P-adic Weierstrass Factor Theorem

Aleksei Lopatin

May 2026

Abstract

In this paper, we will go over the p-adic analogue of Weierstrass's Factor Theorem in the field of Newton Polygons.

1 Introduction

We will begin by explaining the theory of Newton Polygons in essence what they are and various concepts around them alongside motivations for using. We will introduce key lemmas that center around the construction crucial to the p-adic analogue Weierstrass Factor Theorem, including the "Translation Lemma" below:

Lemma 1.1. *Suppose λ_1 is the first slope of $N(f)$, $c \in \Omega$, and $v_p(c) = \lambda \leq \lambda_1$. Suppose that f converges on the closed disk $D(p^\lambda)$, and let*

$$g(x) = (1 - cx)f(x)$$

Then $N(g)$ is obtained by translating $N(f)$ by $(1, \lambda)$ and joining $(0, 0)$ and $(1, \lambda)$.

Suppose finally that f has last slope λ_f. Then $f(x)$ converges on the closed disk $D(p^\lambda f)$ if and only if $g(x)$ does.

After building several more lemmas on showing the "existence" of a construction, we will prove by induction the p-adic analogue of the Weierstrass Factor Theorem below

Theorem 1.2. *Let $f(x) = 1 + \sum_{i=0}^{\infty} a_i x^i$, and suppose that f converges on the closed disk $D(p^\lambda)$. Let N be the total length of all segments $N(f)$ of slope $\leq \lambda$. We implicitly assume that N is not infinity.*

Then there exists a unique polynomial $h(x) \in 1 + x\mathbb{C}_p[x]$ of degree N and $g(x) \in 1 + x\mathbb{C}_p[x]$ convergent and non-zero in $D(p^\lambda)$ such that

$$h(x) = f(x)g(x)$$

Moreover, $N(h)$ is equal to the part of $N(f)$ between 0 and $(N, v_p(a_N))$.

2 Newton Polygons

Typically we work in Ω, the complex p-adic field, for these Polygons. Moreover, $v_p(a)$ is defined as the highest power of p that divides a. A less colloquial definition is

$$v_p(a) = -\log_p |\cdot|$$

where $|\cdot|$ is the p-adic metric. Recall that a Newton Polygon $N(f)$ is

Definition 2.1. The Newton Polygon $N(f)$ of f is the lower convex hull of the points $(0,0)$,$(1, v_p(a_1), \cdots (n, v_p(a_n)))$. That is, it's the highest polygonal line such that all the points $(i, v_p(a_i))$ lie on or above it.

A rather abstract way to understand this is to imagine putting nails at each point, and then letting a rubberband stretch around them. This rubberband touches the boundary of the convex hull of the points, the outside not inside. Then, the convex hull of the points $(i, v_p(a_i))$ makes two paths from $(0, v_p(a_0))$ to $(n, v_p(a_n))$, the one that lies below all the points and the one that lies above all the points. By convention, we take the one that lies below all the points.

Example. Suppose that we have some polynomial of the form in taking a general prime p:

$$f(x) = 1 + px + p^2x^2 + px^3 + p^4x^4$$

We look at each x^i, find the corresponding $v_p(a_i)$ and plot these points. Here, we plot $(0,0)$,$(1,1)$,$(2,2)$, $(3,1)$,$(4,4)$.

The polygonal line is the path along $(0, v_p(a_0))$ to $(i, v_p(a_i))$ that goes both above all the points and below all the points. We then take the one that goes below all the points.

Example. A less superficial example is taking the polynomial $f(x) = 2x^2+4x+1$ with $p = 2$. We plot the points $(0,0)$,$(1,2)$, and $(2,1)$ below:

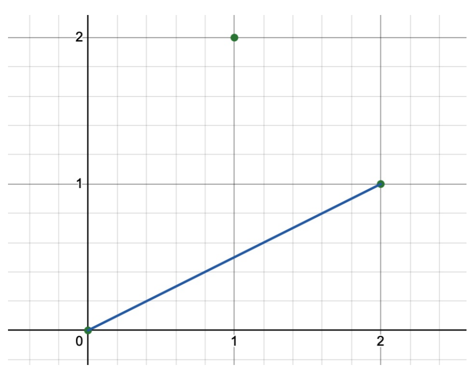

Figure 1: Example

2.1 Motivational Interlude

We may ask what is the point of using Newton Polygons. A Newton Polygon is just the construction of certain points from a polynomial. Why exactly do these points carry any useful information?

The next lemma will answer this question nicely, but we can provide a rough answer before. Suppose I have a polynomial in $\mathbb{R}$ with roots 3 and 5. I may wonder about the coefficients of this polynomial, which can be found through Vieta's Relations. But, what happens when I take a large number of roots. All I know is that each coefficient is the sum of x of these roots multiplied. This does not give any concrete information about these roots however.

In the world of p-adics, the situation is different. It is different because we really only care about the v_p's. Effectively in Q_2 (for instance), 3 and 5 are the same. But that is not all. The most important thing is that v_p is a super nice function. It especially has the property that

$$v_p(a+b) \geq \min(v_p(a), v_p(b))$$

Translating to the Vieta's sense, this is good because then we have some idea of what $v_p(a_i)$ becomes. It gives us some concrete information on these coefficients, and it enables us to go back and forth. As you'll soon see, Newton Polygons cleanly encase this information.

Definition 2.2. The length of a slope s of a segment in $N(f)$ is defined to be the horizontal projection of the segment onto the x-axis.

Theorem 2.3. *Let $f(x) \in \mathbb{Q}_p[x]$ be a polynomial. Suppose its Newton Polygon consists of segments of slope s_1 to s_m and length ℓ_1 to ℓ_m that correspond to their respective slopes. Then for all i in $[1, m]m$ the polynomial $f(x)$ has exactly ℓ_i roots with v_p $-s_i$.*

Proof. The main idea of the proof is using Vieta's formulas to compare the v_p of the roots. The proof can be found here: [RS26]. ■

Funnily enough, there is an analogue of this theorem in Ω, stated below:

Lemma 2.4. *Let $f(x) \in 1 + \Omega[[x]]$. Suppose that there is a segment in the Newton Polygon of f with finite length N and slope λ, then f has exactly N roots of valuation $-\lambda$.*

Going further, we may be in situations where we would like to consider "infinite" polynomials, effectively power series, so we have the following definition.

Definition 2.5. A power series $f(x)$ in $1 + x\Omega[[x]]$ is $f(x) = 1 + \sum_{i=1}^{\infty} a_i x^i$.

Example. Consider the power series $1 + \sum_{i=1}^{\infty} p^2 x^i$. For instance, this looks like

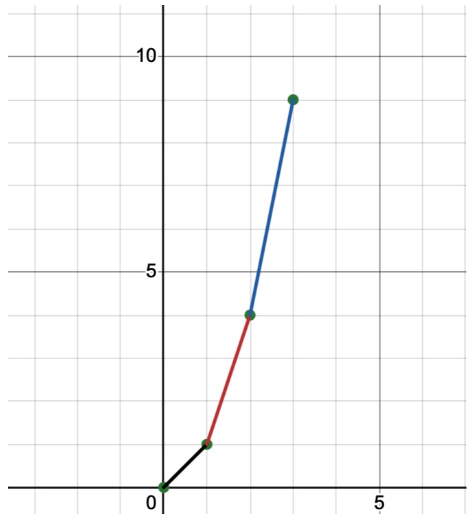

Figure 2: Power Series Example

The Newton Polygon of a power series may contain infinitely many segments or a segment of infinite length, which deserves a new definition.

Definition 2.6. For each $n \geq 1$, let $f_n(x) = 1 + \sum_{i=0}^{n} a_i x^i$. Then we define the Newton Polygon of f to be the limit of the $N(f_n)$ as $n \to \infty$. There are three cases of interest to us:

- $N(f)$ has infinitely many segments of finite length.
- $N(f)$ has finitely many segments, and the last segment, which is infinitely long, has infinitely many of the points $(i, v_p(a_i)$ on it.
- $N(f)$ has finitely many segments, and the last segment, which is infinitely long, has finitely many of the points $(i, v_p(a_i))$ on it.

We may ask what the radius of convergence of a polynomial f may be, which the following lemma answers.

Lemma 2.7. *Let $b = \sup_i \lambda_i$, the supremum over the slopes λ_i of a polynomial f. Then the radius of convergence of f is p^b.*

Moreover, f converges on the boundary of the disk $D(p^b)$ if and only if we are in case 3 above and the distance between the final slope and $(i, v_p(a_i))$ goes to infinity with i.

Proof. We give a sketch where the rest of the details can be found in [BER].

Suppose that we can bound $v_p(x)$ by $-b$ in that $v_p(x) > -b$, then there exists a value c such that $v_p(x) = -c$ and $c < b$. Now we consider the v_p of a general variable of the power series

$$v_p(a_i x^i) = v_p(a_i) + i v_p(x) = v_p(a_i) - ic$$

Notice that $v_p(a_i) - ic$ is clearly monotonic and strictly increasing, so we are done. The converse direction is similar and the final claim is quite clear from this. ■

This begs the question of what we can do to a Newton Polygon while still having convergence. It turns out that translation upholds convergence with a few conditions. The following lemma is quite key.

Lemma 2.8. *Suppose λ_1 is the first slope of $N(f)$, $c \in \Omega$, and $v_p(c) = \lambda \leq \lambda_1$. Suppose that f converges on the closed disk $D(p^\lambda)$, and let*

$$g(x) = (1 - cx)f(x)$$

Then $N(g)$ is obtained by translating $N(f)$ by $(1, \lambda)$ and joining $(0, 0)$ and $(1, \lambda)$.

Suppose finally that f has last slope λ_f. Then $f(x)$ converges on the closed disk $D(p^\lambda f)$ if and only if $g(x)$ does.

Proof. By a mapping of $x \to \frac{x}{c}$, we can assume that $c = 1$ and so $\lambda = 0$. Let $g(x) = 1 + \sum_{i=0}^{\infty} b_i x^i$ in a general form. We know what the b_i are by definition:

$$g(x) = f(x) - xf(x) = 1 + \sum_{i=0}^{\infty} a_i x^i - x(1 + \sum_{i=0}^{\infty} a_i x^i) = 1 + \sum_{i=0}^{\infty} (a_i - a_{i-1}) x^i$$

where the a_{i-1} comes from the fact that we use the exponent x^{i-1} instead of x^i (because we have a factor of x outside).

It remains to show that $N(g)$ is $N(f)$ translated one unit to the right, with an extra segment between the $(0, 0)$ and the endpoint.

Because $v_p(b_i) \geq v_p(a_{i-1})$, the points $(i, v_p(b_i))$ all lie above the Newton polygon of $N(f)$ translated one to the right (since the $(i-1, v_p(a_{i-1}))$ make the Newton polygon of $N(f)$ translated one to the right). Let us assume that $(i-1, v_p(a_{i-1}))$ is a corner of $N(f)$. Then $v_p(a_i) > v_p(a_{i-1})$, so $v_p(b_i) = v_p(a_{i-1})$ and $(i, v_p(b_i)$ is a corner of $N(g)$.

Hence, $N(g)$ is as claimed in the lemma, but we need to verify the last segment. The above lines imply that the last slopes satisfy $\lambda_g \geq \lambda_f$. We wish to show equality, so for the sake of contradiction, suppose $\lambda_g > \lambda_f$.

For a large i, the point $(i+1, v_p(a_i))$ lies below the Newton Polygon of g. But, $v_p(a_{i+1}) = v_p(a_i + b_{i+1}) = v_p(a_i)$. We can induct to get that $v_p(a_j) = v_p(a_i)$ for all $j \geq i$. But, then this Newton Polygon does not converge in $D(p^\lambda) = D(p^0) = D(1)$. Since we have a contradiction, this implies $\lambda_g = \lambda_f$. ■

Example. Recall the Newton Polygon of $2x^2 + 4x + 1$. Consider the Newton Polygon of $-2x^3 - 2x^2 + 3x + 1$ or $(1 - x)(x^2 + 4x + 1)$ as below

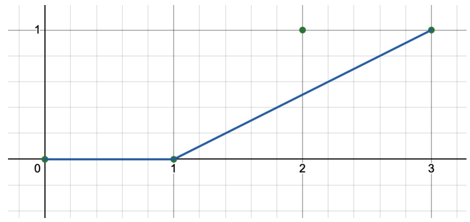

Figure 3: Translation Example

It indeed is as if the Newton Polygon of $2x^2 + 4x + 1$ is translate left by $(1, 0)$, which matches the fact that $v_2(c) = 0$.

3 Proving the Factor Theorem

We are now ready to build up the lemmas that will be involved in the proof of the p-adic analogue of the Weierstrass Factor Theorem.

The first lemma we will need is an existence proof that gives us one part of the construction for the Factor Theorem.

Lemma 3.1. *Suppose that $f(x)$'s first slope is λ_1, and that f has at least two distinct slopes. Then there exists $y \in \Omega$ with $v_p \frac{1}{y} = \lambda_1$ and $f(y) = 0$.*

Proof. A proof is given at [Tho11]. ∎

Eventually, our goal is to construct a polynomial of the form $(1 - \frac{x}{\alpha})^{-1} f(x)$. This is because $(1 - \frac{x}{\alpha})^{-1}$ can be re-arranged with the "Translation" lemma and by a geometric series:

$$\frac{1}{1 - \frac{x}{\alpha}} = 1 + \sum_{i=1}^{\infty} (\frac{x}{\alpha})^i$$

which makes it easier to come up with a power series for $g(x)$.

We may run into convergence issues, which the following lemma helps to settle.

Lemma 3.2. *Suppose that $f(\alpha) = 0$, and that*

$$g(x) = (1 - \frac{x}{\alpha})^{-1} f(x) = (1 + \frac{x}{\alpha} + \frac{x^2}{\alpha^2} + \cdots) f(x)$$

Then $g(x)$ converges on $D(|\alpha|)$.

Proof. Let $f_n(x)$ be the generic power series of form $1 + \sum_{i=1}^{n} a_i x^i$. Similarly, let $g(x) = 1 + \sum_{i=1}^{\infty} b_i x^i$. We directly solve for b_i using the form of $g(x)$ in the lemma above:

$$b_i = a_i + \frac{a_{i-1}}{\alpha} + \cdots + \frac{a_1}{\alpha^{i-1}} + \frac{1}{\alpha^i}$$

Hence, taking $g(\alpha)$, we have $b_i \alpha^i = f_i(\alpha)$. However, clearly as $i \to \infty$, $|f_i(\alpha)| \to 0$ as $f(x)$ is convergent. This implies that $g(\alpha)$ converges, so we are done. ∎

We eventually want to prove the Factor Theorem by induction, so we handle the base case below. This corresponds to the $h(x)$ in $h(x) = f(x)g(x)$ being 1.

Lemma 3.3. *There exists a power series $g(x) \in 1+x\Omega[[x]]$ such that $f(x)g(x) = 1$. Suppose that the slopes of the Newton Polygon of f are strictly greater than λ; Then the same is true of g. In particular, f and g both converge int he closed disk $D(p^\lambda)$ and are non-zero there.*

Proof. Let $g(x) = 1 + \sum_{i=0}^{\infty} b_i x^i$. We make a similar form for $f(x)$.

$$f(x)g(x) = (1 + a_0 + a_1 x + a_2 x^2 + \cdots)(1 + b_0 + b_1 x + b_2 x^2 + \cdots)$$

Consider the coefficient of x^i after the multiplication. This coefficient can come from a_i in that $a_i x^i \cdot 1$, from $a_{i-1}b_1$ in that $a_{i-1}x^{i-1} \cdot b_1 x$, and so on until $b_{i-1}x^{i-1} \cdot a_1 x$ and $b_i x^i \cdot 1$. We must have this coefficient of x_i be 0, hence

$$0 = -(b_i + b_{i-1}a_1 + \cdots + b_1 a_{i-1} + a_i) \implies$$

$$b_i = -(b_{i-1}a_1 + \cdots + b_1 a_{i-1} + a_i)$$

We can then recursively solve for the b_i, and thus $g(x)$ exists as a power series.

For the second part, we can apply the "Translation" lemma to assume that $\lambda = 0$. The hypothesis is that for all i, $v_p(a_i) > 0$ and $v_p(a_i) \to \infty$ as $i \to \infty$. It suffices to show that the same behavior carries onto the b_i. By induction on the relationship for b_i above, we will find that $v_p b_i > 0$.

It remains to show that $v_p(b_i)$ tends to 0 as $i \to \infty$. Let $M > 0$, and choose m such that for all $i > m$, $v_p(a_i) > M$. Let $\epsilon = \min(v_p(a_1), \cdots v_p(a_m))$.

We claim that for all $i > nm$, $v_p(b_i) > \min(M, n\epsilon)$. Specifically, when $i > \frac{mM}{\epsilon}$, $v_p(b_i) > M$. We prove this by induction on n. $n = 0$ is trivial. Suppose we have shown this for $n-1$. Then, from the b_i expression above, we have

$$b_i = -(b_{i-1}a_1 + \cdots + b_{i-m}a_m + b_{i-m-1}a_{m+1} + \cdots + a_i)$$

Notice that $b_{i-1}a_1 + \cdots + b_{i-m}a_m$ corresponds to $v_p > \min(M, (n-1)\epsilon) + \epsilon$ by the $n-1$ hypothesis. Moreover, $b_{i-m-1}a_{m+1} + \cdots + a_i$ has $v_p > M$ by the hypothesis above. ■

We are now ready to prove the Factor Theorem through induction. Before we do so however, let us take a step back and understand why induction is a viable strategy apart from it being a notable proof method.

4 Interlude

Recall that the Weierstrass Factor Theorem (in the p-adic sense) is saying something quite general. It states that we can factor a general polynomial, so this is quite global. We don't enforce too many restrictions on these polynomials. In a sense, that gives us a bit of hope in that induction may work because induction is essentially letting you take a theorem higher and higher into more generalizations (in this case a higher degree). What we want is that these "generalizations" keep the structure that we initialized. Induction lets us do so (most of the time).

However, consider if these factors would need more care as in different variables, etc. This would make the "problem" more localized in a sense that these characteristics arise from one set polynomial, and if we change it largely (by inducting to higher dimension for instance), it may not hold the same properties, which is bad.

Hence, induction is a good idea to try because these polynomials that we describe below work in general cases, and we have some confidence that creating more polynomials will not alter the characteristics we enforced at the beginning.

Now, we go all in. The theorem is stated below for completeness.

Theorem 4.1. *Let $f(x) = 1 + \sum_{i=0}^{\infty} a_i x^i$, and suppose that f converges on the closed disk $D(p^\lambda)$. Let N be the total length of all segments $N(f)$ of slope $\leq \lambda$. We implicitly assume that N is not infinity.*

Then there exists a unique polynomial $h(x) \in 1 + x\mathbb{C}_p[x]$ of degree N and $g(x) \in 1 + x\mathbb{C}_p[x]$ convergent and non-zero in $D(p^\lambda)$ such that

$$h(x) = f(x)g(x)$$

Moreover, $N(h)$ is equal to the part of $N(f)$ between 0 and $(N, v_p(a_N))$.

Proof. The following proof is due to [Tho11].

The core of this proof is induction on N. The case $N = 0$ in which $h(x) = 1$ has been shown above. We assume that the theorem has been proven in the case $N-1$, and let $\lambda_1 \leq \lambda$ be the first slope of f. By a previous lemma, there exists $\alpha \in \Omega$ such that $v_p(\frac{1}{\alpha}) = \lambda_1$ and $f(\alpha) = 0$. We introduce $f_1(x) = (1-\frac{x}{\alpha})^{-1} f(x)$. By a previous lemma, $f_1(x)$ converges on the disk $D(|\alpha|) = D(p^{\lambda_1})$.

Let $c = \frac{1}{\alpha}$, so $f(x) = (1-cx)f_1(x)$. Let $f_1'(x)$ have first slope λ_1'. If this is less than λ_1, then $f_1(x$ has a root of slope λ_1', and so does $f(x)$.However, the previous lemma says that $f(x)$ has no zeroes in the disk $D(\lambda_1^-)$, so we have a contradiction and $\lambda_1' \geq \lambda_1$.

We now apply the "Translation" lemma, which lets us relate $N(f)$ and $N(f_1)$. This says that $N(f)$ is obtained from $N(f_1)$ by translating by $(1, \lambda_1)$ and joining the points $(0, 0)$ and $(1, \lambda_1)$. By induction, there exists $h_1(x)$ of degree $N-1$ (since we assume the theorem has been shown in the case of $N-1$). Moreover there is a power series $g(x)$ convergent in $D(p^\lambda)$ such that

$$h_1(x) = f_1(x)g(x)$$

Crucially, we set $h(x) = (1-cx)h_1(x)$ and we are done. ■

References

[BER] DANA BERMAN. Newton polygons and p-adic power series. *URL https://www.cs.mcgill.ca/ dberma3/notes/Newton$_P$olygons.pdf.*

[RS26] Simon Rubinstein-Salzedo. *P-adic Analysis*. 1 edition, 2026.

[Tho11] Jack A. Thorne. p-adic analysis, p-adic arithmetic. Lecture notes, P-adic Analysis, University of Connecticut, 2011. Available at `https://kconrad.math.uconn.edu/math5020f11/jackthornenotes.pdf`, accessed June 2026.