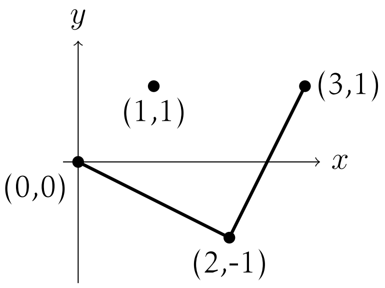
# NEWTON POLYGONS AND P-ADIC POWER SERIES

DANA BERMAN

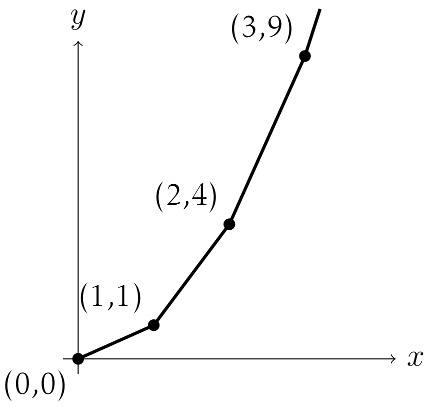
## 1. Introduction

In this paper we will study how Newton Polygons can be used to easily extract information about the roots of polynomials and power-series over $\Omega$. Given a polynomial $f(X) = \sum_{i=0}^{n} a_i X^i$ Newton Polygon is defined as the convex hull[1] of the sets of points $\{(i, \nu_p(a_i))\}_{i=1}^{n}$. In the case of a power series, the Newton Polygon is defined in the same way but with $n = \infty$. Consider the following examples:

(1) Consider the polynomial $f(X) = 1 + pX + p^{-1}X^2 + pX^3$. The Newton polygon will be as follows:




(2) Secondly, suppose we have the power series $f(X) = \sum_{i=0}^{\infty} p^{i^2} X^i$, then our polygon will be as follows:




*Date*: Thursday 10th August, 2017.

[1]The slopes of the segments of the polygon strictly increase as one moves along the x-axis.

(3) Finally, consider the power series $f(X) = \sum_{i=0}^{\infty} pX^i$. Then notice that the associated Newton polygon is exactly the non-negative side of the x-axis.

In remainder of the paper, if $f(X)$ is a polynomial, than it will be assumed that $f(X) \in 1 + \Omega[X]$. If $f(X)$ is a power series, than it will be assumed that $f(X) \in 1 + \Omega[[X]]$. Note that under these assumptions, 0 may not be a root of any polynomial or power series. If $\xi$ is a root of $f(X)$, it therefore follows that $\nu_p(\xi)$ is finite.

## 2. Polynomials

Here we shall prove a lemma which relates Newton polygons to polynomials in $\Omega[X]$. This lemma shall be generalized to power series in the next section.

**Lemma 2.1.** *Let $f(X) \in 1 + \Omega[X]$ and let $\xi_1, \ldots, \xi_n$ be the roots of the polynomial.*[2] *Denote $-\nu_p(\xi_i)$ by $\lambda_i$ for each $i \in \{1, 2, \ldots, n\}$ and suppose that one of the segments of the polygon has slope $\lambda$ and length $l$. Then there exists a set $S \subseteq \{1, \ldots, n\}$ of size $l$ such that $\lambda_j = \lambda$ for all $j \in S$.*

*Proof.* Without loss of generality, we write

$$f(X) = \prod_{i=1}^{n} (1 - X/\xi_i)$$

such that $\{\lambda_i = -\nu_p(\xi_i)\}_{i=1}^{n}$ is listed in increasing order.

Set $\lambda_{n+1} = \infty$ and let $r \leq n$ be the least natural number such that $\lambda_1 = \cdots = \lambda_r < \lambda_{r+1}$. For the proof, it is sufficient to show that the first segment of the Newton polygon is indeed from $(0, 0)$ to $(r, r \cdot \lambda_1)$. This would imply that the lemma holds true for the first segment of the polygon. Then it is easy to see by induction that the statement also holds true for every following segment of the polygon.

We now prove the above claim. We may write the function $f(X)$ as $1 + \sum_{i=1}^{n} a_i x^i$ for some $\{a_i\}_{i=1}^{n}$.

First, we claim that for all $i \in \{1, \ldots, r\}$, the points $(i, \nu_p(a_i))$ lie above or on the segment with ends at $(0, 0)$ and $(n, n \cdot \lambda_1)$. For each such $i$, we may express $a_i$ as the sum of all possible products combining $i$ many distinct elements from the set $\{1/\xi_1, \ldots, 1/\xi_n\}$. It follows that for each $i$, we have

$$\nu_p(a_i) \geq \min_{k_1,\ldots,k_i \in \{1,\ldots,n\}} -\nu_p\left(\xi_{k_1} \cdot \xi_{k_2} \cdots \xi_{k_i}\right) \geq i \cdot \lambda_1$$

We now claim that for $i \in \{r + 1, \ldots, n\}$, the points $(i, \nu_p(a_i))$ lie strictly above the segment with ends at $(0, 0)$ and $(n, n \cdot \lambda_1)$. This is because in this case, we

[2] Note that these are not necessarily distinct, but such that $f(X) = (1 - X/\xi_1) \cdot (1 - X/\xi_2) \cdots (1 - X/\xi_n)$.

have

$$\nu_p(a_i) \geq \min_{k_1,\ldots,k_i \in \{1,\ldots,n\}} -\nu_p\left(\xi_{k_1}\cdots\xi_{k_i}\right) \geq (i-1)\cdot\lambda_1 + \lambda_{r+1} > i\lambda_1.$$

Finally, we claim that $\nu_p(a_r) = r\cdot\lambda_1$. We have

$$\nu_p(a_r) = -\nu_p\left[\sum_{k_1,\ldots,k_r\in\{1,\ldots,n\}} (\xi_{k_1}\cdots\xi_{k_r})\right]$$

The first term of the abive sum, $\xi_1\cdot\xi_2\cdots\xi_r$ has valuation $r\cdot\lambda_1$. Every other term has some element $\xi_k$ with $k\geq r+1$, and therefore has valuation strictly greater than $r\cdot\lambda_1$. It follows by the isosceles triangle principle that $\nu_p(a_r) = r\cdot\lambda_1$.

Putting the above claims, together, it follows that the first segement of the polygon has indeed ends $(0,0)$ to $(r, r\cdot\lambda_1)$ as recquired. □

## 3. Power Series

*Remark* 1. Note that there may be three types of Newton polygons. There may be infinitely many finite segments, finitely many segments such that the last segment hits points which are arbitrarily far out or finitely many segments such that the last segment does not hit points which are arbitrarily far our (see 3). Consider where each of these converge.[3]

**Lemma 3.1.** *Let $f(X) = 1+\sum_{i=1}^{\infty} a_i X^i$. Let $c\in\Omega$ be such that $\nu_p(c) = \lambda \leq \lambda_1$ where $\lambda_1$ is the first slope in the Newton polygon of $f$. Assume furthermore that $f$ converges in the closed disc $D(p^\lambda)$[4] and define the power series*

$$g(X) = (1-cX)f(X)$$

*Then the Newton polygon of $g$ is exactly the segment $(0,0)$ to $(1,\lambda)$ joined by the newton polygon of $f$. Moreover, $f$ and $g$ have the same radius of convergence.*

*Proof.* Note that by showing that $f$ and $g$ have the same final slope, denoted $\lambda_f$ in their respective Newton polygons, it follows that both converge in $D(\lambda_f^-)$. For the convergence part of the proof it is therefore sufficient to show that $f$ converges on the boundary if and only if $g$ converges on the boundary.

We begin by proving the lemma for the special case $c=1$, $\lambda=0$. We may write $g(X) = 1+\sum_{i=1}^{\infty} b_i X^i$ where $b_i = a_{i+1} - a_i$ for $i\in\mathbb{N}$ where $a_0 = 1$. It follows that

$$\nu_p(b_i) \geq \min\left\{\nu_p(a_i), \nu_p(a_{i-1})\right\}$$

with equality holding if and only if $\nu_p(a_i)\neq\nu_p(a_{i-1})$. It follows that the points $(i, b_{i+1})$ lie on or above the Newton polygon of $f$. Moreover, if $(i, a_i)$ is a vertex, then since we have assumed the first slope to be greater or equal to 0,[5] we must

[3]Consider the Newton polygon of $f(X) = 1+\sum_{i=1}^{\infty} p^i X^i$.
[4]Note that $f$ must converge in $D(\lambda_1^-)$.
[5]We have $0=\lambda\leq\lambda_1$.

have that $\nu_p(a_{i+1}) > \nu_p(a_i)$. It follows that if $a_i$ corresponds to a vertex, then $\nu_p(b_{i+1}) = \nu_p(a_i)$. The Newton polygon of $g$ must therefore be as described by the lemma until the last vertex, if such a vertex exists.

In the case where the Newton polygon of $f$ has a last vertex, it remains only to prove that the final slope of the Newton polygons $f$ and $g$, which we denote $\lambda_f$ and $\lambda_g$ respectively, are equal. By the above argument, we already know that $\lambda_g \geq \lambda_f$. Suppose for the sake of contradiction that $\lambda_g > \lambda_f$. Then for some large $i$, the point $(i+1, \nu_p(a_i))$ would lie below the Newton polygon of $g$. It follows that $\nu_p(a_{i+1}) = \nu_p(a_i + b_{i+1}) = \nu_p(a_i)$. By induction, we obtain that $\nu_p(a_j) = \nu_p(a_i)$ for all $j \geq i$. But this is a contradiction to the assumption that $f(X)$ converges in $D(p^\lambda) = D(1)$.[6]

Since the points $(i, b_{i+1})$ lie on or above the Newton polygon of $f$, it is clear that if $f$ converges in some disc, then so does $g$. It remains only to show that if $g(X)$ converges in $D(p^{\lambda_f})$ then so does $f(X)$. This can be shown by a similar argument as above.

Recall that the proof so far only covers the special case $c = 1$, $\lambda = 0$. If we are not in that case, then the lemma may be applied to $f_1(X) = f(X/c)$ and $g_1(X) = (1 - X)f_1(X)$. One may then write $g(X) = g_1(cX)$. □

**Lemma 3.2.** *Let $f = 1 + \sum_{i=1}^{\infty} a_i X^i$ have Newton polygon with first slope $\lambda_1$ and suppose the for some $i$, $(i, \nu_p(a_i))$ lies on the first segment. Suppose furthermore that $f$ converges in $D(p^{\lambda_1})$. Then there exists an $x \in \Omega$ such that $f(x) = 0$ and $\nu_p(x) = -\lambda_1$.*

*Proof.* As in the proof of the previous lemma, we begin by considering the case $\lambda_1 = 0$. Note that this implies that $\nu_p(a_i) \geq 0$ for all $i \in \mathbb{N}$ and that $\nu_p(a_i) \to \infty$ since $f(X)$ in the disk $D(1)$ by assumption. We know that there exists some $i$ such that $\nu_p(a_i) = 0$, we therefore let $N$ be the maximal such $i$.

For any $n \in \mathbb{N}$, consider the polynomial $f_n(X) = 1 + \sum_{i=1}^{n} a_i X_i$. By Lemma 2.1, we know that for $n \geq N$, each such polynomial has exactly $N$ zeros which we denote $S_n = \left\{\xi_1^{(n)}, \dots \xi_N^{(n)}\right\}$, each with valuation (order) $\lambda_1 = 0$.

We know build a sequence $(x_n)_{n=N}^{\infty}$ and claim that it converges to some $x \in \Omega$ such that $f(x) = 0$ and $\nu_p(x) = -\lambda_1 = 0$ thereby proving the lemma for the case $\lambda_1 = 0$.

The sequence is constructed as follows: let $x_N = \xi_1^(N)$. The for each $n \geq N$, we define $x_{n+1} = \xi_j^{(n+1)}$ where $j \in \{1, \dots, N\}$ is such that $|x_n - x_{n+1}|_p$ is minimal. In order to prove that the sequence is convergent, it is sufficient to show that it is Cauchy. To see this, first note that for any $n \geq N$ one has

$$|f_{n+1}(x_n) - f_n(x_n)|_p = |f_{n+1}(x_n)|_p = \prod_{\xi \text{ is a root of } f_{n+1}} \left|1 - \frac{x_n}{\xi}\right|_p \tag{1}$$

[6] Consider $x = 1$.

The Newton polygon of each $f_{n+1}$ tells us that, other than the roots in $S_{n+1}$, i.e. the terms $\xi_j^{(n+1)}$, any root has valutaion strictly greater than zero, and therefore has norm strictly less than 1. It follows that for each root $\xi \notin S_{n+1}$, one has $|x_n/\xi|_p > 1$ and therefore $|1 - x_n/\xi|_p = 1$. Returning to equation 1, one can see that

$$|f_n(x_n) - f_n(x_{n+1})|_p = \prod_{i=1}^{N} \left|1 - \frac{x_n}{\xi_i^{(n+1)}}\right|_p = \prod_{i=1}^{N} \left|\xi_i^{(n+1)} - x_n\right|_p$$

since $\left|\xi_i^{(n+1)}\right| = 1$ for any $i \in \{1, \ldots, N\}$. Finally, it follows from our choice of $x_n$ that

$$|f_n(x_n) - f_n(x_{n+1})|_p \geq |x_{n+1} - x_n|_p^N$$

Using the above, we can show that we indeed have a Cauchy sequence since:

$$|x_{n+1} - x_n|_p^N \leq |f_n(x_n) - f_n(x_{n+1})|_p = \left|a_{n+1}x_n^{n+1}\right|_p = |a_{n+1}|_p \xrightarrow{n\to\infty} 0.$$

It follows that our sequence is Cauchy, and therefore $x_n \to x$ for some $x \in \Omega$. It is clear that $\nu_p(x) = 0$ and remains only to show that $f(x) = 0$. To see this, we write

$$|f_n(x)|_p = |f_n(x) - f_n(x_n)|_p = \left|\sum_{i=1}^{n} a_i(x^i - x_n^i)\right|_p$$

Now note that $|a_i|_p \leq 1$ and $\left|\sum_{k=0}^{i-1} x^k x_n^{i-1-k}\right| \leq max_{k\in\{0,\ldots,i-1\}} \left|x^k x_n^{i-1-k}\right| = 1$. Furthermore, we have that $x^i - x_n^i = (x - x_n)\sum_{k=0}^{i-1} x^k x_n^{i-1-k}$ and therefore,

$$|f_n(x)|_p = \left|\sum_{i=1}^{n} a_i(x^i - x_n^i)\right|_p \leq |x - x_n| \xrightarrow{n\to\infty} 0.$$

We therefore have $\lim_{n\to\infty} f_n(x) = f(x) = 0$ as required.

For the general case, i.e. $\lambda_1 \neq 0$, we may pick $c \in \Omega$ such that $\nu_p(c) = \lambda_1$. We consider the function $g(X) = f(X/c)$ which satisfies the conditions with $\lambda_1 = 1$ and find $x$ such that $\nu_p(x) = 0$ and $g(x) = 0$. Then for the function $f$, $x/c$ satisfies the requirements of the lemma. □

**Lemma 3.3.** *Consider $f(X) = 1 + \sum_{i=1}^{\infty} a_i X^i$ and let $\alpha$ be such that $f(\alpha) = 0$. Then*

$$g(X) = \frac{f(X)}{1 - X/\alpha}$$

*converges in $D(|\alpha|_p)$*

*Proof.* By assumption, we know that $f(X)$ converges at $\alpha$ and therefore on $D(|\alpha|_p)$. Note also that dividing by $(1-X/\alpha)$ is equivalent to multiplying by $\sum_{i=0}^{\infty} X^i/\alpha^i$.

It follows that we may writes $g(X) = 1 + \sum_{i=1}^{\infty} b_i X^i$ where (definiting $a_0 = 1$)

$$b_i = \sum_{k=0}^{i} a_k / \alpha^{i-k}.$$

Then to see that $g(X)$ converges in $D(|\alpha|_p)$ we note that

$$\left|b_i \alpha^i\right| = \sum_{k=0}^{i} a_i \alpha^k = f_i(\alpha) \xrightarrow{i \to \infty} f(\alpha) = 0.$$

□

**Theorem 3.4** (p-adic Weierstrass Preparation Theorem). *Let $f(X) = 1 + \sum_{i=1}^{\infty} a_i X^i$ and suppose $f(X)$ converges on $D(p^\lambda)$. Then let $N$ be the total length of all segments of the Newton polygon of slope $\leq \lambda$ if it is finite. Otherwise, let $N$ be the greatest $i$ such that $(i, \nu_p(a_i))$ lies on the Newton polygon.*[7] *Then there exists a unique polynomial $h(X) \in 1 + X\Omega[X]$ of degree $N$ and a power series $g \in 1 + \Omega[[X]]$ convergent and non-zero in $D(p^\lambda)$ such that*

$$h(X) = g(X) f(X).$$

*Furthermore, the Newton polygon of $h$ is uniquely determined by those properties and coincides with the Newton polygon of $f$ up to $(N, \nu_p(a_N))$.*

*Proof.* As in Lemma 3.1 and Lemma 3.2, we reduce our problem to the case $\lambda = 0$.

We prove by induction on $N$. Consider first the base case $N = 0$. This case implies that we are looking for $g(X)$ such that

$$g(X) f(X) = 1$$

since $h$ has degree 0. We therefore already know that $g(X)$ is non-zero on $D(p^\lambda)$. We may write $g(X) = 1 + \sum_{i=1}^{\infty} b_i X^i$ and solve for the terms $b_i$:

$$g(X) f(X) = \left(1 + \sum_{i=1}^{\infty} a_i X^i\right)\left(1 + \sum_{i=1}^{\infty} b_i X^i\right) = 1 + \sum_{i=1}^{\infty}\left(\sum_{k=0}^{i} a_k b_{i-k}\right) X^i$$

It follows that for each $i \geq 1$, we have

$$\begin{aligned} & b_i + a_1 b_{i-1} + a_2 b_{i-2} + \cdots + a_{i-1} b_1 + a_i = 0 \\ \iff & b_i = -\left(a_1 b_{i-1} + a_2 b_{i-2} + \cdots + a_{i-1} b_1 + a_i\right) \end{aligned}$$

where $a_0 = b_0 = 1$. For the theorem in the case $N = 0$, it remains only to show that $g(X)$ must converge in $D(p^\lambda) = D(1)$. One notes that it is sufficient to show that $\nu_p(b_n) \to \infty$ as $n \to \infty$. To prove this, first note that $\nu_p(a_i) > \lambda = 0$ for all

[7] Note that this second case implies that there is an infinite segment with slope $\lambda$. Since we converge on the closed disc $D(p^\lambda)$, such a greatest $i$ must exist.

$i \geq 1$ by hypothesis. It is therefore easy to verify by induction that $\nu_p(b_i) > 0$ for all $i \geq 1$.

Futhermore, since $f(X)$ converges in $D(1)$, we know that $\nu_p(a_i) \to \infty$. We may therefore pick any natural number $M$ and find $m$ such that $\nu_p(a_i) \geq M$ whenever $i \geq m$.

We fix $M$ and claim that for any $n \in \mathbb{N}$, if $i > nm$ then $\nu_p(b_i) \geq \min\{M, n\varepsilon\}$ where

$$\varepsilon = \min_{1 \leq k \leq m} \{\nu_p(a_k)\}$$

Then letting $n, M \to \infty$, we have that $\nu_p(b_i) \to \infty$ as desired. We now prove our claim by induction. When $n = 0$, the claim is trivially true as it is equivalent to $\nu_p(b_i) \geq 0$. Suppose now that the claim is true for $n - 1$. Then let $i > nm$ and write

$$b_i = -\left(a_1 b_{i-1} + a_2 b_{i-2} + \cdots + a_m b_{i-m} + a_{m+1} b_{i-m-1} + \cdots + a_{i-1} b_1 + a_i\right).$$

We now consider each individual term of the form $a_k b_{i-k}$. First suppose that $k > m$, then $\nu_p(a_k n_{i-k}) \geq \nu_p(a_k) \geq M$. Suppose now that $k \leq m$. Then $\nu_p(a_k b_{i-k}) \geq \nu_p(b_{i-k}) + \varepsilon$ and by induction hypothesis, $\nu_p(a_k b_{i-k}) \geq \min\{M, (n-1)\varepsilon\} + \varepsilon \geq \min\{M, n\varepsilon\}$. This proves the claim, and hence the theorem for the case $N = 0$.

We now suppose $N \geq 1$ and assume that the theorem holds true for $N - 1$. By Lemma 3.2 we may find $\alpha \in \Omega$ such that $f(\alpha) = 0$ and $\nu_p(\alpha) = \lambda_1 \leq \lambda$ where $\lambda_1$ is the first slope of the Newton polygon. We now define

$$f_1(X) = \frac{f(X)}{1 - X/\alpha}.$$

Then by Lemma 3.3, $f_1(X)$ converges in $D(p^\lambda)$. Moreover, letting $c = 1/\alpha$, we have $f(X) = (1 - cX)f_1(X)$. Denote now the first slope of the Newton polygon of $f_1$ by $\eta_1$. We first claim that $\eta_1 \geq \lambda_1$. Suppose not, then by 3.2, there exists $\beta \in \Omega$ such that $f_1(\beta) = 0$ and $\nu_p(\beta) = -\eta_1$. It follows that $f(\beta) = 0$. But then, the first slope of the polygon of $f(X)$ would have to be strictly less than $\eta_1$ (easy to verify). This is clearly a contradiction, since the first slope is $\lambda_1$. It follows that we indeed have $\eta_1 \geq \lambda_1$.

We will now use Lemma 3.1 on $f(X) = (1 - cX)f_1(X)$. The Lemma tells us that the Newton polygon of $f_1(X)$ is exactly that of $f(X)$, minus the segment $(0, 0)$ to $(1, \lambda_1)$. Moreover, $f_1(X)$ converges on $D(p^\lambda)$.

We may now apply our induction hypothesis on $f_1(X)$ and we obtain:

$$h_1(X) = f_1(X)g(X) \implies h(X) = (1 - cX)h_1(X) = f(X)g(X).$$

where $h_1(X) \in 1 + \Omega[X]$ is of degree $N - 1$, and therefore $h(X)$ is as required.

Note that since the power series $g$ is left unchanged, it follows from the induction hypothesis that it is indeed non-zero in $D(p^\lambda)$.

We now prove that $h(X)$ is unique. Suppose $\tilde{h}(X) = f(X)\tilde{g}(X)$ where $g(X), \tilde{g}(X)$ converge in $D(p^\lambda)$ and $\tilde{h}(X), h_1(X) \in 1 + \Omega[X]$ are polynomials of degree $N$.

Then, since $h_1(X)$ is uniquely determined by the induction hypothesis,

$$\tilde{h}(X) = (1 - cX)f_1(X)\tilde{g}(X) = (1 - cX)h_1(X) = h(X).$$

$\square$

We now state the equivalent of our Lemma 2.1 for power series.

**Corollary 3.5.** *Let $f(X) \in 1 + \Omega[[X]]$. Suppose that there is a segment in the Newton polygon of $f$ which has finite length $N$ and slope $\lambda$, then $f$ has exactly $N$ roots (counting multiplicity) of valutaion $-\lambda$.*

Note that there are further conclusions that can be easily derived from the theorem. For instance, if $f$ converges everywhere, then it can be factored into an ifinite product of $(1 - X/r)$ over its roots. If a function converges everywhere and has no zeros, it must therefore be constant.