when it uses the absolute value on K.

Theorem 3.7. *Let $f(x) = \sum_{n\geq 0} a_n x^n$ where $a_n \in K$. Define $R \in [0, \infty]$ by the formula*

$$\frac{1}{R} = \varlimsup_{n\to\infty} \sqrt[n]{|a_n|},$$

[1] In LaTeX, $\limsup$ = `\limsup` and $\varlimsup$ = `\varlimsup`.

where $1/0 = \infty$ *and* $1/\infty = 0$.[2] *The set of* $x \in K$ *where* $f(x)$ *converges is*

$$\{x \in K : |x| < R\} \quad \text{or} \quad \{x \in K : |x| \leq R\}.$$

Proof. First we will show the series converges if $|x| < R$ and not if $|x| > R$. We will break up the proof into three cases: (1) $R = 0$, (2) $R = \infty$, and $0 < R < \infty$.

Case 1: $R = 0$. We want to show $f(x)$ does not converge for $x \neq 0$ in K. From $R = 0$ we have $\overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|} = \infty$, so some subsequence of $\sqrt[n]{|a_n|}$ tends to ∞. For nonzero x in K we want to show $f(x)$ does not converge:

$$\begin{aligned} |x| > 0 &\implies \sqrt[n]{|a_n|} > \frac{1}{|x|} \text{ infinitely often} \\ &\implies |a_n| > \frac{1}{|x^n|} \text{ infinitely often} \\ &\implies |a_n x^n| > 1 \text{ infinitely often,} \end{aligned}$$

so $\sum_{n\geq 0} a_n x^n$ doesn't converge since the general term does not tend to 0.

Case 2: $R = \infty$. We want to show $f(x)$ converges for all x in K. From $R = \infty$ we have $\overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|} = 0$, so $\sqrt[n]{|a_n|} \to 0$. Convergence of $f(x)$ for $x = 0$ is obvious. For $x \in K - \{0\}$,

$$\begin{aligned} \sqrt[n]{|a_n|} < \frac{1}{2|x|} \text{ for } n \gg 0 &\implies |a_n| < \frac{1}{2^n |x^n|} \text{ for } n \gg 0 \\ &\implies |a_n x^n| < \frac{1}{2^n} \text{ for } n \gg 0. \end{aligned}$$

Therefore the convergence of $\sum 1/2^n$ in $\mathbf{R}$ implies convergence of $\sum a_n x^n$ in K by the comparison test (Theorem 2.3).

Case 3: $0 < R < \infty$. Convergence of $f(x)$ at $x = 0$ is obvious. For $0 < |x| < R$ we want to show $f(x)$ converges. Rewriting the inequalities as $0 < |x|/R < 1$, there's an ε in $(0, 1)$ for which $|x|/R < 1 - \varepsilon < 1$, so $1/R < (1 - \varepsilon)/|x|$. Thus

$$\overline{\lim_{n\to\infty}} \sqrt[n]{|a_n|} < \frac{1-\varepsilon}{|x|} \implies \sqrt[n]{|a_n|} < \frac{1-\varepsilon}{|x|} \text{ for } n \gg 0 \implies |a_n x^n| < (1-\varepsilon)^n \text{ for } n \gg 0.$$

Convergence of $\sum_{n\geq 0}(1-\varepsilon)^n$ in $\mathbf{R}$ implies convergence of $\sum_{n\geq 0} a_n x^n$ in K by the comparison test (Theorem 2.3).

For $|x| > R$ we show $f(x)$ does not converge. Since $1/|x| < 1/R = \overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|}$, we have $1/|x| < \sqrt[n]{|a_n|}$ infinitely often, so $1 < |a_n x^n|$ infinitely often, and thus $f(x)$ does not converge since its general term does not tend to 0.

If $x \in K$ and $|x| = R$, when does $f(x)$ converge? (If $R \notin |K^\times|$ then there is no such x, *e.g.*, if $K = \mathbf{Q}_p$ and $R \notin p^{\mathbf{Z}}$; we'll see an example of this later with the p-adic exponential series for $p \neq 2$.) Since $|\cdot|$ is nonarchimedean, convergence of $\sum_{n\geq 0} a_n x^n$ in K is equivalent to $|a_n x^n| = |a_n| R^n$ tending to 0, which depends on x only through $|x|$, so either $f(x)$ converges for all $x \in K$ with $|x| = R$ or $f(x)$ does not converge for all such x. □

If $R \in |K^\times|$, to determine whether or not the "circumference" $|x| = R$ is in the disc of convergence of $f(x)$ it suffices to check convergence at just one element of K with absolute value R.

[2]The numbers $\sqrt[n]{|a_n|}$ don't include a term for $n = 0$.

Example 3.8. Let $f(x) = \sum_{n\geq 0} x^n$. All coefficients are 1, so $R = 1$. At $x = 1$ the series does not converge, so the geometric series has disc of convergence $\{x \in K : |x| < 1\}$, which we already knew.

In our next three examples, let K be a complete extension field of $\mathbf{Q}_p$, so $|\cdot| = |\cdot|_p$ on the subfield $\mathbf{Q}_p$, and thus in particular on rational numbers in K. We call such K a *p-adic field*.

Example 3.9. Let $f(x) = \sum_{k\geq 0} p^k x^{p^k} = x + px^p + p^2x^{p^2} + \cdots$, so this series has nonzero coefficients only when the exponent is a power of p. Then

$$\frac{1}{R} = \varlimsup_{n\to\infty} \sqrt[n]{|a_n|} = \varlimsup_{k\to\infty} \sqrt[p^k]{|p^k|_p} = \varlimsup_{k\to\infty} \frac{1}{p^{k/p^k}}.$$

In $\mathbf{R}$, $k/p^k \to 0$ as $k \to \infty$, so $R = 1$. At $x = 1$ the series converges, so the disc of convergence is $\{x \in K : |x| \leq 1\}$.

Example 3.10. In $\mathbf{R}$, $\log(1 + x) = \sum_{n\geq 1}(-1)^{n-1}x^n/n$ converges for $-1 < x \leq 1$. How does this series behave in a p-adic field K? We have

$$\frac{1}{R} = \varlimsup_{n\to\infty} \sqrt[n]{\left|\frac{(-1)^{n-1}}{n}\right|} = \varlimsup_{n\to\infty} \frac{1}{\sqrt[n]{|n|_p}} = \varlimsup_{n\to\infty} p^{\mathrm{ord}_p(n)/n}.$$

Since $1 \leq p^{\mathrm{ord}_p(n)} \leq n$, we have $1 \leq p^{\mathrm{ord}_p(n)/n} \leq \sqrt[n]{n}$, and $\sqrt[n]{n} \to 1$ as $n \to \infty$, so $R = 1$. (In particular, $\sqrt[n]{|n|_p} \to 1$ as $n \to \infty$.) At $x = 1$ the terms $(-1)^{n-1}x^n/n = \pm 1/n$ have absolute value $1/|n|_p$, which does not tend to 0 (it is 1 when n is not divisible by p), so the series doesn't converge at 1 and therefore the disc of convergence of $\sum_{n\geq 1}(-1)^{n-1}x^n/n$ in K is the open unit disc $\{x \in K : |x| < 1\}$.

Example 3.11. In $\mathbf{R}$, $e^x = \sum_{n\geq 0} x^n/n!$ has an infinite radius of convergence, where the $n!$ in the denominator helps. But p-adically $n!$ is very small when n gets large, so having $n!$ in the denominator of the power series hurts the convergence. Might this series for $x \in K$ (ignoring its interpretation in $\mathbf{R}$ as $(2.718\ldots)^x$) have p-adic radius of convergence 0?

Letting R be the radius of convergence,

$$\frac{1}{R} = \varlimsup_{n\to\infty} \sqrt[n]{|1/n!|} = \varlimsup_{n\to\infty} \sqrt[n]{|1/n!|_p} = \varlimsup_{n\to\infty} p^{\mathrm{ord}_p(n!)/n}, \tag{3.1}$$

so we need to know the p-divisibility of $n!$. A formula for $\mathrm{ord}_p(n!)$ was given by Legendre two hundred years ago, and it can be described in two ways:

$$\mathrm{ord}_p(n!) = \sum_{k\geq 1} \left\lfloor \frac{n}{p^k} \right\rfloor = \frac{n - s_p(n)}{p-1}, \tag{3.2}$$

where $s_p(n)$ is the sum of the base p digits of n: if $n = a_0 + a_1p + \cdots + a_rp^r$ then $s_p(n) = a_0 + a_1 + \cdots + a_r$. The series in (3.2) is finite since the terms are zero once $p^k > n$. For example, if $n = p^2 + 1$ then the series is $\lfloor (p^2+1)/p \rfloor + \lfloor (p^2+1)/p^2 \rfloor = p + 1$ and $((p^2+1) - s_p(p^2+1))/(p-1) = (p^2+1-2)/(p-1) = p+1$: both say $\mathrm{ord}_p((p^2+1)!) = p+1$.

The formulas in (3.2) are clear if $n = 0$, since they are 0. To derive this formula for $n \geq 1$ count the highest power of p in $n! = 1 \cdot 2 \cdot 3 \cdot \cdots \cdot n$ by counting how many integers up to n are divisible by p, by p^2, by p^3, and so on. Integers divisible by p just once are counted only once, those divisible by p just twice are counted twice (they are first counted when

we count numbers divisible by p, and then they are counted again when we count numbers divisible by p^2), and so on. The number of integers up to n divisible by p^k is $\lfloor n/p^k \rfloor$, so

$$\mathrm{ord}_p(n!) = \sum_{k \geq 1} \left\lfloor \frac{n}{p^k} \right\rfloor.$$

Now we use the base p expansion $n = a_0 + a_1 p + \cdots + a_r p^r$ to write the sum in another way: $n/p^k < 1$ if $k > r$, so the sum in Legendre's formula runs from $k = 1$ to $k = r$. Since

$$\frac{n}{p^k} = \frac{a_0}{p^k} + \cdots + \frac{a_{k-1}}{p} + a_k + a_{k+1}p + \cdots + a_r p^{r-k},$$

we have $\lfloor n/p^k \rfloor = a_k + a_{k+1}p + \cdots + a_r p^{r-k}$. Therefore

$$\begin{aligned} \mathrm{ord}_p(n!) = {} & a_1 + a_2 p + a_3 p^2 + \cdots + a_r p^{r-1} + \\ & a_2 + a_3 p + \cdots + a_r p^{r-2} + \\ & a_3 + \cdots + a_r p^{r-3} + \ \cdots \end{aligned}$$

and summing terms vertically,

$$\begin{aligned} \mathrm{ord}_p(n!) &= a_1 + a_2(1+p) + a_3(1+p+p^2) + \cdots + a_r(1+p+\cdots+p^{r-1}) \\ &= a_1 \frac{p-1}{p-1} + a_2 \frac{p^2-1}{p-1} + a_3 \frac{p^3-1}{p-1} + \cdots + a_r \frac{p^r-1}{p-1} \\ &= \frac{(a_1 p + a_2 p^2 + \cdots + a_r p^r) - (a_1 + a_2 + \cdots + a_r)}{p-1} \\ &= \frac{(a_0 + a_1 p + a_2 p^2 + \cdots + a_r p^r) - (a_0 + a_1 + a_2 + \cdots + a_r)}{p-1} \\ &= \frac{n - s_p(n)}{p-1}. \end{aligned}$$

From the second formula for $\mathrm{ord}_p(n!)$ in (3.2), for $n \geq 1$ we have the upper bound[3]

$$\mathrm{ord}_p(n!) < \frac{n}{p-1}.$$

We will use the second formula in (3.2) to get a lower bound. The base p expansion is $n = a_0 + a_1 p + \cdots + a_r p^r$, with $a_r \neq 0$. Since $a_i \leq p-1$,

$$\mathrm{ord}_p(n!) = \frac{n - (a_0 + a_1 + a_2 + \cdots + a_r)}{p-1} \geq \frac{n - (r+1)(p-1)}{p-1} = \frac{n}{p-1} - (r+1).$$

To get a lower bound on this purely in terms of n, what is an upper bound on r in terms of n? Since $a_r \neq 0$ we have $p^r \leq n < p^{r+1}$, so $r \leq \log_p n < r+1$, whhere $\log_p$ is the classical base p logarithm. Thus

$$\mathrm{ord}_p(n!) \geq \frac{n}{p-1} - (\log_p n + 1),$$

so

$$\frac{n}{p-1} - (\log_p n + 1) \leq \mathrm{ord}_p(n!) < \frac{n}{p-1}.$$

Dividing by n,

$$\frac{1}{p-1} - \frac{\log_p n + 1}{n} \leq \frac{\mathrm{ord}_p(n!)}{n} < \frac{1}{p-1}.$$

[3]The first formula in (3.2) also gives this upper bound since $\mathrm{ord}_p(n!) < \sum_{k \geq 1} n/p^k$; sum the series.

Therefore $\mathrm{ord}_p(n!)/n \to 1/(p-1)$ as $n \to \infty$, so

$$\lim_{n\to\infty} p^{\mathrm{ord}_p(n!)/n} = p^{1/(p-1)},$$

which means the radius of convergence of $\sum_{n\geq 0} x^n/n!$ in K is

$$\frac{1}{p^{1/(p-1)}} = \left(\frac{1}{p}\right)^{1/(p-1)}.$$

The real number $(1/p)^{1/(p-1)}$ is quite special in p-adic analysis, showing up a lot. Let's observe that it lies strictly between $1/p$ and 1 for $p \neq 2$, while it is $1/2 = 1/p$ for $p = 2$.

We need to see if the disc of convergence includes $x \in K$ with $|x| = (1/p)^{1/(p-1)}$ (if there are such x in K, *e.g.*, there are for $K = \mathbf{Q}_2$ when $(1/p)^{1/(p-1)} = 1/2$). For such x,

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

where $s_p(n)$ is the sum of the base p digits of n. When n is a power of p this absolute value is $(1/p)^{1/(p-1)}$, which does not tend to 0. So the disc of convergence of $\sum_{n\geq 0} x^n/n!$ in K is $\{x \in K : |x| < (1/p)^{1/(p-1)}\}$.

For example, when $K = \mathbf{Q}_p$ the inequality $|x|_p < (1/p)^{1/(p-1)}$ is the same as $|x|_p \leq 1/p$ for $p \neq 2$ and $|x|_2 \leq 1/4$ for $p = 2$. Thus $\sum_{n\geq 0} x^n/n!$ in $\mathbf{Q}_p$ has disc of convergence $p\mathbf{Z}_p$ when $p \neq 2$ and $4\mathbf{Z}_2$ when $p = 2$.

It might seem strange to be fussy about the radius $(1/p)^{1/(p-1)}$ when it can be replaced with $1/p$ or $1/4$ on the p-adic numbers. The point is that $\mathbf{Q}_p$ can be enlarged to bigger complete fields (an analogue of going from $\mathbf{R}$ to $\mathbf{C}$, say) and in such fields there can be elements with absolute value between $1/p$ and $(1/p)^{1/(p-1)}$ when $p \neq 2$.

Theorem 3.12. *If $f(x) = \sum_{n\geq 0} a_n x^n$ converges on a disc $\{x \in K : |x| \leq r\}$ where $r \in |K^\times|$ then on this disc the power series is uniformly continuous and bounded.*

In this theorem we are not insisting that $\{x : |x| \leq r\}$ is the maximal disc of convergence for the series, just that the series converges on such a disc. It will be clear in the proof why it matters that r is the absolute value of something in K.

Proof. By hypothesis there is some $t \in K$ with $|t| = r$ and the series converges at t, so $|a_n t^n| = |a_n| r^n \to 0$ as $n \to \infty$. Thus $|x| \leq r \Longrightarrow |f(x)| \leq \max_{n\geq 0} |a_n||x|^n \leq \max_{n\geq 0} |a_n| r^n$, so f is bounded on $\{x \in K : |x| \leq r\}$.

For $|x| \leq r$ and $|y| \leq r$ in K,

$$\begin{aligned} f(x) - f(y) &= \sum_{n\geq 1} a_n(x^n - y^n) \\ &= \sum_{n\geq 1} a_n(x-y)(x^{n-1} + x^{n-2}y + \cdots + xy^{n-2} + y^{n-1}) \\ &= (x-y)\sum_{n\geq 1} a_n(x^{n-1} + x^{n-2}y + \cdots + xy^{n-2} + y^{n-1}). \end{aligned}$$

(The summation starts at $n = 1$ since the constant terms cancel in the difference.) From the strong triangle inequality,

$$|x^{n-1} + x^{n-2}y + \cdots + xy^{n-2} + y^{n-1}| \leq \max_{0\leq i\leq n-1} |x|^i |y|^{n-1-i} \leq r^{n-1},$$

so

$$\left|\sum_{n\geq 1} a_n(x^{n-1} + x^{n-2}y + \cdots + xy^{n-2} + y^{n-1})\right| \leq \max_{n\geq 1} |a_n|r^{n-1} = \frac{1}{r}\max_{n\geq 1}|a_n|r^n.$$

This maximum exists since $|a_n|r^n \to 0$. Thus

$$|f(x) - f(y)| \leq |x - y| \max_{n\geq 1} |a_n|r^{n-1} = A_r|x - y|$$

where $A_r = \max_{n\geq 1} |a_n|r^{n-1}$.

To show f is uniformly continuous on $\{x \in K : |x| \leq r\}$, we consider two cases.

(1) If $a_n = 0$ for all $n \geq 1$, so $f(x)$ is the constant function a_0, it is obviously uniformly continuous.
(2) If some a_n for $n \geq 1$ is not 0, then $A_r > 0$, so for $\varepsilon > 0$ setting $\delta = \varepsilon/A_r$ shows $|x - y| \leq \delta \Longrightarrow |f(x) - f(y)| \leq A_r|x - y| \leq A_r\delta = \varepsilon$. □

Corollary 3.13. *If $f(x) = \sum_{n\geq 0} a_n x^n$ converges on a disc $\{x \in K : |x| \leq r\}$ where $r \in |K^\times|$ then there is a number A_r such that $|f(x) - a_0| \leq A_r|x|$ when $|x| \leq r$.*

Proof. Set $y = 0$ in Theorem 3.12. □

Corollary 3.14. *A power series $\sum_{n\geq 0} a_n x^n$ with coefficients in K is continuous on its disc of convergence in K.*

This is not a restatement of the continuity from Theorem 3.12 since the disc of convergence might be of the form $\{x \in K : |x| < R\}$ rather than of the form $\{x \in K : |x| \leq R\}$.

Proof. If the radius of convergence is 0 then the series only converges at $x = 0$ and the desired result is obvious (and boring). If the radius of convergence is positive and x_0 is a number in $K^\times$ at which the power series converges, then the series converges on the closed disc $\{x \in K : |x| \leq r\}$ where $r = |x_0| > 0$. By Theorem 3.12 the series is uniformly continuous on this disc, which is an open subset of K, and thus in particular the series is continuous at x_0 and at 0. As x_0 was arbitrary in the disc of convergence, we are done. □

Corollary 3.15. *A power series with a positive radius of convergence has only one choice of coefficients: if $\sum_{n\geq 0} a_n x^n = \sum_{n\geq 0} b_n x^n$ for all x in a disc of positive radius then $a_n = b_n$ for all n.*

Proof. Using the power series $\sum c_n x^n := \sum_{n\geq 0}(a_n - b_n)x^n$, we will show that if $\sum c_n x^n = 0$ for all small x in K then $c_n = 0$ for all n.

Setting $x = 0$ in the power series, we get $c_0 = 0$. Suppose $c_n = 0$ for $n < N$, so $\sum_{n\geq N} c_n x^n = 0$ for all small x in K. We want to show $c_N = 0$. For small nonzero x in K, divide by x:

$$\sum_{n\geq N} c_n x^n = 0 \Longrightarrow c_N x^N + c_{N+1}x^{N+1} + c_{N+2}x^{N+2} + \cdots = 0$$

$$\Longrightarrow c_N + c_{N+1}x + c_{N+2}x^2 + \cdots = 0 \text{ since } x \neq 0.$$

Thus $\sum_{n\geq 0} c_{N+n}x^n = 0$ for small nonzero x in K. This power series also converges at $x = 0$ with value c_N. A power series is continuous on its disc of convergence, so $c_N = \lim_{x\to 0} \sum_{n\geq 0} c_{N+n}x^n = \lim_{x\to 0} 0 = 0$. □

In real analysis, the usual proof that a power series $f(x) = \sum a_n x^n$ has unique coefficients uses Taylor's formula $a_n = f^{(n)}(0)/n!$. We'll give such a proof later in Corollary 5.6.

4. The p-adic exponential series

In this section, K is a p-adic field: a complete extension field of $\mathbf{Q}_p$.

We saw in Example 3.11 that the classical exponential series $\sum_{n\geq 0} x^n/n!$ converges in K only at those x with $|x| < (1/p)^{1/(p-1)}$.

Definition 4.1. For $x \in K$ with $|x| < (1/p)^{1/(p-1)}$, the *p-adic exponential* at x is

$$e^x = \exp(x) = \sum_{n\geq 0} \frac{x^n}{n!}.$$

We determined the disc of convergence of e^x in K as part of Example 3.11: it is the disc

$$D_p = D_p(K) = \{x \in K : |x| < (1/p)^{1/(p-1)}\}. \tag{4.1}$$

Example 4.2. On $\mathbf{Q}_p$, the p-adic exponential series is defined on

$$D_p(\mathbf{Q}_p) = \begin{cases} p\mathbf{Z}_p, & \text{if } p \neq 2, \\ 4\mathbf{Z}_2, & \text{if } p = 2. \end{cases}$$

Since 1 is not in D_p, the series e^x does not converge at $x = 1$: we denote the series as e^x by analogy with the real exponential series, but *there is no p-adic number e*. The number e^p for $p \neq 2$ (resp., e^4 for $p = 2$) exists in $\mathbf{Q}_p$, but it has no pth root (resp., no fourth root when $p = 2$) in $\mathbf{Q}_p$. Even if we pass to a larger field than $\mathbf{Q}_p$ in which such a root can be found, that number (after deciding *which* pth root or 4th root should be preferred!) has no known significance in p-adic analysis. The bottom line is that specific transcendental real numbers can't be productively interpreted as p-adic numbers.

Example 4.3. In $\mathbf{Q}_2$ let's compute e^4 mod 2^7. By definition $e^4 = \sum_{n\geq 0} 4^n/n!$, and the nth term tends to 0 as n grows. How large must n be so that $|4^n/n!|_2 \leq 1/2^7$, or equivalently $\mathrm{ord}_2(4^n/n!) \geq 7$? Since

$$\mathrm{ord}_2\left(\frac{4^n}{n!}\right) = 2n - \mathrm{ord}_2(n!) = 2n - (n - s_2(n)) = n + s_2(n),$$

and $s_2(n) \geq 1$ for $n \geq 1$, we have $\mathrm{ord}_2(4^n/n!) \geq 7$ if $n \geq 6$. The table below displays $\mathrm{ord}_2(4^n/n!)$ for smaller n.

n	0	1	2	3	4	5
$s_2(n)$	0	1	1	2	1	2
$n + s_2(n)$	0	2	3	5	5	7

From the table $\mathrm{ord}_2(4^n/n!) \geq 7$ for $n \geq 5$, so $4^n/n! \in 2^7\mathbf{Z}_2$ for $n \geq 5$. Thus

$$e^4 = 1 + 4 + \frac{4^2}{2!} + \frac{4^3}{3!} + \frac{4^4}{4!} + \text{element of } 2^7\mathbf{Z}_2.$$

The sum of the first five terms is $103/3 \equiv 77 \bmod 2^7$, so $e^4 \equiv 77 \equiv 1011001 \bmod 2^7$.

Theorem 4.4. *For x and y in D_p, $e^x e^y = e^{x+y}$.*

Since D_p is a disc centered at 0 in a nonarchimedean field it is an additive group, so $x + y \in D_p$ when x and y are in D_p.

Proof. We use Corollary 2.11:

$$\begin{aligned} e^x e^y &= \sum_{m\geq 0} \frac{x^m}{m!} \sum_{n\geq 0} \frac{y^n}{n!} \\ &= \sum_{\ell \geq 0} \sum_{m+n=\ell} \frac{1}{m!}\frac{1}{n!} x^m y^n \\ &= \sum_{\ell \geq 0} \frac{1}{\ell!} \sum_{m+n=\ell} \frac{\ell!}{m!n!} x^m y^n \\ &= \sum_{\ell \geq 0} \frac{1}{\ell!} \sum_{m=0}^{\ell} \frac{\ell!}{m!(\ell-m)!} x^m y^{\ell-m} \\ &= \sum_{\ell \geq 0} \frac{1}{\ell!} (x+y)^\ell \\ &= e^{x+y}. \end{aligned}$$

□

Since $e^x e^{-x} = e^0 = 1$, the p-adic exponential series on K does not take the value 0, just like the classical exponential series.

A striking contrast with real analysis is that the p-adic exponential series has a finite radius of convergence, as we saw. A more striking contrast is that it preserves distances!

Theorem 4.5. *If $t \in D_p$ then $|e^t - 1| = |t|$. For x and y in D_p, $|e^x - e^y| = |x - y|$.*

Proof. If $t = 0$ then $e^t - 1 = 0$, so we can assume $0 < |t| < (1/p)^{1/(p-1)}$. To show $|e^t - 1| = |t|$, extract the first two terms in the power series for e^t:

$$e^t = 1 + t + \sum_{n\geq 2} \frac{t^n}{n!} \Longrightarrow e^t - 1 = t + \sum_{n\geq 2} \frac{t^n}{n!}.$$

We will show $|t^n/n!| < |t|$ for $n \geq 2$. Then for finite $N \geq 2$ we have $|\sum_{n=1}^N t^n/n!| = |t|$, and letting $N \to \infty$ we get $|e^t - 1| = |t|$. For $t \neq 0$, the condition $|t^n/n!| < |t|$ is equivalent to $|t|^{n-1} < |n!| = |n!|_p$, which is the same as $|t| < |n!|_p^{1/(n-1)}$, so for $|t| < (1/p)^{1/(p-1)}$ it suffices to show $(1/p)^{1/(p-1)} \leq |n!|_p^{1/(n-1)}$:

$$\begin{aligned} \left(\frac{1}{p}\right)^{1/(p-1)} \leq |n!|_p^{1/(n-1)} &\iff \left(\frac{1}{p}\right)^{1/(p-1)} \leq \left(\frac{1}{p}\right)^{(n-s_p(n))/(p-1)(n-1)} \\ &\iff p^{1/(p-1)} \geq p^{(n-s_p(n))/(p-1)(n-1)} \\ &\iff 1 \geq \frac{n - s_p(n)}{n-1}, \\ &\iff s_p(n) \geq 1, \end{aligned}$$

and this last inequality is true for $n \geq 2$.

For $|t| < (1/p)^{1/(p-1)}$, $|e^t| = |e^t - 1 + 1| = 1$ since $|e^t - 1| = |t| < 1$, so for $x, y \in D_p$,

$$|e^x - e^y| = |e^y(e^{x-y} - 1)| = |e^y||e^{x-y} - 1| = 1|e^{x-y} - 1| = |x - y|$$

using $t = x - y \in D_p$. □

Corollary 4.6. *If $e^x = e^y$ for $x, y \in D_p$ then $x = y$.*

Proof. If $e^x = e^y$ then $0 = |e^x - e^y| = |x - y|$, so $x = y$. □

Since the e^x preserves distances and sends 0 to 1, it sends D_p to $1 + D_p$. In particular,

- for $p \neq 2$, $x \in p\mathbf{Z}_p \Longrightarrow e^x \in 1 + p\mathbf{Z}_p$,
- $x \in 4\mathbf{Z}_2 \Longrightarrow e^x \in 1 + 4\mathbf{Z}_2$.

The set $1 + D_p$ (which for $K = \mathbf{Q}_p$ is $1 + p\mathbf{Z}_p$ for $p \neq 2$ and $1 + 4\mathbf{Z}_2$ for $p = 2$) is a multiplicative group, as is every set of the form $\{x \in K : |x - 1| < r\}$ where $r \in (0, 1)$, so e^x is a homomorphism from the additive group D_p to the multiplicative group $1 + D_p$.

5. Differentiability of power series

As in real analysis, the derivative of a function $f\colon U \to K$ for an open subset U of K is defined as a limit of Newton quotients: for $x \in U$,

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h},$$

where the limit is over h in K tending to 0. Familiar differentiation rules in $\mathbf{R}$ work in K.

Theorem 5.1. *Let U be an open subset of K and $f, g\colon U \to K$ be differentiable functions. Then $f + g$ and fg are differentiable at each $x \in U$, with*

$$(f+g)'(x) = f'(x) + g'(x), \quad (fg)'(x) = f(x)g'(x) + f'(x)g(x).$$

At each $x \in U$ where $g(x) \neq 0$, f/g is differentiable with

$$(f/g)'(x) = \frac{g(x)f'(x) - f(x)g'(x)}{g(x)^2}.$$

If $f\colon U \to K$ and $g\colon V \to K$ are differentiable functions with $g(V) \subset U$, then $f \circ g\colon V \to K$ is differentiable with derivative at each $x \in V$ given by

$$(f \circ g)'(x) = f'(g(x))g'(x).$$

Proof. Exercise. □

Example 5.2. For $c \in K$, if $g(x) = f(x + c)$ then $g'(x) = f'(x + c)$. If $g(x) = f(cx)$ then $g'(x) = cf(cx)$. These are very simple uses of the chain rule that could be checked directly.

Theorem 5.3. *If $f(x) = \sum_{n \geq 0} a_n x^n$ has a positive radius of convergence in K, then for each x in the disc of convergence we have $f'(x) = \sum_{n \geq 1} n a_n x^{n-1}$.*

Proof. Let R be the radius of convergence (possibly ∞). The disc of convergence for the series is filled up by discs of the form $\{x \in K : |x| \leq r\}$ for $0 < r < R$, where $r \in |K^\times|$, and possibly also for $r = R$ if $R \in |K^\times|$, so it suffices to prove the derivative formula on discs where f converges that have the form $\{x \in K : |x| \leq r\}$ with $r \in |K^\times|$.

Fix $x \in K$ with $|x| \leq r$. For $0 < |h| \leq r$ we have $|x + h| \leq \max(|x|, |h|) \leq r$, so $f(x + h)$ is defined and

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

Each h appears in h^m with $m \geq 1$, so dividing through by h gives us

$$\frac{f(x+h)-f(x)}{h} = \sum_{n\geq 1} a_n \left(\sum_{m=1}^{n} \binom{n}{m} h^{m-1} x^{n-m} \right) = \sum_{n\geq 1} \sum_{m \geq 1} c_{mn},$$

where

$$c_{mn} = \begin{cases} \binom{n}{m} h^{m-1} a_n x^{n-m}, & \text{if } 1 \leq m \leq n, \\ 0, & \text{if } m > n. \end{cases}$$

We want to change the order of the double series. To justify this by Theorem 2.9 we check $c_{mn} \to 0$ in K as $\max(m,n) \to \infty$. If $m > n$ then c_{mn} is 0. If $m \leq n$ then since $|\binom{n}{m}| = |\binom{n}{m}|_p \leq 1$,

$$|c_{mn}| \leq \left| \binom{n}{m} \right| |h|^{m-1} |a_n| |x|^{n-m} \leq r^{m-1} |a_n| r^{n-m} = |a_n| r^{n-1} = \frac{1}{r} |a_n| r^n,$$

which tends to 0 as $n \to \infty$ since the power series converges at radius r. Thus $c_{mn} \to 0$ as $\max(m,n) \to \infty$, so we can swap the order of the double series:

$$\begin{aligned} \frac{f(x+h)-f(x)}{h} &= \sum_{n\geq 1} \sum_{m \geq 1} c_{mn} \\ &= \sum_{m\geq 1} \sum_{n \geq 1} c_{mn} \\ &= \sum_{m\geq 1} \sum_{n \geq m} \binom{n}{m} h^{m-1} a_n x^{n-m} \\ &= \sum_{m\geq 1} \left(\sum_{n \geq m} \binom{n}{m} a_n x^{n-m} \right) h^{m-1} \\ &= \sum_{n\geq 1} n a_n x^{n-1} + \left(\sum_{n \geq 2} \binom{n}{2} a_n x^{n-2} \right) h + \dots, \end{aligned}$$

and this is a power series in h. We assumed $0 < |h| \leq r$, so this series converges for $|h| \leq r$ (it obviously converges at $h = 0$). Since a power series is continuous on its disc of convergence (Corollary 3.14), as $h \to 0$ the power series in h tends to its constant term $\sum_{n\geq 1} n a_n x^{n-1}$. Thus $f'(x)$ exists and equals the termwise derivative $\sum_{n\geq 1} n a_n x^{n-1}$. □

Remark 5.4. While differentiability implies continuity, just like in real analysis, our proof of the differentiability of power series relied on already knowing power series are continuous functions, at least at the origin.

Example 5.5. As in $\mathbf{R}$, $(e^x)' = e^x$ for x in the disc of convergence of the exponential series in K.

Write the higher derivatives of a function as f'', f''', and $f^{(n)}$, as in real analysis.

Corollary 5.6. *If $f(x) = \sum_{n\geq 0} a_n x^n$ has a positive radius of convergence in K then f is infinitely differentiable on its disc of convergence in K and $a_n = f^{(n)}(0)/n!$ for all $n \geq 0$. In particular, f has only one choice of power series coefficients.*

The uniqueness of power series coefficients was proved earlier without derivatives in Corollary 3.15.

Proof. Let D be the disc of convergence of $f(x)$. By Theorem 5.3, $f'(x) = \sum_{n\geq 1} na_n x^{n-1}$ for all $x \in D$. This is again a power series, so by Theorem 5.3 it is differentiable on D with derivative $f''(x) = \sum_{n\geq 2} n(n-1)a_n x^{n-2}$. By induction, the kth derivative is

$$f^{(k)}(x) = \sum_{n\geq k} n(n-1)\cdots(n-(k-1))a_n x^{n-k}$$

for all $x \in D$. In particular, the constant term is $f^{(k)}(0) = k(k-1)\cdots(k-(k-1))a_k = k!a_k$, so $a_k = f^{(k)}(0)/k!$. This shows the coefficients of the power series are determined by f as a function around 0. □

Remark 5.7. Since $\sqrt[n]{|n|_p} \to 1$ (see Example 3.10), the derivative of a p-adic power series has the same radius of convergence as the original series, just like in real analysis. However, while in real analysis the derivative of a power series can converge on a smaller set than the original series – $\sum_{n\geq 1} x^n/n^2$ has interval of convergence $-1 \leq x \leq 1$ while its derivative $\sum_{n\geq 1} x^{n-1}/n$ has interval of convergence $-1 \leq x < 1$ – in p-adic analysis the derivative of a power series might converge on a larger set than the original series: on $\mathbf{Q}_p$, $\sum_{k\geq 0} x^{p^k}$ has disc of convergence $p\mathbf{Z}_p$ while its derivative $\sum_{k\geq 0} p^k x^{p^k-1}$ has disc of convergence $\mathbf{Z}_p$.

Corollary 5.8. *If two power series on a disc of positive radius in K centered at the origin have the same derivative on the disc then they differ by a constant on that disc.*

Proof. Let $f(x) = \sum_{n\geq 0} a_n x^n$ and $g(x) = \sum_{n\geq 0} b_n x^n$ converge on a common disc D around 0 with positive radius. If $f'(x) = g'(x)$ on D then $f^{(n)}(x) = g^{(n)}(x)$ on D for all $n \geq 1$. Therefore when $n \geq 1$, $a_n = f^{(n)}(0)/n! = g^{(n)}(0)/n! = b_n$, so

$$f(x) = a_0 + \sum_{n\geq 1} a_n x^n = a_0 + (g(x) - b_0) = g(x) + (a_0 - b_0),$$

so f and g differ by a constant on D. □

As a special case of Corollary 5.8, if a power series on a disc centered at 0 in K has derivative 0 then the power series is a constant function (all higher-degree coefficients are 0). In calculus, the property "derivative $= 0 \Rightarrow$ constant" for functions on an interval is proved using the Mean Value Theorem, without assuming the functions are representable as power series. There is no Mean Value Theorem in p-adic analysis, and in fact p-adic functions can have derivative 0 on $\mathbf{Z}_p$ without being constant functions.

Example 5.9. Let $f\colon \mathbf{Z}_p \to \mathbf{Q}_p$ by

$$f(x) = \begin{cases} 0, & \text{if } x \in p\mathbf{Z}_p, \\ 1, & \text{if } x \in \mathbf{Z}_p^\times. \end{cases}$$

This function is not constant but it is locally constant since $p\mathbf{Z}_p$ and $\mathbf{Z}_p^\times$ are both open in $\mathbf{Q}_p$: for each $x \in \mathbf{Z}_p$ the disc $x + p\mathbf{Z}_p = \{y \in \mathbf{Z}_p : |y - x|_p < 1\}$ is in $p\mathbf{Z}_p$ if x is in $p\mathbf{Z}_p$ and it is in $\mathbf{Z}_p^\times$ if x is in $\mathbf{Z}_p^\times$. Since f is locally constant on $\mathbf{Z}_p$, $f'(x) = 0$.

You might think the right correction to the failure of "derivative $= 0 \Rightarrow$ constant" in p-adic analysis should be "derivative $= 0 \Rightarrow$ locally constant", but even that has counterexamples, like the following.

Example 5.10. Let $f\colon \mathbf{Z}_p \to \mathbf{Q}_p$ by

$$f\left(\sum_{n\geq 0} a_n p^n\right) = \sum_{n\geq 0} a_n p^{2n}$$

where $a_n \in \{0, 1, \ldots, p-1\}$. This function, which spreads apart p-adic digits, is not constant or even locally constant: if $|x-y|_p = 1/p^m$ then $|f(x) - f(y)|_p = 1/p^{2m}$, so

$$|f(x) - f(y)|_p = |x-y|_p^2$$

for all x and y in $\mathbf{Z}_p$ (it is obvious if $x = y$ and we just calculated it for $x \neq y$). Thus $|f(x+h) - f(x)|_p = |h|_p^2$, so $|(f(x+h) - f(x))/h|_p = |h|_p$. Letting $h \to 0$ implies $f'(x) = 0$ for all $x \in \mathbf{Z}_p$ even though f is not locally constant anywhere on $\mathbf{Z}_p$.

6. A delicate p-adic proof with π

When a geometric series converges, the formula for its value is "universal" in all fields:

$$|x| < 1 \Longrightarrow \sum_{n\geq 0} x^n = \frac{1}{1-x}.$$

For example, $2/3$ has absolute value less than 1 in $\mathbf{R}$ and in $\mathbf{Q}_2$, so

$$\sum_{n\geq 0} \left(\frac{2}{3}\right)^n = \frac{1}{1-2/3} = 3 \text{ in } \mathbf{R} \text{ and } \mathbf{Q}_2.$$

The termwise differentiation of series in Theorem 5.3 looks the same in real and p-adic analysis. For example, if $|x| < 1$ then

$$\sum_{n\geq 0} x^n = \frac{1}{1-x} \Longrightarrow \sum_{n\geq 1} nx^{n-1} = \frac{1}{(1-x)^2} \Longrightarrow \sum_{n\geq 1} nx^n = \frac{x}{(1-x)^2}.$$

Thus

$$\sum_{n\geq 1} \frac{n2^n}{3^n} = \frac{2/3}{(1-2/3)^2} = 6 \text{ in } \mathbf{R} \text{ and } \mathbf{Q}_2.$$

Such "universal" series formulas can give the impression that when an infinite series of rational numbers converges to a rational number in $\mathbf{R}$ and in some $\mathbf{Q}_p$, the rational number in both cases has to be the same. In that light, consider the following p-adic proof of the irrationality of π.

Theorem 6.1. *The real number π is irrational.*

Proof. Assume π is rational, so we can write π in reduced form as a/b where a and b are relatively prime positive integers. Since we know $\pi > 3$, π is not of the form $1/b$ or $2/b$, so $a \geq 3$. Thus a has an odd prime factor or a is divisible by 4, so either $\pi \in p\mathbf{Z}_p$ for some odd prime p or $\pi \in 4\mathbf{Z}_2$.

The way we will get information about the real number π is from the relation $\sin(\pi) = 0$:

$$\sum_{k\geq 0} (-1)^k \frac{\pi^{2k+1}}{(2k+1)!} = 0. \tag{6.1}$$

Assuming π is rational, the series on the left side is an infinite series of rational numbers, and its value is 0.

Consider now the sine series of a p-adic variable: $\sum_{k\geq 0}(-1)^k x^{2k+1}/(2k+1)!$ for $x \in \mathbf{Q}_p$. The proof that the exponential series converges on $p\mathbf{Z}_p$ for $p \neq 2$ and on $4\mathbf{Z}_2$ for $p = 2$ applies also to the p-adic sine series (whose terms are just the odd-degree terms in the series e^x, up to sign). If π is rational then it belongs to $p\mathbf{Z}_p$ for an odd prime p or to $4\mathbf{Z}_2$, so there is a prime p for which the series on the left side of (6.1) converges in $\mathbf{Q}_p$ and then we can interpret (6.1) as an equation in $\mathbf{Q}_p$ for that p.

Since $\sin x$ has linear term x, just like $e^x - 1$, the proof of the first part of Theorem 4.5 carries over to the sine series: $|\sin x|_p = |x|_p$ for $x \in p\mathbf{Z}_p$ for odd p or $x \in 4\mathbf{Z}_2$ for $p = 2$. Therefore the equation $\sin(\pi) = 0$ in $\mathbf{Q}_p$ implies $0 = |\sin(\pi)|_p = |\pi|_p$, so $\pi = 0$. This is a contradiction, so π is irrational. □

That was slick! Unfortunately it is also wrong. The error is interpreting (6.1) as an equation in some $\mathbf{Q}_p$ just because it is valid in $\mathbf{R}$ and the left side makes sense in $\mathbf{Q}_p$. If $r_n \in \mathbf{Q}$ and $\sum_{n\geq 1} r_n$ converges in both $\mathbf{R}$ and some $\mathbf{Q}_p$ with a rational value in $\mathbf{R}$, this does not imply the value in $\mathbf{Q}_p$ is the same number.

Example 6.2. Pick a prime p and consider the rational sequence $a_n = 1/(1+p^n)$ for $n \geq 1$. We have $a_n \to 0$ in $\mathbf{R}$ and $a_n \to 1$ in $\mathbf{Q}_p$. Set $r_1 = a_1 = 1/(1+p)$ and $r_n = a_n - a_{n-1}$ for $n \geq 2$, so $r_1 + r_2 + \cdots + r_n = a_n$. Thus $\sum_{n\geq 1} r_n = 0$ in $\mathbf{R}$ and $\sum_{n\geq 1} r_n = 1$ in $\mathbf{Q}_p$.

Example 9.5 is more concrete example of this phenomenon: the same infinite series of rational numbers can converge to $8/7$ in $\mathbf{R}$ and to $-8/7$ in $\mathbf{Q}_3$.

In 1873, Hermite proved e is transcendental by an argument using integrals. In 1905 Hensel [4, pp. 555–557] devised the following incorrect p-adic proof of that fact.

Theorem 6.3. *The number e is transcendental.*

Proof. The argument will be by contradiction, and requires some familiarity with field theory and the Eisenstein irreducibility criterion.

Assume e is algebraic over $\mathbf{Q}$, say of degree n. Then for each prime p, e has degree at most n over $\mathbf{Q}_p$. We are going to show (incorrectly) that e has degree p over $\mathbf{Q}_p$ when $p > 2$. Therefore $p = [\mathbf{Q}_p(e) : \mathbf{Q}_p] \leq n$, so we get a contradiction by choosing $p > n$, which we can do since there are infinitely many primes.

For $p > 2$, $e^p = \sum_{n\geq 0} p^n/n! = 1 + p\sum_{n\geq 1} p^{n-1}/n!$ and $p^{n-1}/n!$ is 1 at $n = 1$ and it is in $p\mathbf{Z}_p$ for $n \geq 2$ (the reader should check that), so $e^p = 1 + pu$ where $u \in 1 + p\mathbf{Z}_p \subset \mathbf{Z}_p^\times$. That shows e is a root of $X^p - (1 + pu) \in \mathbf{Q}_p[X]$, so $e - 1$ is a root of

$$(X+1)^p - (1+pu) = X^p + \sum_{k=1}^{p-1} \binom{p}{k} X^k - pu, \tag{6.2}$$

which is an Eisenstein polynomial with respect to p (its constant term is divisible by p exactly once). The proof of the Eisenstein irreducibility criterion at p in $\mathbf{Q}[X]$ works in the same way in $\mathbf{Q}_p[X]$ (only for the prime p, not for other primes). Therefore (6.2) is irreducible over $\mathbf{Q}_p$, so $[\mathbf{Q}_p(e) : \mathbf{Q}_p] = [\mathbf{Q}_p(e-1) : \mathbf{Q}_p] = p$. Now we get a contradiction by using $p > n$. □

The error in the above proof is confusing the meaning of e^p in $\mathbf{R}$ and in $\mathbf{Q}_p$. They both are written as $\sum_{n\geq 0} p^n/n!$, so they appear to be the same number, but that is a mistake: the real and p-adic limits of the finite sums $\sum_{n=0}^N p^n/n!$ as $N \to \infty$ have no reason to be equal to each other in any real sense (pun intended).

7. General power series and analytic functions

So far all of our power series have been centered at the origin. We can also consider power series centered at other numbers: for $\alpha \in K$, a *power series centered at* α is an infinite series

$$f(x) = \sum_{n \geq 0} a_n (x - \alpha)^n. \tag{7.1}$$

Theorem 7.1. *Every power series $f(x)$ in (7.1) has the following properties.*

1. *It has a disc of convergence $\{x \in K : |x - \alpha| < R\}$ or $\{x \in K : |x - \alpha| \leq R\}$, where R is described by the Cauchy–Hadamard formula in Theorem 3.7 using the coefficients a_n from (7.1).*
2. *When $R > 0$, the series is continuous on its disc of convergence and uniformly continuous on $\{x \in K : |x - \alpha| \leq r\}$ inside the disc of convergence when $r \in |K^\times|$.*
3. *When $R > 0$, the series can be differentiated termwise and is infinitely differentiable, with coefficients given by Taylor's formula at α: $a_n = f^{(n)}(\alpha)/n!$.*
4. *Two power series centered at α that converge on a common disc centered at α and have the same derivative on that disc differ on that disc by a constant.*

Proof. Since $f(x + \alpha)$ is a power series centered at 0, proofs reduce to the case $\alpha = 0$, which we have already worked out: (1) is Theorem 3.7, (2) is Theorem 3.12 and Corollary 3.14, (3) is Theorem 5.3 and Corollary 5.6, and (4) is Corollary 5.8. □

The third property above says a power series centered at α has only one choice of coefficients, based on Taylor's formula using repeated derivatives. We next show the uniqueness of the coefficients can be proved by a method that makes *no use* of differentiation.

Theorem 7.2. *If D is a disc in K with positive radius and $f : D \to K$ is a function on D expressible as a power series centered at $\alpha \in D$, then it is so expressible in only one way: if $\sum_{n \geq 0} a_n (x - \alpha)^n = \sum_{n \geq 0} b_n (x - \alpha)^n$ for all $x \in D$ then $a_n = b_n$ for all n.*

Proof. Setting $x = \alpha$ in both series we get $a_0 = b_0$. Assume for some $N \geq 1$ that $a_n = b_n$ for $0 \leq n \leq N - 1$. We want to show $a_N = b_N$. For all $x \in D$,

$$\begin{aligned} \sum_{n \geq 0} a_n (x - \alpha)^n = \sum_{n \geq 0} b_n (x - \alpha)^n &\implies \sum_{n \geq N} a_n (x - \alpha)^n = \sum_{n \geq N} b_n (x - \alpha)^n \\ &\implies \sum_{n \geq N} (a_n - b_n)(x - \alpha)^n = 0 \\ &\implies (x - \alpha)^N \sum_{n \geq N} (a_n - b_n)(x - \alpha)^{n - N} = 0. \end{aligned} \tag{7.2}$$

The series $\sum_{n \geq N} (a_n - b_n)(x - \alpha)^{n-N} = (a_N - b_N) + (a_{N+1} - b_{N+1})(x - \alpha) + \cdots$ converges because it's obvious for $x = \alpha$ while for $x \neq \alpha$ in D and $n > N$,

$$|(a_n - b_n)(x - \alpha)^n| \to 0 \implies \frac{|(a_n - b_n)(x - \alpha)^n|}{|x - \alpha|^N} \to 0 \implies |(a_n - b_n)(x - \alpha)^{n-N}| \to 0.$$

When $x \neq \alpha$ in D, canceling the factor $(x - \alpha)^N$ in (7.2) tells us that

$$\sum_{n \geq N} (a_n - b_n)(x - \alpha)^{n - N} = 0 \text{ for } x \neq \alpha.$$

This is a convergent power series on D that vanishes on $D - \{\alpha\}$. A power series is continuous on its disc of convergence, so the power series must vanish at α too, where its value is $a_N - b_N$. Therefore $a_N = b_N$. □

Allowing power series centered at numbers besides 0 opens up an important way to study the local behavior of a power series: recentering.

Example 7.3. In $\mathbf{R}$, $\sum_{n\geq 0} x^n = 1/(1-x)$ for $|x| < 1$. For $|a| < 1$, if $|x-a| < |1-a|$ then

$$\frac{1}{1-x} = \frac{1}{1-a}\frac{1}{1-(x-a)/(1-a)} = \frac{1}{1-a}\sum_{n\geq 0}\left(\frac{x-a}{1-a}\right)^n = \sum_{n\geq 0}\frac{1}{(1-a)^{n+1}}(x-a)^n,$$

which has interval of convergence $\{x \in \mathbf{R} : |x-a| < |1-a|\}$ and radius of convergence $|1-a|$. For example, the power series for $1/(1-x)$ centered at $a = -1/2$ has interval of convergence $(-2, 1)$, not $(-1, 1)$ like the power series for $1/(1-x)$ centered at $a = 0$.

Example 7.4. In K, $\sum_{n\geq 0} x^n = 1/(1-x)$ for $|x| < 1$. For $|a| < 1$, if $|x-a| < |1-a|$ then

$$\frac{1}{1-x} = \sum_{n\geq 0}\frac{1}{(1-a)^{n+1}}(x-a)^n$$

by the same reasoning as in Example 7.3. The disc of convergence of this power series is $\{x \in K : |x-a| < |1-a|\}$. Since $|a| < 1$, $|1-a| = 1$ and $|x-a| < 1$ if and only if $|x| < 1$, by the strong triangle inequality, so the power series for $1/(1-x)$ centered at $x = a$ has the same disc of converges as the power series for $1/(1-x)$ centered at 0: the open unit disc.

Theorem 7.5. *If $f(x) = \sum_{n\geq 0} a_n(x-\alpha)^n$ in K has disc of convergence D and $\beta \in D$ then we can write f as a power series centered at β: $f(x) = \sum_{m\geq 0} b_m(x-\beta)^m$ where $b_m \in K$ and $x \in D$, and this power series centered at β also has disc of convergence D.*

Proof. For each $x \in D$,

$$\begin{aligned} f(x) &= \sum_{n\geq 0} a_n(x-\alpha)^n \\ &= \sum_{n\geq 0} a_n(x-\beta+\beta-\alpha)^n \\ &= \sum_{n\geq 0} a_n \sum_{m=0}^{n}\binom{n}{m}(x-\beta)^m(\beta-\alpha)^{n-m} \\ &= \sum_{n\geq 0}\sum_{m\geq 0} c_{mn}, \end{aligned}$$

where

$$c_{mn} = \begin{cases} \binom{n}{m}a_n(x-\beta)^m(\beta-\alpha)^{n-m}, & \text{if } m \leq n, \\ 0, & \text{if } m > n. \end{cases}$$

For $m \leq n$, $|c_{mn}| \leq |a_n|\max(|x-\beta|, |\beta-\alpha|)^n \leq |a_n|\max(|x-\alpha|, |\beta-\alpha|)^n$. Since f converges at x and at β, both $|a_n||x-\alpha|^n$ and $|a_n||\beta-\alpha|^n$ tend to 0 as $n \to \infty$. Also $c_{mn} = 0$ when $m > n$, so $c_{mn} \to 0$ as $\max(m, n) \to \infty$. That is sufficient to justify switching the order of

summation in the double series above:

$$\begin{aligned} f(x) &= \sum_{m\geq 0}\sum_{n\geq 0} c_{mn} \\ &= \sum_{m\geq 0}\left(\sum_{n\geq m}\binom{n}{m}a_n(\beta-\alpha)^{n-m}\right)(x-\beta)^m \\ &= \sum_{m\geq 0} b_m(x-\beta)^m, \end{aligned}$$

where

$$b_m = \sum_{n\geq m}\binom{n}{m}a_n(\beta-\alpha)^{n-m} = a_m + (m+1)a_{m+1}(\beta-\alpha) + \cdots, \tag{7.3}$$

which depends on the centers α and β but not on x.

Our computations show that if a power series centered at α has disc of convergence D in K, then the power series can be recentered at each β in D and this recentered series converges on all of D. If D' is the maximal disc of convergence of the recentered series at β, then we showed $D \subset D'$. Everything we did is symmetric in the roles of α and β: the recentered series for f at β converging on D' can be re-recentered to α, and that must be $\sum_{n\geq 0} a_n(x-\alpha)^n$ by uniqueness of coefficients. Thus $D' \subset D$ too, so $D' = D$. □

The idea of recentering a power series is a basic procedure in real and complex analysis, because the recentered series will often converge at new points not covered by the original series, thereby extending the domain of the function. Theorem 7.5 says this idea does not work in complete nonarchimedean fields: a power series defined on a disc in K can't have its domain extended by recentering the series "near the edge of the disc". There is no actual edge of a disc in K, since every point in a nonarchimedean disc can be regarded as the center, and Theorem 7.5 is consistent with that fact in its own weird way.

Theorem 7.6. *Let $f(x)$ be a power series converging on a disc D in K centered at 0. If $f(\alpha) = 0$ for some $\alpha \in D$ then we can factor out $x - \alpha$: $f(x) = (x-\alpha)g(x)$ where $g(x)$ is a power series centered at 0 converging on D.*

Proof. Recenter the power series for $f(x)$ at α by Theorem 7.5: $f(x) = \sum_{n\geq 0} a_n(x-\alpha)^n$ on D with $a_0 = f(\alpha) = 0$. Therefore $f(x) = (x-\alpha)g(x)$ where $g(x) = \sum_{n\geq 1} a_n(x-\alpha)^{n-1}$. The power series $g(x)$ converges for each $x \in D$: this is obvious at $x = \alpha$, and for $x \neq \alpha$ in D we have $|a_n(x-\alpha)^n| \to 0$ when $n \to \infty$, so $|a_n(x-\alpha)^{n-1}| = |a_n(x-\alpha)^n|/|x-\alpha| \to 0$ when $n \to \infty$.

Although g was constructed as a power series centered at α, since $0 \in D$ we can express g as a power series centered instead at 0 and the new series converges on D. □

This theorem resembles a property of polynomials: a polynomial vanishing at α can be written as $x - \alpha$ times another polynomial. The proof of that is purely algebraic, but the proof of Theorem 7.6 needs care to check the series involved converge and uses recentering, which depends on being able to interchange the order of a double series.

Thanks to Theorem 7.5, the property of a function being representable as a power series on a disc in K is independent of the choice of the point in the disc around which the series is expanded.

Definition 7.7. A function $f\colon D \to K$ defined on a disc D in K with positive radius is called *analytic* if it is a power series centered at a number in D and converging on D. A *p-adic analytic function* is an analytic function on a disc in some p-adic field.

Example 7.8. On $\mathbf{Q}_p$, the p-adic exponential function is analytic on $p\mathbf{Z}_p$ for $p \neq 2$ and on $4\mathbf{Z}_2$ for $p = 2$.

The disc in the definition of an analytic function may have the form $\{x \in K : |x-\alpha| < R\}$ or $\{x \in K : |x - \alpha| \leq R\}$, where $R > 0$. Both such discs are open and closed subsets of K.

Theorem 7.9. *For a disc D in K let $f\colon D \to K$ be analytic. If $f(\alpha) = 0$ for some $\alpha \in D$ then $f(x) = (x - \alpha)g(x)$ where $g\colon D \to K$ is analytic.*

Proof. As in the proof of Theorem 7.6 we can write f as a power series centered at α to show $f(x) = (x-\alpha)g(x)$ on D where g is a power series centered at α that converges on D. The power series g is analytic by definition (independent of the choice of center in D). □

Theorem 7.10. *When K is a p-adic field, two analytic functions on a common disc in K that have the same derivative differ by a constant on that disc.*

Proof. Let $f, g\colon D \to K$ both be analytic. By Theorem 7.5 we can express them as power series centered at the *same point* α in D. Then $f' = g'$ on D implies $f^{(n)} = g^{(n)}$ on D for all $n \geq 1$, so the coefficients of their power series at α are equal in all positive degrees: $f^{(n)}(\alpha)/n! = g^{(n)}(\alpha)/n!$ for all $n \geq 1$. The only difference between the two power series is in their constant terms ($n = 0$), so f and g differ on D by a constant. □

Theorem 7.11. *If $f\colon D \to K$ is analytic and not identically zero, its zeros in D are isolated: if $f(\alpha) = 0$ for an $\alpha \in D$ then for an $r > 0$, $f(x) \neq 0$ if $0 < |x - \alpha| < r$.*

Proof. We carry out a more careful version of the proof of Theorem 7.6. Writing $f(x) = \sum_{n\geq 0} a_n(x-\alpha)^n$ for $x \in D$, $a_0 = f(\alpha) = 0$ and some a_n is not 0, since otherwise f would be identically zero on D. Let $a_N \neq 0$ with $N \geq 1$ minimal, so $f(x) = \sum_{n\geq N} a_n(x-\alpha)^n = (x-\alpha)^N g(x)$ where $g(x) = \sum_{n\geq N} a_n(x-\alpha)^{n-N}$. The series $g(x)$ converges for all $x \in D$ by the same reason used in the proof of Theorem 7.6 (there we essentially treated N as 1).

Since $g(\alpha) = a_N$ and a power series is continuous on the disc where it converges, $\lim_{x\to\alpha} g(x) = a_N \neq 0$. Therefore there is a small $r > 0$ such that $|x - \alpha| < r \Longrightarrow g(x) \neq 0$. Then $0 < |x - \alpha| < r \Longrightarrow f(x) = (x-\alpha)^N g(x) \neq 0$. □

Corollary 7.12. *For a sequence $c_n \in \mathbf{Q}_p$ such that the series $f(x) = \sum_{n\geq 0} c_n x^n$ converges on $\mathbf{Z}_p$, if the coefficients are not all zero then f has only finitely many zeros in $\mathbf{Z}_p$.*

Proof. We will prove the contrapositive: if f has infinitely many zeros $x_1, x_2, x_3, \ldots$ in $\mathbf{Z}_p$ where the x_k are distinct then every c_n is 0. Since $\mathbf{Z}_p$ is compact, the sequence $\{x_k\}$ has a convergent subsequence, say $x_{k_i} \to x \in \mathbf{Z}_p$. Then $f(x) = \lim_{i\to\infty} f(x_{k_i}) = \lim_{i\to\infty} 0 = 0$, and the zero x is not isolated since it is a limit of the zeros x_{k_i}. Theorem 7.11 implies $f(x) = 0$ for all $x \in \mathbf{Z}_p$, so all of the power coefficients of f are 0 by uniqueness of those coefficients (Corollary 3.15 or 5.6). □

Theorems 7.9, 7.10, 7.11, and Corollary 7.12 are all true for real power series converging on a closed and bounded interval $[a, b]$. In some cases the proofs are the same, *e.g.*, using compactness of $[a, b]$ in place of compactness of $\mathbf{Z}_p$ in Corollary 7.12.

The following corollary shows that the phenomenon of nonconstant power series being periodic functions, which is well-known in real analysis (*e.g.*, trigonometric functions), can't exist in p-adic analysis.

Corollary 7.13. *Let K be a p-adic field and $f\colon D \to K$ be analytic. If f is periodic, meaning there is a $t \neq 0$ in K such that $f(x) = f(x+t)$ for all $x \in D$, then f is constant on D.*

Part of the hypothesis is that $x \in D \Longrightarrow x + t \in D$.

Proof. From $f(x) = f(x+t)$ for all $x \in D$ we get $f(x) = f(x+nt)$ for every positive integer n. In particular, for each x we have $f(x) = f(x + p^r t)$, where $r \geq 1$.

Fix an $x_0 \in D$, so $f(x_0) = f(x_0 + p^r t)$ for all $r \geq 1$. The function $F(x) = f(x) - f(x_0)$ is a power series on D such that $F(x_0) = 0$ and $F(x_0 + p^r t) = 0$ for all $r \geq 1$. Since $p^r t \to 0$ as $r \to \infty$ and $t \neq 0$, the point x_0 is not an isolated zero of F. Therefore F is identically zero on D by Theorem 7.11, so f is a constant function. □

Remark 7.14. While nonconstant additively periodic functions do not exist over the p-adics, there are nonconstant *multiplicatively* periodic functions: $f(qx) = f(x)$ for some $q \neq 0$ and $x \in K^\times$ with $x \notin q^{\mathbf{Z}}$. The point is that a p-adic field K does not have discrete additive subgroups other than $\{0\}$, but the multiplicative group $K^\times$ has many discrete subgroups, namely $q^{\mathbf{Z}}$ when $|q| \neq 1$.

8. The p-adic logarithm

In this section K is a p-adic field.

We saw in Example 3.10 that the power series $\sum_{n\geq 1}(-1)^{n-1}x^n/n$ has disc of convergence $\{x \in K : |x| < 1\}$. In real analysis this is the power series for $\log(1+x)$, so we adopt this notation in K as well.

Definition 8.1. For $x \in K$ with $|x| < 1$, the *p-adic logarithm* at $1 + x$ is

$$\log(1+x) = \sum_{n\geq 1}(-1)^{n-1}\frac{x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots.$$

Equivalently, for $y \in K$ with $|y - 1| < 1$,

$$\log y = \sum_{n\geq 1}(-1)^{n-1}\frac{(y-1)^n}{n} = (y-1) - \frac{(y-1)^2}{2} + \frac{(y-1)^3}{3} - \cdots.$$

This "log" function does *not* have a base in K. It is defined by its series.

Example 8.2. The logarithm on $\mathbf{Q}_p$ has domain $1 + p\mathbf{Z}_p$ (including when $p = 2$).

Example 8.3. In $\mathbf{Q}_5$ we will compute the first five digits of $\log 11$, so we are seeking $\log 11$ modulo 5^5. As a series,

$$\log 11 = \log(1+10) = \sum_{n\geq 1}(-1)^{n-1}\frac{10^n}{n}.$$

We know the general term in the series tends to 0, so $|(-1)^{n-1}10^n/n|_5 \leq 1/5^5$ for all large n, but when does this bound occur and remain true thereafter? The bound is equivalent to $\mathrm{ord}_5(10^n/n) \geq 5$, which says $n - \mathrm{ord}_5(n) \geq 5$. A table suggests this works for $n \geq 6$.

n	1	2	3	4	5	6	7	8	9	10
$\mathrm{ord}_5(n)$	0	0	0	0	1	0	0	0	0	1
$n - \mathrm{ord}_5(n)$	1	2	3	4	4	6	7	8	9	9

To prove $n - \mathrm{ord}_5(n) \geq 5$ for $n \geq 6$ we will work out a lower bound on $n - \mathrm{ord}_5(n)$ from an upper bound on $\mathrm{ord}_5(n)$:

$$5^{\mathrm{ord}_5(n)} \leq n \Longrightarrow \mathrm{ord}_5(n) \leq \log_5(n),$$

so $n - \mathrm{ord}_5(n) \geq n - \log_5(n)$. By calculus, the function $g(x) = x - \log_5 x$ on $(0, \infty)$ is increasing for $x \geq 1/\ln 5 \approx .378$, so from $g(6) \approx 4.8$ and $g(7) \approx 5.7$, we have $n - \mathrm{ord}_5(n) \geq g(n) \geq 5$ for $n \geq 7$. Since also $6 - \mathrm{ord}_5(6) = 6 > 5$ we have proved $\mathrm{ord}_5(10^n/n) \geq 5$ for $n \geq 6$. Thus

$$\log 11 = \log(1 + 10) = 10 - \frac{10^2}{2} + \frac{10^3}{3} - \frac{10^4}{4} + \frac{10^5}{5} \bmod 5^5.$$

The sum of the first 5 terms in the series is $53380/3 \equiv 255/3 \equiv 85 \bmod 5^5$, so

$$\log 11 \equiv 85 \equiv 02300 \bmod 5^5.$$

Theorem 8.4. *If $|x - 1| < 1$ in K then $\log'(x) = 1/x$.*

Proof. We can write $\log x$ as a power series in $x - 1$, and such series when $|x - 1| < 1$ can be differentiated termwise, so

$$\begin{aligned}
\log'(x) &= \left(\sum_{n \geq 1} (-1)^{n-1} \frac{(x-1)^n}{n}\right)' \\
&= \sum_{n \geq 1} (-1)^{n-1} \frac{n}{n} (x-1)^{n-1} \\
&= \sum_{n \geq 1} (1-x)^{n-1} \\
&= \frac{1}{1 - (1-x)} \\
&= \frac{1}{x}.
\end{aligned}$$

□

The real logarithm sends products to sums: it is a group homomorphism $(0, \infty) \to \mathbf{R}$. The p-adic logarithm's domain $\{x \in K : |x - 1| < 1\}$ is also a multiplicative group: by the strong triangle inequality, $|x - 1| < 1 \Longrightarrow |x| = |(x - 1) + 1| = 1$, so $|x - 1| < 1$ and $|y - 1| < 1$ imply $|xy - 1| = |(x - 1)y + (y - 1)| \leq \max(|x - 1||y|, |y - 1|) < 1$ and $|1/x - 1| = |(1 - x)/x| = |1 - x| < 1$. The values of the p-adic logarithm are in K, an additive group. Is the p-adic logarithm a group homomorphism?

Theorem 8.5. *If $|x-1| < 1$ and $|y-1| < 1$ in a p-adic field K then $\log(xy) = \log x + \log y$.*

Proof. Fix a choice of y. We consider the two expressions $\log(xy)$ and $\log x + \log y$ as functions of x on the disc $\{x \in K : |x - 1| < 1\}$. To prove they are equal, look at the x-derivative of both sides. By Theorem 8.4, $(\log x + \log y)' = \log'(x) = 1/x$. By the chain rule, $(\log(xy))' = (1/(xy))y = 1/x$.

Since the two derivative formulas match, we expect the functions $\log(xy)$ and $\log x + \log y$ differ by a constant: $\log(xy) = \log x + \log y + C_y$ for some $C_y \in K$. Setting $x = 1$, we would then see that $\log y = \log y + C_y$, so $C_y = 0$ and thus $\log(xy) = \log x + \log y$. Since y was fixed but arbitrary, and x was arbitrary, we have proved the desired identity.

However, this reasoning has a gap. The argument was based on knowing that two p-adic functions on a disc with the same derivative on that disc differ by a constant, and we have

shown this holds for two functions on a disc represented by power series that meet the hypotheses of Theorem 7.1(4) or Theorem 7.10. Both $\log x + \log y$ and $\log(xy)$, as functions of x, have power series expansions:

$$\log x + \log y = \log y + \sum_{n \geq 1} (-1)^{n-1} \frac{(x-1)^n}{n}$$

and

$$\log(xy) = \sum_{n \neq 1} (-1)^{n-1} \frac{(xy-1)^n}{n} = \sum_{n \neq 1} (-1)^{n-1} y^n \frac{(x-1/y)^n}{n}.$$

The first series is centered at 1 and the second is centered at $1/y$. Both series converge on the disc $D = \{x \in K : |x-1| < 1\}$, and D contains $1/y$ since $|1/y - 1| = |1-y|/|y| = |1-y| < 1$. Thus $\log x + \log y$ and $\log(xy)$ are power series in x centered at different points of D (unless $y = 1$: a trivial case), so we could use Theorem 7.10. Or by recentering, we can write the two functions of x as power series on D centered at the same point and then use Theorem 7.1(4) (equality of derivatives does not depend on the center). □

Corollary 8.6. *If $|x-1| < 1$ in K then $\log(1/x) = -\log x$ and $\log(x^n) = n \log x$ for all $n \in \mathbf{Z}$.*

Proof. Taking the logarithm on both sides of $x(1/x) = 1$ we get $\log x + \log(1/x) = \log(1) = 0$, so $\log(1/x) = -\log x$. We get $\log(x^n) = n \log x$ for $n \in \mathbf{Z}^+$ from Theorem 8.5 by induction. The equation is trivial for $n = 0$, and the case $n < 0$ follows from the case $n > 0$ since $\log(x^{-n}) = \log((x^n)^{-1}) = -\log(x^n)$. □

The p-adic exponential is defined near 0 and the p-adic logarithm is defined near 1. We saw in Theorem 4.5 that the p-adic exponential is an isometry on the disc where it converges. The p-adic logarithm is an isometry too, but only on numbers close enough to 1.

Theorem 8.7. *If $|t-1| < (1/p)^{1/(p-1)}$ then $|\log t| = |t-1|$. If $|x-1| < (1/p)^{1/(p-1)}$ and $|y-1| < (1/p)^{1/(p-1)}$ then $|\log x - \log y| = |x-y|$.*

Proof. If $t = 1$ then $\log t = t - 1 = 0$. For $0 < |t-1| < (1/p)^{1/(p-1)}$, extract the first term from the power series for $\log t$:

$$\log t = (t-1) + \sum_{n \geq 2} (-1)^{n-1} \frac{(t-1)^n}{n}.$$

We will show $|(-1)^{n-1}(t-1)^n/n| < |t-1|$ for $n \geq 2$. Then $|t - 1 + \sum_{n=2}^{N} (-1)^{n-1}(t-1)^n/n| = |t-1|$ for $N \geq 2$, and letting $N \to \infty$ we get $|\log t| = |t-1|$.

The condition $|(-1)^{n-1}(t-1)^n/n| < |t-1|$ is equivalent to $|t-1|^{n-1} < |n| = |n|_p$, which for $n \geq 2$ says

$$|t-1| < |n|_p^{1/(n-1)}. \tag{8.1}$$

Why is this inequality true? Since $|t-1| < (1/p)^{1/(p-1)}$, we can verify (8.1) by showing

$$\left(\frac{1}{p}\right)^{1/(p-1)} \leq |n|_p^{1/(n-1)}.$$

It is left to the reader to check this inequality holds for all $n \geq 2$ (and equality occurs if and only if $n = p$). This completes the proof of the first part.

To show $|\log x - \log y| = |x-y|$, rewrite $|\log x - \log y|$ as $|\log(x/y)|$ by Corollary 8.6. To apply the first part to $|\log(x/y)|$, we check

$$\left|\frac{x}{y} - 1\right| = \frac{|x-y|}{|y|} = |x-y| \leq \max(|x-1|, |y-1|) < \left(\frac{1}{p}\right)^{1/(p-1)}.$$

Therefore $|\log(x/y)| = |x/y - 1| = |x-y|/|y| = |x-y|$. □

Recall the disc D_p in (4.1).

Corollary 8.8. *The logarithm is injective on* $\{x \in K : |x-1| < (1/p)^{1/(p-1)}\} = 1 + D_p$.

Proof. If $\log x = \log y$ for x and y in $1 + D_p$ then $|x-y| = |\log x - \log y| = 0$, so $x = y$. □

Example 8.9. For prime $p \neq 2$,

$$\{x \in \mathbf{Q}_p : |x-1|_p < (1/p)^{1/(p-1)}\} = \{x \in \mathbf{Q}_p : |x-1|_p \leq 1/p\} = 1 + p\mathbf{Z}_p,$$

which is the full domain on which the p-adic logarithm is defined in $\mathbf{Q}_p$. Thus the logarithm on $1 + p\mathbf{Z}_p$ is injective and an isometry here: $|\log x - \log y|_p = |x-y|_p$. Its image is in $p\mathbf{Z}_p$ since $|\log x|_p = |x-1|_p \leq 1/p$.

For $p = 2$,

$$\{x \in \mathbf{Q}_2 : |x-1|_2 < (1/2)\} = \{x \in \mathbf{Q}_2 : |x-1|_2 \leq 1/4\} = 1 + 4\mathbf{Z}_2,$$

so the 2-adic logarithm is injective and an isometry on $1 + 4\mathbf{Z}_2$, with image in $4\mathbf{Z}_2$. The set $1 + 4\mathbf{Z}_2$ is not the whole domain for the 2-adic log, whose power series converges on $1 + 2\mathbf{Z}_2$. It turns out that on $1 + 2\mathbf{Z}_2$ the logarithm is *not* an isometry or injective. The reason is explained most easily with -1. Since $-1 \in 1 + 2\mathbf{Z}_2$ and $(-1)^2 = 1$, taking the logarithm of both sides shows $2\log(-1) = \log(1) = 0$ by Theorem 8.5, so $\log(-1) = 0$. Hence $\log(-1) = \log(1)$ even though $-1 \neq 1$. More generally, $\log(-a) = \log a$ for all $a \in 1 + 2\mathbf{Z}_2$. So if $a \equiv 3 \bmod 4$ then $|\log a|_2 = |\log(-a)|_2 = |-a-1|_2 = |a+1|_2$ since $-a \equiv 1 \bmod 4$.

Example 8.10. The 2-adic equation $\log(-1) = 0$, which says $\log(1-2) = 0$, is equivalent to

$$\sum_{n \geq 1} (-1)^{n-1} \frac{(-2)^n}{n} = -\sum_{n \geq 1} \frac{2^n}{n} = 0,$$

so in $\mathbf{Q}_2$

$$\sum_{n \geq 1} \frac{2^n}{n} = 0. \tag{8.2}$$

Using the partial sums $s(n) = 2 + 2^2/2 + 2^3/3 + \cdots + 2^n/n$, (8.2) says $\mathrm{ord}_2(s(n)) \to \infty$ as $n \to \infty$. In the table below, the 2-adic valuations of $\mathrm{ord}_2(s(n))$ grow irregularly.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
$\mathrm{ord}_2(s(n))$	1	2	2	5	8	5	5	13	9	10	10	12	12	12	12	22	17	18

Since $\mathrm{ord}_2(s(n))$ doesn't directly mention 2-adic logarithms, it's natural to ask whether we can prove $\mathrm{ord}_2(s(n)) \to \infty$ as $n \to \infty$ without using 2-adic logarithms. This can be done. See the accepted answer at https://math.stackexchange.com/questions/2816184. The idea is to show when n is even that $\sum_{k=1}^{n} 2^k/k$ is n times a 2-adic integer, so $\sum_{k=1}^{2^m} 2^k/k$ when $m \geq 1$ is 2^m times a 2-adic integer. Thus $\sum_{k=1}^{2^m} 2^k/k \to 0$ in $\mathbf{Z}_2$ as $m \to \infty$, so the series $\sum_{k \geq 1} 2^k/k$ has a subsequence among its partial sums that tends to 0. Since the whole

series converges in $\mathbf{Q}_2$, it equals every limit taken along a subsequence of its partial sums, and thus the whole series is 0.

Based on a lot of numerical data, that $\mathrm{ord}_2(s(n)) \to \infty$ as $n \to \infty$ appears to be expressible as the following lower bound.

Conjecture 8.11. *When $2^r \leq n < 2^{r+1}$, $\mathrm{ord}_2(s(n)) \geq n - r$, with equality if and only if $n = 2^{r+1} - 1$.*

Here is a corresponding upper bound based on numerical data.

Conjecture 8.12. *When $2^r \leq n < 2^{r+1}$, $\mathrm{ord}_2(s(n)) \leq n + 2r - 2$ except at $n = 8$, and there is equality if $n = 2^r$ except if n is 4 or 8.*

In Conjecture 8.12 we write "if" rather than "if and only if" since sometimes $\mathrm{ord}_2(s(n)) = n + 2r - 2$ when n is not a power of 2: it holds at $n = 40$ and $n = 296$.

Rewriting $2^r \leq n < 2^{r+1}$ as $r \leq \log_2 n < r + 1$ (where $\log_2$ is the high school base-2 logarithm), the bounds in Conjectures 8.11 and 8.12 say

$$n - \lfloor \log_2 n \rfloor \leq \mathrm{ord}_2(s(n)) \leq n + 2\lfloor \log_2 n \rfloor - 2$$

for all $n \geq 1$ except when $n = 8$ in the upper bound.

Returning to general primes, the p-adic exponential and logarithm are continuous homomorphisms $\exp\colon p\mathbf{Z}_p \to 1 + p\mathbf{Z}_p$ and $\log\colon 1 + p\mathbf{Z}_p \to p\mathbf{Z}_p$ for $p \neq 2$ and $\exp\colon 4\mathbf{Z}_2 \to 1 + 4\mathbf{Z}_2$ and $\log\colon 1 + 4\mathbf{Z}_2 \to 4\mathbf{Z}_2$ for $p = 2$. We will show next that these are inverse functions. More generally, on a p-adic field these functions are inverses on the discs where we proved they are isometries.

Theorem 8.13. *The functions $\exp\colon D_p \to 1 + D_p$ and $\log\colon 1 + D_p \to D_p$ are inverses, where $D_p = \{x \in K : |x| < (1/p)^{1/(p-1)}\}$.*

Before getting into the proof of Theorem 8.13, some comments are in order to get the right perspective about the result and its proof.

You might think the proof should be easy using derivatives. After all, setting $f(x) = \log(e^x)$, we have by the chain rule that $f'(x) = (\log'(e^x))(e^x)' = (1/e^x)(e^x) = 1$, so f has constant derivative 1. That means $f(x) = x + c$ for some constant c. Then we can find c by setting $x = 0$: $f(0) = \log(e^0) = \log(1) = 0$, so c is 0 and $f(x) = x$, *i.e.*, $\log(e^x) = x$, right? This is how a proof in real analysis could go. But such reasoning in p-adic analysis has a *gap*: why must a function with derivative 1 be linear? Recall by Example 5.9 that not all functions $\mathbf{Z}_p \to \mathbf{Q}_p$ with derivative 0 have to be constant! We know that on a disc in a p-adic field, an *analytic* function on the disc with derivative 1 must have the form $x + c$, since the derivative determines the function up to an additive constant. But that begs the question: why is $\log(e^x)$ analytic? That $\log x$ and e^x are each analytic is not good enough: the composition of analytic functions does not have to be analytic!

Example 8.14. Consider the 2-adic exponential and logarithm. Both $\log\colon 1 + 2\mathbf{Z}_2 \to \mathbf{Q}_2$ and $\exp\colon 4\mathbf{Z}_2 \to \mathbf{Q}_2$ are analytic (each is a power series on a disc). Since $1+2\mathbf{Z}_2 = \pm(1+4\mathbf{Z}_2)$ and $\log(-1) = 0$, we have $\log(1 + 2\mathbf{Z}_2) = \log(1 + 4\mathbf{Z}_2) \subset 4\mathbf{Z}_2$. Also $\exp(4\mathbf{Z}_2) \subset 1 + 4\mathbf{Z}_2$, so if for $x \in 1 + 2\mathbf{Z}_2$ we set $h(x) = e^{\log x}$ then $h(1 + 2\mathbf{Z}_2) \subset 1 + 4\mathbf{Z}_2$. Therefore we can't have $h(x) = x$ for $x \in 1 + 2\mathbf{Z}_2$ with $x \notin 1 + 4\mathbf{Z}_2$. (We know it's false at $x = -1$!) In fact it turns out that for $x \in 1 + 2\mathbf{Z}_2$,

$$e^{\log x} = \begin{cases} x, & \text{if } x \equiv 1 \bmod 4, \\ -x & \text{if } x \equiv 3 \bmod 4. \end{cases}$$

This composite of analytic functions is not an analytic function since it is not given by a single power series on $1 + 2\mathbf{Z}_2$.

There is a theorem giving general conditions under which a formal composition of power series $h(X) = f(g(X))$ satisfies the numerical identity $h(x) = f(g(x))$ for x in the domain of convergence of $g(X)$. (See [2, pp. 99-101] for p-adic power series and [6, pp. 180-181] for real power series; the real case uses hypotheses of absolute convergence.) The only instance of composition of power series that we will use is for the p-adic exponential and logarithm, so we will bypass proving a general theorem justifying formal-to-numerical composition and give instead a proof tailor-made to the p-adic exponential and logarithm series. The following proof of Theorem 8.13 can be safely skipped and returned to later.

Proof. Since $\exp x$ is an isometry on D_p and $\log x$ is an isometry on $1 + D_p$, we know $\exp(D_p) \subset 1 + D_p$ and $\log(1 + D_p) \subset D_p$. Therefore it makes sense to ask if $\log(e^x) = x$ for all $x \in D_p$ and if $e^{\log y} = y$ for all y in $1 + D_p$.

If we prove $\log(e^x) = x$ for all $x \in D_p$ or $e^{\log y} = y$ for all $y \in 1 + D_p$ then the other follows. For example, if $\log(e^x) = x$ for all $x \in D_p$ then for $y \in 1 + D_p$ set $x = \log y \in D_p$ and thus $\log(e^{\log y}) = \log y$. Both $e^{\log y}$ and y are in $1 + D_p$ and log is injective on this disc since it is an isometry there, so from $\log(e^{\log y}) = \log y$ on $1+D_p$ we get $e^{\log y} = y$ on $1+D_p$.

We will now prove $\log(e^x) = x$ for $x \in D_p$. *From now on, x always runs over D_p.*

The formula $\log(e^x) = \sum_{n\geq 1}(-1)^{n-1}(e^x - 1)^n/n$, is not directly a power series in x, but rather in $e^x - 1$. For $N \geq 1$, define the truncated log polynomial

$$L_N(T) = \sum_{n=1}^{N} \frac{(-1)^{n-1}}{n} T^n = T - \frac{1}{2}T^2 + \cdots + \frac{(-1)^{N-1}}{N} T^N.$$

For $y \in 1 + D_p$, $L_N(y - 1) \to \log y$ as $N \to \infty$. When $x \in D_p$ and we set $y = e^x$, we get $L_N(e^x - 1) \to \log(e^x)$. We will now prove that $L_N(e^x - 1) \to x$ as $N \to \infty$, so $\log(e^x) = x$.

Since $L_N(e^x - 1)$ is a *polynomial* in $e^x - 1$, it is a finite sum of powers of analytic functions on D_p and therefore it is analytic on D_p (Theorem 3.1). Let the expansion of $L_N(e^x - 1)$ as a power series in x have coefficients $c_{m,N}$ for $m \geq 0$: for $x \in D_p$,

$$L_N(e^x - 1) = \sum_{n=1}^{N} \frac{(-1)^{n-1}}{n}(e^x - 1)^n = \sum_{m\geq 0} c_{m,N} x^m. \tag{8.3}$$

We want to understand what is happening to the coefficients $c_{m,N}$ as N grows.

Because $e^x - 1 = x + x^2/2! + \cdots$ has constant term 0 and first term x, $(e^x - 1)^n$ is a power series in x with first term x^n. Therefore the expression of

$$L_N(e^x - 1) = (e^x - 1) - \frac{(e^x - 1)^2}{2} + \cdots + (-1)^{N-1}\frac{(e^x - 1)^N}{N}$$

as a polynomial in $e^x - 1$ shows $c_{0,N} = 0$ and $c_{1,N} = 1$, so

$$L_N(e^x - 1) = x + \sum_{m\geq 2} c_{m,N} x^m. \tag{8.4}$$

We will use derivatives to show $c_{m,N} = 0$ for $2 \leq m \leq N$. (Since $L_N(e^x - 1) \to \log(e^x)$ as $N \to \infty$, and anticipating that $\log(e^x) = x$, we should expect in (8.4) that many of the $c_{m,N}$'s for $m \geq 2$ should vanish as N grows.)

Differentiate $L_N(e^x - 1)$ as a function of x. Since $L_N(T)$ is a polynomial, the rules of differential calculus (basically the chain rule) tell us that for $x \in D_p$,

$$(L_N(e^x - 1))' = L_N'(e^x - 1)(e^x - 1)' = L_N'(e^x - 1)e^x.$$

Since $L_N(T)$ has derivative $\sum_{n=1}^{N}(-1)^{n-1}T^{n-1}$,

$$\begin{aligned} L_N'(e^x - 1)e^x &= \sum_{n=1}^{N}(-1)^{n-1}(e^x - 1)^{n-1}e^x \\ &= \sum_{n=1}^{N}(1 - e^x)^{n-1}e^x \\ &= \frac{1 - (1 - e^x)^N}{1 - (1 - e^x)}\, e^x \\ &= 1 - (1 - e^x)^N. \end{aligned}$$

The power series for $1 - e^x$ starts with $-x$, so $1 - (1 - e^x)^N = 1 + (-1)^{N+1}x^N + \cdots$, which has *no* terms in degrees 1 through $N - 1$. By (8.3), $(L_N(e^x - 1))' = \sum_{m\geq 1} mc_{m,N}x^{m-1}$. Therefore by uniqueness of coefficients in power series expansions (Corollary 3.15 or 5.6), the coefficients of $\sum_{m\geq 1} mc_{m,N}x^{m-1}$ are 0 for $1 \leq m - 1 \leq N - 1$, so $c_{m,N} = 0$ for $2 \leq m \leq N$. Feeding this into (8.4),

$$L_N(e^x - 1) = x + \sum_{m\geq N+1} c_{m,N}x^m. \tag{8.5}$$

Now we compute $1 - (1 - e^x)^N$ as a power series more explicitly:

$$\begin{aligned} 1 - (1 - e^x)^N &= 1 - \sum_{k=0}^{N}\binom{N}{k}(-e^x)^k \\ &= \sum_{k=1}^{N}\binom{N}{k}(-1)^{k+1}e^{kx} \quad \text{by Theorem 4.4} \\ &= \sum_{k=1}^{N}\binom{N}{k}(-1)^{k+1}\sum_{m\geq 0}\frac{k^m}{m!}x^m. \end{aligned}$$

We can interchange the order of summation since the sum of finitely many convergent series can always be added termwise (you could apply Theorem 3.1 if you wish). Then

$$1 - (1 - e^x)^N = \sum_{m\geq 0}\left(\sum_{k=1}^{N}\binom{N}{k}(-1)^{k+1}k^m\right)\frac{x^m}{m!} = \sum_{m\geq 0} b_{m,N}\frac{x^m}{m!},$$

where $b_{m,N}$ is an *integer*. Since $(L_N(e^x - 1))' = 1 - (1 - e^x)^N$, and power series in x have uniquely determined coefficients, equating coefficients of x^{m-1} on both sides gives us $mc_{m,N} = b_{m-1,N}/(m - 1)!$ for $m \geq 1$. Therefore when $m \geq 1$, $c_{m,N} = b_{m-1,N}/m!$ is an integer divided by $m!$. Place this formula for $c_{m,N}$ into (8.5):

$$L_N(e^x - 1) = x + \sum_{m\geq N+1}\frac{b_{m-1,N}}{m!}x^m.$$

Subtract x and estimate:

$$|L_N(e^x - 1) - x| \leq \max_{m \geq N+1} \left| \frac{b_{m-1,N}}{m!} x^m \right| \leq \max_{m \geq N+1} \left| \frac{x^m}{m!} \right|,$$

where in the last estimate we used the fact that $b_{m-1,N}$ is an integer. Because $x \in D_p$, $x^m/m! \to 0$ as $m \to \infty$ from convergence of the p-adic exponential series at x. Therefore our upper bound on $|L_N(e^x - 1) - x|$ tends to 0 as $N \to \infty$, so $L_N(e^x - 1) \to x$ as $N \to \infty$. We already saw at the start that $L_N(e^x - 1) \to \log(e^x)$ as $N \to \infty$, so $\log(e^x) = x$ by uniqueness of limits. □

Example 8.15. For p odd, e^x is an isomorphism from $p\mathbf{Z}_p$ onto $1 + p\mathbf{Z}_p$ with inverse $\log y$. These isomorphisms are isometries, so e^x maps $p^k\mathbf{Z}_p$ onto $1 + p^k\mathbf{Z}_p$ if $k \geq 1$. For $p = 2$, e^x is an isomorphism from $4\mathbf{Z}_2$ onto $1 + 4\mathbf{Z}_2$, and the image of $2^k\mathbf{Z}_2$ under e^x is $1 + 2^k\mathbf{Z}_2$ if $k \geq 2$.

In $\mathbf{R}$ the exponential function with its infinite radius of convergence defines a *global* group isomorphism from all of $\mathbf{R}$ to the positive reals $(0, \infty)$. In the p-adic world, the exponential function sets up a *local* group isomorphism between a neighborhood of 0 (additive) and a neighborhood of 1 (multiplicative).

Example 8.16. We saw in Example 8.3 that in $\mathbf{Q}_5$, $\log 11 \equiv 85 \bmod 5^5$. Since the 5-adic exponential is an isometry $5\mathbf{Z}_5 \to 1 + 5\mathbf{Z}_5$, we must have $e^{85} \equiv 11 \bmod 5^5$. Let's check. As a series, $e^{85} = \sum_{n \geq 0} 85^n/n!$. The reader should check that $85^n/n! \equiv 0 \bmod 5^5$ for $n \geq 6$, so

$$\begin{aligned} e^{85} &\equiv 1 + 85 + \frac{85^2}{2} + \frac{85^3}{6} + \frac{85^4}{24} + \frac{85^5}{5!} \bmod 5^5 \\ &\equiv 1 + 85 + 2050 + 2875 + 2500 + 1875 \bmod 5^5 \\ &\equiv 11 \bmod 5^5. \end{aligned}$$

9. Power series and root extractions

In this section we look at the p-adic version of power series for power functions.

One of Newton's great discoveries was the extension of the binomial theorem from integral to real exponents:

$$(1 + x)^t = \sum_{n \geq 0} \binom{t}{n} x^n = 1 + tx + \binom{t}{2} x^2 + \cdots$$

for $t \in \mathbf{R}$, which is absolutely convergent for $|x| < 1$ in $\mathbf{R}$.[4] In particular, taking $t = 1/m$ for $m \in \mathbf{Z}^+$,

$$(1 + x)^{1/m} = \sum_{n \geq 0} \binom{1/m}{n} x^n \tag{9.1}$$

is $\sqrt[m]{1 + x}$, the positive real mth root of $1 + x$, when $-1 < x < 1$.

We will look at Newton's binomial series for p-adic x and p-adic integer exponents t, and in particular show when $t = 1/m$ that the mth root series (9.1) describes p-adic mth roots, but it might not converge to the mth root you'd expect.

Theorem 9.1. *For $t \in \mathbf{Z}_p$, the series $\sum_{n \geq 0} \binom{t}{n} x^n$ converges for $|x|_p < 1$.*

[4]For $x = \pm 1$ the series converges for some t, and of course it converges for all x when $t \in \mathbf{Z}^+$ since then the series is a polynomial.

Proof. The binomial coefficients $\binom{t}{n}$ are all in $\mathbf{Z}_p$: we can write $t = \lim_{k\to\infty} t_k$ where $t_k \in \mathbf{Z}^+$ (*e.g.*, t_k is the kth truncation of the p-adic expansion of t), so $\binom{t}{n} = \lim_{k\to\infty}\binom{t_k}{n}$ for each n since $\binom{x}{n}$ is a polynomial in x (of degree n) and polynomials are p-adically continuous. Each $\binom{t_k}{n}$ is a positive integer by combinatorics if $t_k \geq n$ and is 0 if $0 \leq t_k < n$, so $\binom{t}{n}$ is a p-adic limit of integers and thus is a p-adic integer.

Since $|\binom{t}{n}|_p \leq 1$ for all n, the power series $\sum_{n\geq 0}\binom{t}{n}x^n$ converges when $|x|_p < 1$.[5] □

Definition 9.2. For $t \in \mathbf{Z}_p$ and $|x|_p < 1$, define $(1+x)^t := \sum_{n\geq 0}\binom{t}{n}x^n$.

When t is a positive integer, the series $\sum_{n\geq 0}\binom{t}{n}x^n$ is a polynomial and equals the t-th power of $1+x$ by the binomial theorem. We want to justify the notation $(1+x)^t$ as a power for other exponents t, particularly when t is a rational number in $\mathbf{Z}_p$.

Theorem 9.3. *For t and t' in $\mathbf{Z}_p$, and $|x|_p < 1$, $(1+x)^t(1+x)^{t'} = (1+x)^{t+t'}$.*

Proof. By Theorem 3.1,

$$(1+x)^t(1+x)^{t'} = \sum_{n\geq 0}\binom{t}{n}x^n\sum_{n\geq 0}\binom{t'}{n}x^n = \sum_{n\geq 0}\left(\sum_{k=0}^{n}\binom{t}{k}\binom{t'}{n-k}\right)x^n.$$

The coefficient here of x^n equals $\binom{t+t'}{n}$ as a special case of the polynomial identity in X and Y called Vandermonde convolution:

$$\binom{X+Y}{n} = \sum_{k=0}^{n}\binom{X}{k}\binom{Y}{n-k}. \tag{9.2}$$

Thus $(1+x)^t(1+x)^{t'} = (1+x)^{t+t'}$. □

Corollary 9.4. *If m is a positive integer not divisible by p and $x \in p\mathbf{Z}_p$ then the series $(1+x)^{1/m} := \sum_{n\geq 0}\binom{1/m}{n}x^n$ equals the unique mth root of $1+x$ in $1+p\mathbf{Z}_p$.*

Proof. Since m is not divisible by p, Hensel's lemma for the polynomial $T^m - (1+x)$ with approximate root 1 shows $1+x$ has a unique mth root in $1+p\mathbf{Z}_p$. We want to show this mth root is the series $\sum_{n\geq 0}\binom{1/m}{n}x^n$.

Since $1/m \in \mathbf{Z}_p$, for $n \geq 1$ we have $|\binom{1/m}{n}x^n|_p \leq |x|_p < 1$, so $\sum_{n\geq 0}\binom{1/m}{n}x^n \equiv 1 \bmod p$. By Theorem 9.3, the mth power of the series $(1+x)^{1/m}$ is

$$(1+x)^{\overbrace{1/m+\cdots+1/m}^{m \text{ times}}} = (1+x)^1 = 1+x,$$

so $\sum_{n\geq 0}\binom{1/m}{n}x^n$ is the unique mth root of $1+x$ that lies in $1+p\mathbf{Z}_p$. □

Example 9.5. Let's look at square roots. If $p \neq 2$ and $x \in 1+p\mathbf{Z}_p$, the infinite series

$$\sum_{n\geq 0}\binom{1/2}{n}x^n$$

[5]For $t \in \mathbf{Z}_p$ with $t \notin \mathbf{N}$, $\binom{t}{n} \in \mathbf{Z}_p^\times$ infinitely often: if $t = \sum_{k\geq 0}a_kp^k$ then $\binom{t}{p^k} \equiv a_k \bmod p$ for all k, and $1 \leq a_k \leq p-1$ infinitely often. So the disc of convergence of the series $(1+x)^t$ is the open unit disc.

is the unique square root of $1+x$ in $1+p\mathbf{Z}_p$. In $\mathbf{Q}_3$ and $\mathbf{Q}_5$ for instance, $\sum_{n\geq 0}\binom{1/2}{n}(15/49)^n$ is a square root of $1+15/49 = 64/49 = (8/7)^2$, so the series is $\pm 8/7$. Since $-8/7 \equiv 1 \bmod 3$ and $-8/7 \equiv 1 \bmod 5$,

$$\sum_{n\geq 0}\binom{1/2}{n}\left(\frac{15}{49}\right)^n = -\frac{8}{7}$$

in $\mathbf{Q}_3$ and $\mathbf{Q}_5$. All terms in this series are rational, and the same series converges in $\mathbf{R}$ to the positive square root of $(8/7)^2$, which is $8/7$. This is a concrete example of an infinite series of rational numbers having different rational values in $\mathbf{R}$ and $\mathbf{Q}_3$.

Example 9.6. In $\mathbf{Q}_{19}$, $\sum_{n\geq 0}\binom{1/3}{n}(19/8)^n$ is a cube root of $1 + 19/8 = 27/8 = (3/2)^3$ and is not $3/2$ since $3/2 \not\equiv 1 \bmod 19$. The cube roots of unity in $\mathbf{Q}_{19}$ are the Teichmüller representatives $1, \omega(7)$, and $\omega(11)$. Since $(3/2)7 \equiv 1 \bmod 19$, the series has value $(3/2)\omega(7)$ in $\mathbf{Q}_{19}$. The series does not converge in $\mathbf{R}$.

Remark 9.7. In $\mathbf{Q}_2$, $\sum_{n\geq 0}\binom{1/2}{n}x^n$ does not converge on all of $2\mathbf{Z}_2$ since $1/2 \notin \mathbf{Z}_2$: $\binom{1/2}{n}$ is not a 2-adic integer for $n \geq 1$, and in fact $|\binom{1/2}{n}|_2 = |1/2|_2^n/|n!|_2 = 2^{2n-s_2(n)}$. The series converges on $8\mathbf{Z}_2$, where it equals the square root of $1+x$ that is $\equiv 1 \bmod 4$. For example, $\sum_{n\geq 0}\binom{1/2}{n}8^n$ is a square root of $1 + 8 = 9 = 3^2$, and its value is -3 since $-3 \equiv 1 \bmod 4$.

More generally, if $t \in \mathbf{Q}_p - \mathbf{Z}_p$, so $|t|_p > 1$, then $|\binom{t}{n}|_p = |t|_p^n/|n!|_p$ and $\sum_{n\geq 0}\binom{t}{n}x^n$ converges if and only if $|x|_p < (1/|t|_p)(1/p)^{1/(p-1)}$.

To relate $(1+x)^t$ to integral powers of $1+x$ when $t \in \mathbf{Z}_p$, we'll show it depends continuously on t.

Theorem 9.8. *If $x \in p\mathbf{Z}_p$ and t and t' are in $\mathbf{Z}_p$ then $|(1+x)^t - (1+x)^{t'}|_p \leq |t-t'|_p|x|_p$.*

Proof. The constant terms of the series $(1+x)^t$ and $(1+x)^{t'}$ are both 1, so

$$(1+x)^t - (1+x)^{t'} = \sum_{n\geq 1}\left(\binom{t}{n} - \binom{t'}{n}\right)x^n$$
$$\Longrightarrow |(1+x)^t - (1+x)^{t'}|_p \leq \max_{n\geq 1}\left|\binom{t}{n} - \binom{t'}{n}\right|_p |x|_p^n.$$

Set $\delta = t - t'$, so by the Vandermonde identity (9.2)

$$\binom{t}{n} = \binom{t'+\delta}{n} = \sum_{k=0}^{n}\binom{t'}{k}\binom{\delta}{n-k} = \sum_{k=0}^{n-1}\binom{t'}{k}\binom{\delta}{n-k} + \binom{t'}{n}.$$

Using the identity $\binom{X}{m} = \frac{X}{m}\binom{X-1}{m-1}$ for $m \geq 1$,

$$\begin{aligned}
\binom{t}{n} - \binom{t'}{n} &= \sum_{k=0}^{n-1} \binom{t'}{k} \frac{\delta}{n-k} \binom{\delta-1}{n-k-1} \\
\implies \left| \binom{t}{n} - \binom{t'}{n} \right|_p &\leq \max_{0 \leq k \leq n-1} \frac{|\delta|_p}{|n-k|_p} \\
&= |\delta|_p \max\left(\frac{1}{|1|_p}, \frac{1}{|2|_p}, \ldots, \frac{1}{|n|_p} \right) \\
\implies \left| \left(\binom{t}{n} - \binom{t'}{n} \right) x^n \right|_p &\leq |\delta|_p \max_{1 \leq k \leq n} \frac{|x|_p^n}{|k|_p} \\
&\leq |\delta|_p \max_{1 \leq k \leq n} \frac{|x|_p^k}{|k|_p} \quad \text{since } |x|_p < 1 \\
&= |t-t'|_p \max_{1 \leq k \leq n} \left| \frac{x^k}{k} \right|_p .
\end{aligned}$$

By the proof of Theorem 8.7, $|x^k/k|_p < |x|_p$ for $k \geq 2$ if $|x|_p < (1/p)^{1/(p-1)}$. Similar reasoning shows $|x^k/k|_p \leq |x|_p$ for $k \geq 1$ if $|x|_p \leq (1/p)^{1/(p-1)}$, so for $x \in p\mathbf{Z}_p$ we have $|x^k/k|_p \leq |x|_p$ for $k \geq 1$, and thus

$$\left| \left(\binom{t}{n} - \binom{t'}{n} \right) x^n \right|_p \leq |t-t'|_p \max_{1 \leq k \leq n} \left| \frac{x^k}{k} \right|_p \leq |t-t'|_p |x|_p. \qquad \square$$

Corollary 9.9. *If $t \in \mathbf{Z}_p$ and $\{t_k\} \subset \mathbf{Z}^+$ satisfies $t_k \to t$ as $k \to \infty$, then for $x \in p\mathbf{Z}_p$, $(1+x)^{t_k} \to (1+x)^t$ as $k \to \infty$. Thus $(1+x)^t$ is a p-adic limit of integral powers of $1+x$.*

Proof. By Theorem 9.8, $|(1+x)^t - (1+x)^{t_k}|_p \leq |t-t_k|_p |x|_p \leq |t-t_k|_p$, and $|t-t_k|_p \to 0$. $\square$

10. Strassmann's theorem

What we have done so far with p-adic power series has been motivated by the real case. Weird things happen, like e^x having a finite p-adic radius of convergence, but p-adic power series are broadly similar to real power series (radius of convergence, termwise differentiability, *etc.*) In this section we will meet a phenomenon without an analogue in real analysis: a way to bound the number of roots of a p-adic power series based on the sizes of its coefficients.

Our focus will be on power series $\sum_{n \geq 0} a_n x^n$ over K that converge on the closed unit disc $\{x \in K : |x| \leq 1\}$. This convergence is equivalent to $a_n \to 0$ (check convergence at 1). Taking $K = \mathbf{Q}_p$, for instance, Theorem 7.12 tells us that a series over $\mathbf{Q}_p$ converging on $\mathbf{Z}_p$ that is not identically zero has finitely many zeros in $\mathbf{Z}_p$. That finiteness comes from compactness of $\mathbf{Z}_p$ and is purely qualitative. We will make the finiteness quantitative by deriving an upper bound on the number of those zeros using a theorem of Strassmann from 1928.

When a power series $f(x) = \sum_{n \geq 0} a_n x^n$ with $a_n \in K$ has $a_n \to 0$ and the a_n's are *not all* 0, the numbers $|a_n|$ have a positive maximum and there is a last time the maximum occurs. The index farthest into the series with a coefficient of maximal absolute value will be denoted $N(f)$. That is,

$$N(f) = \max\{N \geq 0 : |a_n| \leq |a_N| \text{ for all } n \geq 0\}.$$

For the power series f whose coefficients are all 0, $N(f)$ is not defined.

Theorem 10.1 (Strassmann [8]). *Let $a_n \to 0$ in K and set $f(x) = \sum_{n\geq 0} a_n x^n$ for $x \in K$ with $|x| \leq 1$. If the coefficients a_n are not all zero then $f(x) = 0$ for at most $N(f)$ numbers in the closed unit disc of K.*

We can apply this to polynomials, which are power series with finitely many terms.

Example 10.2. By algebra, $f(X) = 1 + pX + X^2 + pX^5$ has at most 5 zeros in $\mathbf{Q}_p$. Since $N(f) = 2$, Strassmann's theorem tells us $f(X)$ has at most 2 zeros in $\mathbf{Z}_p$. The true number of zeros in $\mathbf{Z}_p$ is 0 if $p = 2$ ($a \in \mathbf{Z}_2 \Rightarrow f(a) \equiv 1, 2 \bmod 4$) or $p = 3$ ($a \in \mathbf{Z}_3 \Rightarrow f(a) \equiv 1, 2 \bmod 3$) and 2 if $p = 5$ (use Hensel's lemma with $a = 2$ and $a = 3$).

Example 10.3. Over $\mathbf{Q}_p$, $f(X) = 1 + X + pX^2$ has $N(f) = 1$ and thus at most 1 zero in $\mathbf{Z}_p$. There is a zero in $\mathbf{Z}_p$ from Hensel's lemma with approximate root -1. The second root of $f(X)$ in $\mathbf{Q}_p$ lies outside $\mathbf{Z}_p$.

Example 10.4. Strassmann's theorem provides a second proof that the p-adic exponential function is injective on $p\mathbf{Z}_p$ for $p \neq 2$. View e^x on $p\mathbf{Z}_p$ as e^{px} on $\mathbf{Z}_p$. The power series $e^{px} = \sum_{n\geq 0}(p^n/n!)x^n$ has coefficients $p^n/n!$ that tend to 0. For each $y \in \mathbf{Z}_p$ set $f_y(x) = e^{px} - e^{py} = (1 - e^{py}) + px + \sum_{n\geq 2}(p^n/n!)x^n$ for $x \in \mathbf{Z}_p$. Since $|e^{py} - 1|_p \leq \max_{n\geq 1} |p^n/n!|_p$ when $|y|_p \leq 1$ and $|p^n/n!|_p < 1/p$ for $n \geq 2$ by the proof of Theorem 4.5, $N(f_y) = 1$. An obvious solution to $f_y(x) = 0$ is $x = y$, so by Strassmann's theorem the only solution of $e^{px} = e^{py}$ for $x \in \mathbf{Z}_p$ is $x = y$. A similar argument shows e^x on $4\mathbf{Z}_2$ is injective.

Strassmann's theorem can be regarded as an analogue for nonarchimedean power series of bounding the number of roots of a polynomial over a field by the degree of the polynomial. In the polynomial theorem the key idea is to factor out $x - \alpha$ if α is a root, which lowers the degree of the polynomial by one, and the proof of Strassmann's theorem will have a step just like this where the value of $N(f)$ drops by one after removing a factor corresponding to a root (if one exists). When dealing with power series rather than polynomials we have to be a little more careful at the factoring step, but we have actually done most of that step already in Theorem 7.6.

Proof. To prove Theorem 10.1 we will use induction on $N(f)$.

When $N(f) = 0$, $|a_n| < |a_0|$ for all $n \geq 1$, so $a_0 \neq 0$ and $\max_{n\geq 1} |a_n| < |a_0|$ because the a_n's tend to 0. For $x \in K$ with $|x| \leq 1$,

$$\left|\sum_{n\geq 1} a_n x^n\right| \leq \max_{n\geq 1} |a_n x^n| \leq \max_{n\geq 1} |a_n| < |a_0|,$$

so by the strong triangle inequality $|f(x)| = |a_0 + \sum_{n\geq 1} a_n x^n| = |a_0| > 0$. Thus f has no zero in the closed unit disc of K.

Suppose $N \geq 1$ and the theorem is proved for all power series $g(x)$ over K converging on the closed unit disc of K with $N(g) < N$. If $N(f) = N$ and f has no zero in the closed unit disc of K then we are done ($0 < N$). If f has a zero $\alpha \in K$ with $|\alpha| \leq 1$, then we will show

$$f(x) = (x - \alpha)g(x) \tag{10.1}$$

for a power series g converging on the closed unit disc of K and $N(g) = N(f) - 1 = N - 1$, so by induction g has at most $N - 1$ zeros in the closed unit disc of K. Then by (10.1), for

$x \in K$ with $|x| \leq 1$ we have $f(x) = 0$ if and only if $x = \alpha$ or $g(x) = 0$, so the number of zeros of f in the closed unit disc of K is at most $1 + (N-1) = N$.

One way to prove (10.1) is that it follows directly from Theorem 7.6. For a second way,

$$\begin{aligned} f(x) &= f(x) - f(\alpha) \\ &= \sum_{k \geq 1} a_k(x^k - \alpha^k) \\ &= (x-\alpha)\sum_{k \geq 1} a_k(x^{k-1} + x^{k-1}\alpha + \ldots + x\alpha^{k-2} + \alpha^{k-1}) \\ &= (x-\alpha)\sum_{k \geq 1}\sum_{n=0}^{k-1} a_k x^n \alpha^{k-1-n} \\ &= (x-\alpha)\sum_{k \geq 1}\sum_{n \geq 0} c_{kn}, \end{aligned} \tag{10.2}$$

where $c_{kn} = a_k x^n \alpha^{k-1-n}$ for $n \leq k-1$ and $c_{kn} = 0$ for $n \geq k$. If $n \leq k-1$ then $|c_{kn}| = |a_k x^n \alpha^{k-1-n}| \leq |a_k|$ since $|x|, |\alpha| \leq 1$, and if $n > k$ then $|c_{kn}| = 0$. Since $a_k \to 0$ in K as $k \to \infty$, we get $c_{kn} \to 0$ as $\max(n,k) \to 0$, so we can swap the order of the sums in (10.2) by Theorem 2.9:

$$\begin{aligned} f(x) &= (x-\alpha)\sum_{n \geq 0}\sum_{k \geq 1} c_{kn} \\ &= (x-\alpha)\sum_{n \geq 0}\sum_{k > n} c_{kn} \ \text{ since } c_{kn} = 0 \text{ for } k \leq n \\ &= (x-\alpha)\sum_{n \geq 0}\sum_{k \geq n+1} a_k x^n \alpha^{k-1-n} \ \text{ by the definition of } c_{kn} \\ &= (x-\alpha)\sum_{n \geq 0}\left(\sum_{k \geq n+1} a_k \alpha^{k-n-1}\right) x^n \\ &= (x-\alpha)g(x), \end{aligned}$$

where $g(x) = \sum_{n \geq 0} b_n x^n$ with

$$b_n = \sum_{k \geq n+1} a_k \alpha^{k-n-1} = \sum_{k \geq 1} a_{n+k}\alpha^{k-1}, \tag{10.3}$$

Then $|b_n| \leq \max_{k \geq n+1} |a_k|$, so from $a_k \to 0$ in K as $k \to \infty$ we get $b_n \to 0$ in K as $n \to \infty$. Thus $g(x)$ is a power series converging on the closed unit disc of K and (10.1) holds.

It remains to prove $N(g) = N - 1$. Writing $g(x) = \sum_{n \geq 0} b_n x^n$ to retain the notation above, proving $N(g) = N-1$ means proving

$$|b_n| \leq |b_{N-1}| \text{ for all } n, \quad |b_n| < |b_{N-1}| \text{ for } n \geq N. \tag{10.4}$$

We will do this in two ways, both of which involve showing $|b_{N-1}| = |a_N|$ too.

Method 1 (Power series formula for b_n) Using the definition of b_n in (10.3) as a power series in α, where $|\alpha| \leq 1$,

$$|b_n| \leq \max_{k \geq n+1} |a_k| \leq |a_N|$$

for all n. Since $b_{N-1} = a_N + a_{N+1}\alpha + a_{N+2}\alpha^2 + \cdots$ and $|a_{N+j}\alpha^j| \leq |a_{N+j}| < |a_N|$ for $j \geq 1$, we have $|b_{N-1}| = |a_N|$. Thus $|b_n|$ among all n is maximized at $n = N-1$. For $n \geq N$, $|b_n| \leq \max_{k \geq n+1} |a_k| \leq \max_{k \geq N+1} |a_k| < |a_N| = |b_{N-1}|$, so $|b_n|$ among all n is maximized for the last time at $n = N-1$. Thus $N(g) = N-1$.

Method 2 (Recentering). We can use the formula for recentered coefficients in (7.3) twice to express the coefficients of g in terms of the coefficients of f in (10.1). Recentering f from 0 to α, $f(x) = \sum_{n \geq 0} a_n x^n = \sum_{n \geq 0} c_n (x-\alpha)^n$ where $c_n = \sum_{k \geq n} \binom{k}{n} a_k (\alpha - 0)^{k-n}$. Then $c_0 = \sum_{k \geq 0} a_k \alpha^k = f(\alpha) = 0$ (as expected) so

$$f(x) = (x-\alpha) \sum_{n \geq 1} c_n (x-\alpha)^{n-1} = (x-\alpha) \sum_{n \geq 0} c_{n+1} (x-\alpha)^n.$$

Thus $g(x) = \sum_{n \geq 0} c_{n+1}(x-\alpha)^n = \sum_{n \geq 0} b_n x^n$ where we recenter g from α to 0, so by (7.3)

$$\begin{aligned} b_n &= \sum_{\ell \geq n} \binom{\ell}{n} c_{\ell+1} (0-\alpha)^{\ell-n} \\ &= \sum_{\ell \geq n} \binom{\ell}{n} \left(\sum_{k \geq \ell+1} \binom{k}{\ell+1} a_k \alpha^{k-(\ell+1)} \right) (-\alpha)^{\ell-n} \\ &= \sum_{\ell \geq n} \sum_{k \geq \ell+1} \binom{\ell}{n} \binom{k}{\ell+1} a_k (-1)^{\ell-n} \alpha^{k-1-n} \\ &= \sum_{\ell \geq 0} \sum_{k \geq 0} d_{\ell k}, \end{aligned}$$

where $d_{\ell k} = \binom{\ell}{n}\binom{k}{\ell+1} a_k (-1)^{\ell-n} \alpha^{k-1-n}$ if $\ell \geq n$ and $k \geq \ell+1$, and $d_{\ell k} = 0$ otherwise. We want to change the order of this double series, which by Theorem 2.9 is justified if $d_{\ell k} \to 0$ as $\max(k, \ell) \to \infty$.

We have $|d_{\ell k}| \leq |a_k|$, since this is clear from the formula defining $d_{\ell k}$ if $\ell \geq n$ and $k > \ell$, and it's also clear in the other cases since $d_{\ell k} = 0$. From $|a_k| \to 0$ as $k \to \infty$ we get $|d_{\ell k}| \to 0$ as $\max(k, \ell) \to \infty$ (the only interesting case is when $\ell < k$, as $d_{\ell k} = 0$ otherwise), so we can exchange the order of the double series:

$$\begin{aligned} b_n &= \sum_{k \geq 0} \sum_{\ell \geq 0} d_{\ell k} \\ &= \sum_{k \geq n+1} \sum_{\ell=n}^{k-1} \binom{\ell}{n} \binom{k}{\ell+1} a_k (-1)^{\ell-n} \alpha^{k-1-n} \\ &= \sum_{k \geq n+1} \left(\sum_{\ell=n}^{k-1} \binom{\ell}{n} \binom{k}{\ell+1} (-1)^{\ell-n} \right) a_k \alpha^{k-1-n}. \end{aligned} \tag{10.5}$$

The finite sum in parentheses is an integer, so of absolute value at most 1. Also $|\alpha| \leq 1$, so $|b_n| \leq \max_{k \geq n+1} |a_k| \leq |a_N|$ for all n. We will next prove $|b_{N-1}| = |a_N|$ and $|b_n| < |a_N|$ for $n \geq N$.

The first term in (10.5) at $k = n+1$ is $\sum_{\ell=n}^{n}\binom{\ell}{n}\binom{n+1}{\ell+1}(-1)^{\ell-n}a_{n+1}\alpha^{(n+1)-1-n} = a_{n+1}$, so at $n = N-1$,

$$b_{N-1} = a_N + \sum_{k\geq N+1}\left(\sum_{\ell=N-1}^{k-1}\binom{\ell}{N-1}\binom{k}{\ell+1}(-1)^{\ell-(N-1)}\right)a_k\alpha^{k-1-(N-1)}.$$

Since $|a_k| < |a_N|$ when $k \geq N+1$, by the strong triangle inequality $|b_{N-1}| = |a_N|$. If $n \geq N$ then $|b_n| \leq \max_{k\geq N+1}|a_k| < |a_N|$. Thus $|b_n|$ is maximized for the last time at $n = N-1$. □

The formulas for b_n by the two methods, in (10.3) and (10.5), are actually the same formula. To convert (10.5) into (10.3), rewrite (10.5) as a sum over $k \geq 1$:

$$\begin{aligned} b_n &= \sum_{k\geq 1}\left(\sum_{\ell=n}^{k+n-1}\binom{\ell}{n}\binom{k+n}{\ell+1}(-1)^{\ell-n}\right)a_{n+k}\alpha^{k-1} \\ &= \sum_{k\geq 1}\left(\sum_{\ell=0}^{k-1}\binom{\ell+n}{n}\binom{k+n}{\ell+n+1}(-1)^{\ell}\right)a_{n+k}\alpha^{k-1}. \end{aligned}$$

Comparing this to (10.3), we want to show the inner sum in parentheses is 1:

$$\begin{aligned} \sum_{\ell=0}^{k-1}\binom{\ell+n}{n}\binom{k+n}{\ell+n+1}(-1)^{\ell} &= \sum_{\ell=0}^{k-1}\frac{(\ell+n)!}{n!\ell!}\frac{(k+n)!}{(\ell+n+1)!(k-\ell-1)!}(-1)^{\ell} \\ &= \frac{(k+n)!}{n!}\sum_{\ell=0}^{k-1}\frac{1}{\ell!(k-\ell-1)!}\frac{(-1)^{\ell}}{\ell+n+1} \\ &= \frac{(k+n)!}{(k-1)!n!}\sum_{\ell=0}^{k-1}\frac{(k-1)!}{\ell!(k-\ell-1)!}\frac{(-1)^{\ell}}{\ell+n+1} \\ &= \frac{(k+n)!}{(k-1)!(n+1)!}\sum_{\ell=0}^{k-1}\binom{k-1}{\ell}\frac{(n+1)(-1)^{\ell}}{\ell+n+1} \\ &= \binom{k+n}{n+1}\sum_{\ell=0}^{k-1}\binom{k-1}{\ell}\frac{(n+1)(-1)^{\ell}}{\ell+n+1}. \end{aligned}$$

That this equals 1 was posed as a question on MathOverflow [9], where multiple proofs can be found (set m, n, and k in [9] equal to $k-1$, $n+1$, and ℓ here).

Remark 10.5. The number of roots of a polynomial over a field need not equal its degree, but equality does occur in degree 1: $ax+b = 0$ if and only if $x = -b/a$ (if $a \neq 0$). Similarly, if $N(f) = 1$ in Strassmann's theorem then there really is a root of $f(x)$ in the closed unit disc of K. This can be proved using a version of Hensel's lemma for power series.[6]

Strassmann's theorem extends to closed discs in K of radius different from 1 by a simple scaling argument, as follows.

[6]See Section 8 of https://kconrad.math.uconn.edu/blurbs/gradnumthy/hensel.pdf.

Corollary 10.6. *Let $f(x) = \sum_{n\geq 0} a_n x^n$ converge on the closed disc $\{x \in K : |x| \leq r\}$ where $r \in |K^\times|$. If the coefficients a_n are not all zero then the number of solutions to $f(x) = 0$ with $|x| \leq r$ is at most*

$$N_r(f) = \max\{N \geq 0 : |a_n|r^n \leq |a_N|r^N \text{ for all } n \geq 0\}.$$

The integer $N_r(f)$ exists since, by hypothesis, $|a_n|r^n \to 0$ as $n \to \infty$.

Proof. Let $c \in K$ with $|c| = r$ and set $g(x) = f(cx) = \sum_{n\geq 0} a_n c^n x^n$. This is a power series converging on the closed unit disc in K and $N(g) = N_r(f)$ since $|a_n c^n| = |a_n|r^n$. The zeros of f in $\{x \in K : |x| \leq r\}$ are in bijection with the zeros of g in $\{x \in K : |x| \leq 1\}$ by $x \mapsto x/c$, so we are done by Strassmann's theorem. □

For $c \in K^\times$ such that $|c| > 1$, the discs $\{x \in K : |x| \leq |c|^m\}$ as $m \to \infty$ through $\mathbf{Z}^+$ exhaust K. Therefore a power series $f(x)$ with coefficients in K that has an infinite radius of convergence on K has at most countably many zeros in K: each zero of $f(x)$ in K lies in one of the discs $\{x \in K : |x| \leq |c|^m\}$ for $m \in \mathbf{Z}^+$ and each of these discs contains finitely many zeros of $f(x)$.

Corollary 10.7. *If two power series converge on a closed disc $\{x \in K : |x| \leq r\}$ where $r \in |K^\times|$ and are equal infinitely often on this disc then they are equal everywhere on this disc.*

Proof. The difference of the two power series is a power series with infinitely many zeros in the disc and thus the difference has all coefficients equal to 0 by Corollary 10.6, so the difference is identically zero on the disc. □

While Strassmann's theorem doesn't look like anything in real analysis, it can be regarded as a non-archimedean analogue of a result from complex analysis: Rouché's theorem. In its most basic form, Rouché's theorem says that if $0 < R < R'$ and $f(z)$ and $g(z)$ are complex power series converging on the open disc $\{z \in \mathbf{C} : |z| < R'\}$ and $|f(z) - g(z)| < |f(z)|$ for all $z \in \mathbf{C}$ with $|z| = R$, then f and g have the same number of zeros in the open disc $\{z \in \mathbf{C} : |z| < R\}$. To make Strassmann's theorem resemble this, we introduce some notation.

Definition 10.8. For a power series $f(x) = \sum_{n\geq 0} a_n x^n$ with $a_n \to 0$, set $|f| = \max_{n\geq 0} |a_n|$.

Clearly $|f| \geq 0$ with equality if and only if $f = 0$, and easily $|f + g| \leq \max(|f|, |g|)$, and $|fg| \leq |f||g|$ by the coefficient formulas in Theorem 3.1. Surprisingly, the last inequality is actually an equality! Although it is not needed, we will prove it anyway.

Theorem 10.9. *For power series f and g over K with coefficients tending to 0, $|fg| = |f||g|$.*

Proof. If $f = 0$ or $g = 0$ then the equality is obvious, so we can assume f and g each have some nonzero coefficients: $|f| > 0$ and $|g| > 0$.

Write $f(x) = \sum_{m\geq 0} a_m x^m$ and $g(x) = \sum_{n\geq 0} b_n x^n$. Set $|f| = |a_M|$ with M maximal and $|g| = |b_N|$ with N maximal: $|a_m| < |a_M|$ for $m > M$ and $|b_n| < |b_N|$ for $n > N$. Since $|fg| \leq |f||g|$, to prove $|fg| = |f||g|$ we seek a coefficient in fg with absolute value $|f||g|$ and we will find it in degree $M + N$.

The coefficient of x^{M+N} in fg is $\sum_{m=0}^{M+N} a_m b_{M+N-m}$. The term in this sum at $m = M$ is $a_M b_N$. If $0 \leq m < M$ then $M + N - m > N$, so

$$|a_m b_{M+N-m}| = |a_m||b_{M+N-m}| \leq |f||b_{M+N-n}| = |a_M||b_{M+N-n}| < |a_M||b_N|.$$

For $M < m \leq M + N$,

$$|a_m b_{M+N-m}| = |a_m||b_{M+N-m}| \leq |a_m||g| = |a_m||b_N| < |a_M||b_N|.$$

Thus $|a_m b_{M+N-m}| < |a_M b_N|$ for $0 \leq m \leq M + N$ with $m \neq M$, so by the strong triangle inequality we get

$$\left|\sum_{m=0}^{M+N} a_m b_{M+N-m}\right| = |a_M b_N| = |a_M||b_N| = |f||g|. \qquad \square$$

Theorem 10.10. *Let f and g be nonzero power series over K with coefficients tending to 0. If $|f - g| < |f|$ then $N(g) = N(f)$.*

Proof. Write $f(x) = \sum_{n\geq 0} a_n x^n$ and $g(x) = \sum_{n\geq 0} b_n x^n$. Set $N = N(f)$, so $|f| = |a_N|$.

The inequality $|f - g| < |f|$ says $|a_n - b_n| < |a_N|$ for all n, so $|b_n| = |(b_n - a_n) + a_n| \leq \max(|b_n - a_n|, |a_n|) \leq |a_N|$. If $n > N$ then $|b_n| < |a_N|$ since $|b_n - a_n|$ and $|a_n|$ are both less than $|a_N|$. If $n = N$ then $|b_N| = |(b_N - a_N) + a_N| = |a_N|$ since $|b_N - a_N| < |a_N|$. Thus $|b_n| \leq |b_N|$ for all n and $|b_n| < |b_N|$ for $n > N$, so $N(g) = N = N(f)$. $\square$

Example 10.11. If $f(x) = \sum_{n\geq 0} a_n x^n$ is a nonzero power series and has coefficients tending to 0, and $g(x)$ is the polynomial $\sum_{n=0}^{N(f)} a_n x^n$, then $|f - g| < |f|$ and trivially $N(g) = N(f)$.

For a nonzero polynomial, bounding the number of its roots by its degree turns out to be sharp: if $f(x)$ is a polynomial of degree d with coefficients in a field F, there is a field E containing F in which $f(x)$ has d zeros, if the zeros are counted with multiplicity. In a similar way the bound in Strassmann's theorem is sharp: there is a complete field $L \supset K$ in which $f(x)$ has $N(f)$ zeros in the closed unit disc of L (if zeros are counted with multiplicity). Therefore Theorem 10.10 is saying that if f and g are nonzero power series converging on the closed unit disc of K and $|f-g| < |f|$ then f and g have the same number of zeros in the closed unit disc of some complete field $L \supset K$, which is similar to Rouché's theorem.

References

[1] A-L. Cauchy, "Cours d'Analyse de l'École Royale Polytechnique; 1re Partie. Analyse Algébrique," Debure Frères, Paris, 1821. URL https://books.google.com/books?id=UrT0KsbDmDwC.

[2] F. Gouvea, "p-Adic Numbers: An Introduction," 2nd ed., Springer-Verlag, 1997.

[3] J. Hadamard, *Sur le rayon de convergence des séries ordonnées suivant les puissances d'une variable*, pp. 3–6 in "Oeuvres de Jacques Hadamard," vol. 1, CNRS Paris, 1968.

[4] K. Hensel, Über die arithmetischen Eigenschaften der algebraischen und transzendenten Zahlen, *Jahresbericht der Deutschen Mathematiker-Vereinigung* **14** (1905), 545–558. URL https://eudml.org/doc/144990.

[5] S. Katok, "p-adic Analysis Compared with Real," Amer. Math. Society, 2007.

[6] K. Knopp, "Theory and Application of Infinite Series," Dover, 1990.

[7] N. Koblitz, "p-adic Numbers, p-adic Analysis, and Zeta-functions," 2nd ed., Springer-Verlag, 1984.

[8] R. Strassmann, Über den Wertevorrat von Potenzreihen im Gebiet der $\mathfrak{p}$-adischen Zahlen, *J. Reine Angew. Math.* **159** (1930), 13–28. URL https://eudml.org/doc/149651.

[9] http://mathoverflow.net/questions/193611/binomial-coefficient-identity/193647.