On a p-adic newton method

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Master's thesis
submitted to the faculty for Mathematics and Computer Engineering
Georg-August-Universität Göttingen

Author:
Paul Breiding

First supervisor:
Prof. Dr. Preda Mihailescu

Second supervisor:
Prof. Dr. Peter Bürgisser

Contents

1 Introduction **3**

2 Preliminaries **5**

2.1 The p-adic numbers . . . 5

2.1.1 Definition of $\mathbb{Z}_p$ and $\mathbb{Q}_p$ . . . 5

2.1.2 Extensions of $\mathbb{Q}_p$ . . . 8

2.2 Vector spaces and operator norms . . . 9

2.3 Ultrametric spaces . . . 13

2.3.1 Properties of ultrametric spaces . . . 13

2.3.2 Spherical completeness . . . 14

2.4 Hensel lifting . . . 15

2.5 Analytic functions . . . 17

2.6 Orthogonal sets . . . 23

2.6.1 Best approximation . . . 23

2.6.2 Orthogonal bases in K-Banach spaces . . . 24

2.7 The p-adic projective space . . . 28

3 The p-adic Newton method **32**

3.1 Multivariate Newton method . . . 32

3.2 Smale's γ-theorem: p-adic version . . . 34

4 The p-adic Newton method: projective version **39**

4.1 The projective Newton operator . . . 39

4.2 A higher derivative estimate . . . 47

4.3 A Lipschitz estimate . . . 52

4.4 Multivariate Hensel lifting . . . 53

1 Introduction

Consider the problem of finding a solution for a system of n polynomials in n variables over the $\mathbb{Z}_p$:

$$\begin{aligned} f_1(X_1, \ldots, X_n) &= 0 \\ &\vdots \\ f_n(X_1, \ldots, X_n) &= 0. \end{aligned}$$

If $f := (f_1, \ldots, f_n)$ and $X := (X_1, \ldots, X_n)$, one can write $f(X) = 0$ as well. In [7] Kuhlmann describes a multivariate Hensel lifting to lift a solution $f(x) = 0 \mod p$ to a solution in $\mathbb{Z}_p$. The rate of convergence towards a zero of f, as in the univariate Hensel lifting, is at least linear. However, the problem of solving the system $f(x) = 0$ modulo p is hard. Although there are algorithms that can solve $f(x) = 0 \mod p$ (see for instance [5, sec. 2.7]), the speed of these algorithms will go down very fast with n increasing. In fact, Fraenkel and Yesha in [6] give a proof that solving polynomial systems modulo p is NP-hard. While working in the field $\mathbb{F}_p$ allows to apply the classic model of a Turing machine, in [8] Maller and Whitehead develop for odd primes p a model of computation over the p-adics themselves. They follow an approach of Blum, Shub and Smale, who before in [3] developed a model of computation over arbitrary fields and rings, in particular $\mathbb{R}$ (BSS-Turing machine). As in the classic theory of computational complexity one can define classes such as P or NP in this model. Maller and Whitehead work with BSS-Turing machines over $\mathbb{Q}_p$ and conclude that solving polynomial systems over $\mathbb{Q}_p$ is NP-Hard.

The goal of this work therefore is not to establish an algorithm to solve polynomial systems over $\mathbb{Q}_p$ in general. We will describe a multivariate Hensel lifting, but take another approach than Kuhlmann, as we want to avoid solving systems modulo p.

For a short moment, consider the system $f(x) = 0$ over the complex numbers $\mathbb{C}$. Suppose $\zeta \in \mathbb{C}^n$ is a zero of f. In [2, sec. 8 and 14] Blum, Cucker, Shub and Smale give their definition of an 'approximate zero'. These are points $z \in \mathbb{C}^n$ that are close to ζ and have the property that if one applies the Newton algorithm to f and z, one receives a sequence with limit ζ and the rate of convergence is quadratic. In [4, sec. 15-17] Bürgisser and Cucker use the results of [2] for homogeneous polynomials f_i to develop a homotopy method. If f is a system that one wants to solve, the method takes as input another system g and $\xi \in \mathbb{C}^n$ such that $g(\xi) = 0$ and continues ξ along a path to $\zeta \in \mathbb{C}^n$ with $f(\zeta) = 0$. One question that arises is the following: Does this also work p-adically? Unfortunately, the answer is negative. The p-adics are totally disconnected, making it impossible to even find any path between two points. Nevertheless, if we try to translate the results of [2, sec. 8 and 14] to the p-adic world, some interesting results will appear. It turns out that as in the complex case we can compute neighborhoods of zeros where the Newton algorithm will converge quadratically fast. Surprisingly, the results will lead to a multivariate Hensel lifting that has a linear rate of convergence, just as the classical univariate case and as Kuhlmann's method. For a system $f(x) = 0$ the lifting takes a zero ξ of a system g that is 'close' to f as starting value. This method is inspired by the homotopy method in [4]. The advantage when compared to Kuhlmann is that we avoid solving $f(x) = 0 \mod p$. Nevertheless, our Hensel lifting only works if a solution $g(\xi) = 0$ close to f is known.

The first chapter deals with preliminaries, which will be needed throughout the whole work. A short introduction to the construction of the p-adic numbers and its algebraic extension will be made. Moreover, there are sections on operator norms and analytic functions. A concept of orthogonality will be introduced as well. We will give an explicit way to compute orthogonal bases in p-adic vector space and will define an ultrametric on the p-adic projective space. The second chapter introduces the Newton algorithm to the p-adic world and translates the results of [2, sec. 8] to the p-adic world. The third chapter translates the results of [2, sec. 14] to the p-adic world. At the end of chapter three we present our concept of multivariate Hensel lifting. First we will give a multivariate Hensel lifting for homogeneous polynomial system, which will then be used to state an affine version. As mentioned before and contrary to Kuhlmann in both cases we won't need to solve $f(x) = 0 \mod p$ anymore.

2 Preliminaries

Let p be an arbitrary but fixed prime.

2.1 The p-adic numbers

We shortly want to recall important properties of the p-adic numbers. For a complete introduction to the topic see [9] or [10].

2.1.1 Definition of $\mathbb{Z}_p$ and $\mathbb{Q}_p$

Definition 2.1.1 (Non-archimedean norm)**.** Let K be a field. A non-archimedean norm $|\ |$ on K is a map

$$|\ | : K \to \mathbb{R}_{\geq 0}$$

that satisfies the following properties

1. $\forall x \in K : |x| \geq 0$ and $|x| = 0 \Leftrightarrow x = 0$ (positive definite).
2. $\forall x, y \in K : |xy| = |x|\,|y|$ (multiplicativity).
3. $\forall x, y \in K : |x + y| \leq \max\{|x|, |y|\}$ (satisfies the strong triangle inequality).

If K is endowed with a non-archimedean norm $|\ |$, we call K a non-archimedean valued field.

Remark 2.1.2. The term 'valued field' may cause some irritation as a valuation as defined below is not exactly a norm. Nevertheless we will follow the nomenclature in the literature and will use this definition.

Definition 2.1.3 (The p-adic valuation)**.** The p-adic valuation v on $\mathbb{Q}$ is defined as follows: If $\frac{a}{b} \in \mathbb{Q}^*$, write it as $\frac{a}{b} = p^r \frac{a'}{b'}$, where a', b' are prime to p. We set $v\left(\frac{a}{b}\right) = r$. Also we set $v(0) = \infty$.

Definition 2.1.4 (p-adic norm)**.** We define a norm $|\ |$ on $\mathbb{Q}$ by

$$|x| = \begin{cases} p^{-v(x)} & , x \in \mathbb{Q}^* \\ 0 & , x = 0 \end{cases}$$

One checks easily that $|\ |$ is a non-archimedean norm on $\mathbb{Q}$. Now we can construct the p-adic numbers.

Definition 2.1.5 (The p-adic numbers)**.** We define $\mathbb{Q}_p$ to be the completion of $\mathbb{Q}$ with respect to $|\ |$: Let $\mathcal{F}$ be the cauchy sequences in $\mathbb{Q}$ with respect to $|\ |$. For two cauchy sequences (a_n) and (b_n) we define $(a_n) + (b_n) := (a_n + b_n)$ and $(a_n) \cdot (b_n) := (a_n \cdot b_n)$. This way $(\mathcal{F}, +, \cdot)$ becomes a ring. Let $\mathcal{Z} \subset \mathcal{F}$ be the ideal of sequences with limit zero. We then set

$$\mathbb{Q}_p := \mathcal{F}/\mathcal{Z}.$$

Moreover, for $(a_n) \in \mathcal{F}$ we denote the corresponding class in $\mathbb{Q}_p$ with $\{a_n\}$.

Remark 2.1.6*.*

- The map, that maps $x \in \mathbb{Q}$ to the class containing the constant sequence $(x, x, x, \ldots)$ is an injective homomorphism from $\mathbb{Q}$ into $\mathbb{Q}_p$.

- The zeros sequences $\mathcal{Z}$ build up a maximal ideal in $\mathcal{F}$. Hence $\mathbb{Q}_p$ is a field.

Theorem 2.1.7. *The field $\mathbb{Q}_p$ has sum $+$ and product $\cdot$ given by*

$$\{a_n\} + \{b_n\} = \{a_n + b_n\} \text{ and } \{a_n\} \cdot \{b_n\} = \{a_n \cdot b_n\}.$$

Moreover, there is a unique non-archimedean norm $|\ |$ on $\mathbb{Q}_p$ satisfying $|\{a\}| = |a|$ for a constant cauchy sequence $(a_n) = (a)$ with $a \in \mathbb{Q}$. This norm is defined by

$$|\{a_n\}| := \lim_{n\to\infty} |(a_n)|\,.$$

Proof. See [1, p. 20]. □

Definition 2.1.8 (The p-adic integers)**.** The set of p-adic integers $\mathbb{Z}_p$ is defined as

$$\mathbb{Z}_p = \{x \in \mathbb{Q}_p : |x| \leq 1\} = \left\{ \sum_{i=0}^{\infty} a_i p^i : a_i \in \{0, 1, \ldots, p-1\}\ \forall i \right\}.$$

Due to the strong triangle inequality $\mathbb{Z}_p$ is a ring. One checks that $\mathbb{Q}_p$ is the field of fractions for $\mathbb{Z}_p$.

Theorem 2.1.9 (p-adic digit expansion)**.** *Every $x \in \mathbb{Q}_p$ contains a unique sequence $(a_n)_{n\in\mathbb{N}} \in x$ with*

$$a_n = \sum_{i=k}^{n} \lambda_i p^i,$$

where $k \in \mathbb{Z}$ and $\lambda_i \in \{0, 1, \ldots, p-1\}$. Furthermore, $x \in \mathbb{Z}_p$, if and only if $\lambda_i = 0$ for all $i < 0$ and $|x| = p^{-k}$.

Proof. See [1, p. 26]. □

The preceding theorem for any $x \in \mathbb{Q}_p$ allows us to use the *p-adic digit expansion*: Instead of writing x as a sequence we write

$$x = \sum_{i=v(x)}^{\infty} \lambda_i p^i.$$

For an algorithm to add or multiply p-adic numbers in this representation see [9, p. 2-4]. We will now investigate the ideals in $\mathbb{Z}_p$.

Proposition 2.1.10. *(p) is the maximal ideal in $\mathbb{Z}_p$.*

Proof. First check that $\left|\frac{1}{p}\right| = p > 1$. So the inverse of p is not contained in $\mathbb{Z}_p$. Hence p is not invertible in $\mathbb{Z}_p$. Reduction mod (p) yields, using the rp-adic digit expansion

$$\mathbb{Z}_p/(p) \cong \mathbb{F}_p.$$

Here $\mathbb{F}_p$ is the field with p elements. It follows that (p) is maximal. [9, p. 6] □

Proposition 2.1.10 also shows that the units in $\mathbb{Z}_p$ are exactly

$$\mathbb{Z}_p^* = \left\{ \sum_{i=0}^{\infty} a_i p^i : a_0 \neq 0 \right\}.$$

Therefore every element $x \in \mathbb{Z}_p$ can be written as

$$x = p^k u : k \in \mathbb{N}, u \in \mathbb{Z}_p.$$

It follows that the ideals in $\mathbb{Z}_p$ are of the form $(p^k), k \in \mathbb{N}$. There are other, though equivalent definitions of $\mathbb{Z}_p$ (and $\mathbb{Q}_p$ as it fields of fractions). Among those a very important one is to define $\mathbb{Z}_p$ as a projective limit, see [9, sec. 4.1]. This definition of $\mathbb{Z}_p$ yields the following:

Proposition 2.1.11. *For all $k \in \mathbb{N}$:*

$$\mathbb{Z}_p/(p^k) \cong \mathbb{Z}/(p^k).$$

Proof. See [9, p. 33]. □

2.1.2 Extensions of $\mathbb{Q}_p$

A full introduction to algebraic extension of $\mathbb{Q}_p$ can be found in [9, sec. 2]. We give the most important facts without proofs. Let K be a finite algebraic extension of $\mathbb{Q}_p$. It turns out that the algebraic structure of K is similar to that of $\mathbb{Q}_p$. Here only the important facts (without proofs) shall be revisited: We suppose $[K : \mathbb{Q}_p] = d$. The norm on $\mathbb{Q}_p$ extends uniquely to K via

$$x \mapsto |N(x)|^{\frac{1}{d}},$$

where $N = N_{K/\mathbb{Q}_p} : K^* \to \mathbb{Q}_p$ is the relative norm on K (as defined in field theory). We will use the following notation

Notation 2.1.12.

- $\mathcal{O}_K := \{x \in K : |x| \leq 1\}$.
- $\mathsf{P} = \{x \in K : |x| < 1\}$ is the maximal ideal in $\mathcal{O}_K$.
- For a subset $U \subset K$: $|U| = \{|x| : x \in U\}$.

Proposition 2.1.13. *There is an element $\pi \in \mathcal{O}_K$ with $|\pi| = \max |\mathcal{O}_k| \cap (0, 1)$. We call π an uniformizing element or uniformizer. π has the property that for any $x \in K^*$ there exists $z \in \mathbb{Z}$ such that $|x| = |\pi|^z$ (this fact is abbreviated by $|K^*| = |\pi|^{\mathbb{Z}}$). Furthermore, $\mathsf{P} = \pi\mathcal{O}_K$ and any ideal in $\mathcal{O}_K$ is of the form $(\pi^k), k \in \mathbb{N}$.*

Notation 2.1.14. $\mathbb{Q}_p^{alg}$ is the algebraic closure of $\mathbb{Q}_p$.

As for finite algebraic extensions the norm on $\mathbb{Q}_p$ extends uniquely to $\mathbb{Q}_p^{alg}$. However, $\mathbb{Q}_p^{alg}$ is not complete. So one defines the completion of $\mathbb{Q}_p^{alg}$.

Notation 2.1.15. $\mathbb{C}_p$ is the topological completion of $\mathbb{Q}_p^{alg}$.

$\mathbb{C}_p$ keeps the property of being algebraically closed:

Proposition 2.1.16. *$\mathbb{C}_p$ is algebraically closed.*

2.2 Vector spaces and operator norms

In this section let K be a non-archimedean valued field.

Definition 2.2.1. Let V be a vector space over K. A norm $\| \ \|$ on V is a map

$$\| \ \| : V \to \mathbb{R}_{\geq 0}$$

satisfying

1. $\forall x \in V : \|x\| \geq 0$ and $\|x\| = 0 \Leftrightarrow x = 0$,
2. $\forall \lambda \in K, x \in V : \|\lambda x\| = |\lambda| \, \|x\|$,
3. $\forall x, y \in V : |x + y| \leq \max |x| \, , |y|$.

If V is complete with respect to $\| \ \|$, we call V a K-Banach space.

Remark 2.2.2. $| \ |$ denotes a norm on a non-archimedean valued field and $\| \ \|$ a vector- or operator-norm.

Theorem 2.2.3. *Let V be a finite-dimensional normed vector space over K. Then V is complete.*

Proof. See [9, p. 96]. □

Notation 2.2.4. Let $V = K^n$. $\mathcal{E} = \{e_i : 1 \leq i \leq n\}$ is the standard basis on V (that is $(e_i)_j = \delta_{ij}$ and δ_{ij} is the Kronecker symbol).

Theorem 2.2.5. *Let V be a finite dimensional vector space over K. Then all norms on V are equivalent.*

Proof. See [9, p. 92]. □

Example 2.2.6. Let V be a K-vector space and $\mathcal{B} = \{b_1, b_2, \ldots\}$ be a basis of V. We can define the maximum-norm on V as followes: If $x = \sum\limits_i x_i b_i \in V$, then we set

$$\|x\| := \max_i |x_i|$$

One checks easily that $\| \ \|$ is a norm that satisfies the requirements of definition 2.2.1.

Definition 2.2.7. Let $V = K^n, n \in \mathbb{N}$, $\mathcal{E}$ the standard basis and $x = \sum\limits_{i=1}^{n} x_i e_i \in V$. We call $\|x\| = \max\limits_i |x_i|$ the standard norm on V.

Notation 2.2.8. $\mathbb{S}(V) = \{x \in V : \|x\| = 1\}$ is the sphere in V.

Definition 2.2.9. $\mathcal{L}_k(V_1, \ldots, V_k; W)$ is defined to be the K-vector space of K-multilinear functions from $V_1 \times \ldots \times V_k$ to W. If $V_1 = \ldots = V_k$ we abbreviate $\mathcal{L}_k(V_1, \ldots, V_k; W)$ by $\mathcal{L}(V, W)$. The expression $\psi(v, \ldots, v)$ will simply be denoted by $\psi(v^k)$.

Lemma 2.2.10. *Let $V_i = K^{n_i}$, $n_i \in \mathbb{N}$, $1 \leq i \leq k$ all endowed with the standard norm. Let W be a normed vector space over K and $\psi \in \mathcal{L}_k(V_1, \ldots, V_k; W)$. Then $\max\limits_{\substack{v_i \in \mathbb{S}(V_i) \\ 1 \leq i \leq k}} \|\psi(v_1, \ldots, v_k)\|$ exists and is finite.*

Proof. First let $\psi : V \to W$ be a linear map. Suppose V has dimension n. Let $v = \sum\limits_i \lambda_i e_i \in \mathbb{S}(V)$. Then $\max\limits_i |\lambda_i| = 1$ and

$$\|\psi(v)\| = \left\| \sum_i \lambda_i \psi(e_i) \right\| \leq \max_i |\lambda_i| \, \|\psi(e_i)\| \leq \|\psi(e_i)\| .$$

Since $e_i \in \mathbb{S}(V)$ for all i, we have that $\|\psi\| = \max\limits_i \|\psi(e_i)\|$. This shows that the maximum exists. Now return to the general case: If $v_1, \ldots, v_{k-1}$ are fixed, $\psi(v_1, \ldots, v_{k-1}, \cdot)$ is a linear map from V_k to W. So $\max\limits_{v_k \in \mathbb{S}(V_k)} \|\psi(v_1, \ldots, v_k)\|$ exists. Repeating this with all v_i one sees that the maximum exists and that is finite. □

Definition 2.2.11. Let $V_i = K^{n_i}$, $n_i \in \mathbb{N}$, $1 \leq i \leq k$ all endowed with the standard norm. Let W be a normed vector space over K. We define the induced norm on $\mathcal{L}_k(V_1, \ldots, V_k; W)$ as

$$\|\psi\| := \max_{\substack{v_i \in \mathbb{S}(V_i) \\ 1 \leq i \leq k}} \|\psi(v_1, \ldots, v_k)\| .$$

Lemma 2.2.12. $\| \ \| : \mathcal{L}_k(V_1, \dots, V_k; W) \to \|V\|$ *is a non-archimedean norm.*

Proof.

1. $\| \ \| \geq 0$ is inherited from the norm on W. Let $\|\psi\| = 0$. Then for any $v_i \in \mathbb{S}(V_i)$, $1 \leq i \leq k$, $\psi(v_1, \dots, v_k) = 0$. Hence $\psi = 0$.

2. Let $\lambda \in K$. Then

$$\|\lambda\psi\| = \max_{v_i \in \mathbb{S}(V_i)} \|\lambda\psi(v_1, \dots, v_k)\| = |\lambda| \max_{v_i \in \mathbb{S}(V_i)} \|\psi(v_1, \dots, v_k)\| = |\lambda| \, \|\psi\| .$$

3. Let $\psi, \phi \in \mathcal{L}_k(V_1, \dots, V_k; W)$. Then

$$\begin{aligned} \|\psi + \phi\| &= \max_{v_i \in \mathbb{S}(V_i)} \|(\psi + \phi)(v_1, \dots, v_k)\| \\ &\leq \max_{v_i \in \mathbb{S}(V_i)} \max \left(\|\psi(v_1, \dots, v_k)\| , \|\phi(v_1, \dots, v_k)\|\right) \\ &\leq \max \|\psi\| , \|\phi\| . \end{aligned}$$

This shows that $\| \ \|$ is a non-archimedean norm. □

For the rest of the section we let $V = K^n$, $V_i = K^{n_i}, 1 \leq i \leq k$, all endowed with the standard norm and W be a normed vector space over K.

Lemma 2.2.13. *Let* $\psi \in \mathcal{L}_k(V, W)$, $v \in V$. *Then*

$$\left\|\psi(v^k)\right\| \leq \|\psi\| \, \|v\|^k .$$

Proof. If $v = 0$, the assertion is trivial. Assume $v \neq 0$. The multilinearity of ψ yields for any $\lambda \in K$: $\psi((\lambda v)^k) = \lambda^k \psi(v^k)$. Let $\lambda \in K^*$ such that $|\lambda| = \|v\|^{-1}$. Then $\|\lambda v\| = 1$ and $\left\|\psi((\lambda v)^k)\right\| \leq \|\psi\|$ by the definition of $\|\psi\|$. Combine this with $\left\|\psi((\lambda v)^k)\right\| = |\lambda|^k \left\|\psi(v^k)\right\|$ to complete the argument. □

Lemma 2.2.14. *If* $\phi \in \mathcal{L}_k(V_1, \dots, V_k; W)$ *and* $\psi \in \mathcal{L}_1(W; U)$, *then*

$$\|\psi \circ \phi\| \leq \|\psi\| \, \|\phi\| .$$

Proof. Let $v_i \in \mathbb{S}(V_i)$, $1 \leq i \leq k$, be such that

$$\|\psi \circ \phi\| = \|\psi(\phi(v_1, \dots, v_k))\| .$$

Using lemma 2.2.13 we obtain

$$\|\psi(\phi(v_1,\ldots,v_k))\| \leq \|\psi\| \, \|\phi(v_1,\ldots,v_k)\| \, .$$

By the definition of $\|\phi\|$ we see that

$$\|\phi(v_1,\ldots,v_k)\| \leq \|\phi\| \, .$$

Now combine the three equations. □

Proposition 2.2.15. *Let everything as above. Consider the map*

$$\Phi : \mathcal{L}_{k-1}\left(V_1,\ldots,V_{k-1};\mathcal{L}_1(V,W)\right) \to \mathcal{L}_k(V_1,\ldots,V_k;W),$$

which is defined as follows: For $\psi \in \mathcal{L}_{k-1}\left(V_1,\ldots,V_{k-1};\mathcal{L}_1(V,W)\right)$ we define $\Phi(\psi)$ as

$$\Phi(\psi)(v_1,\ldots,v_k) := \psi(v_1,\ldots,v_{k-1})(v_k).$$

Then Φ is an isometry.

Proof. We have to show that 1. Φ is an isomorphism and that 2. $\|\psi\| = \|\Phi(\psi)\|$.

For 1. it is clear that Φ is a homomorphism. Assume $\Phi(\psi) = 0$. Then for all $v_i \in V_i$, $1 \leq i \leq k-1$, we have $\psi(v_1,\ldots,v_{k-1}) = 0$. Hence $\psi = 0$. This shows injectivity. To show surjectivity let $\phi \in \mathcal{L}_k(V_1,\ldots,V_k,W)$. Define $\varphi \in \mathcal{L}_{k-1}\left(V_1,\ldots,V_{k-1};\mathcal{L}_1(V,W)\right)$ to be the map

$$(v_1,\ldots,v_{k-1}) \mapsto (v_k \mapsto \phi(v_1,\ldots,v_{k-1},v_k)).$$

This is multilinear and its image clearly is in $\mathcal{L}_1(V_1,\ldots,V_k;W)$. Hence $\Phi(\varphi) = \phi$.

For 2. let $\psi \in \mathcal{L}_{k-1}\left(V_1,\ldots,V_{k-1};\mathcal{L}_1(V,W)\right)$. Recall that

$$\|\psi\| = \max_{\substack{v_i \in \mathbb{S}(V_i) \\ 1 \leq i \leq k-1}} \|\psi(v_1,\ldots,v_{k-1})\|$$

and

$$\|\psi(v_1,\ldots,v_{k-1})\| = \max_{v_k \in \mathbb{S}(V_k)} \|\psi(v_1,\ldots,v_{k-1})(v_k)\| \, .$$

So $\|\psi\| = \max\limits_{v_i \in \mathbb{S}(V_i)} \|\psi(v_1,\ldots,v_{k-1})(v_k)\| = \|\Phi(\psi)\|$. □

2.3 Ultrametric spaces

For a full introduction to ultrametric spaces see [10]. We want to recall some facts that will be used in this work.

2.3.1 Properties of ultrametric spaces

Definition 2.3.1. (See [10, p. 46]). A metric space (X, d) is an ultrametric space, if the metric d satisfies the strong triangle inequality. That is for all $x, y, z \in X$:

$$d(x, y) \leq \max\{d(x, z), d(y, z)\}.$$

Remark 2.3.2. Given a subset U of a non-archimedean valued field $(K, |\ |)$ (or of a normed vector space over K) we can make it an ultrametric space by endowing it with the metric

$$d(x, y) := |x - y|$$

In fact every ultrametric space can isometrically embedded into a suitable non-archimedean valued field K. For a proof see [10, p. 293].

Throughout this section (X, d) is an ultrametric space.

Proposition 2.3.3. *Every triangle is isosceles: Let $x, y, z \in X$. Then*

$$d(x, y) \neq d(y, z) \quad \Rightarrow \quad d(x, z) = \max\{d(x, y), d(y, z)\}.$$

Proof. (See [10, p. 47]). Without loss of generality assume $d(x, y) > d(y, z)$. The strong triangle inequality then gives $d(x, z) \leq \max\{d(x, y), d(y, z)\} = d(x, y)$. On the other hand we have $d(y, z) < d(x, y) \leq \max\{d(x, z), d(y, z)\}$. Consequently $\max\{d(x, z), d(y, z)\} = d(x, z)$ and it follows that $d(x, y) = d(x, z)$. □

Corollary 2.3.4. *Let $(V, \|\ \|)$ be a K-Banach space. Proposition 2.3.3 implies for all $x, y \in V$:*

1. *If $\|x\| \neq \|y\|$, then $\|x - y\| = \max\{\|x\|, \|y\|\}$.*

2. *If $\|x - y\| < \|x\|$, then $\|x\| = \|y\|$.*

Definition 2.3.5. (See [10, p. 47]). Let $a \in X$, $0 < r < \infty$. Then open ball of radius r with center a is the set

$$B_r(a) := \{x \in X : d(a, x) < r\}.$$

The closed ball of radius r with center a is the set $\overline{B}_r(a) := \{x \in X : d(a, x) \leq r\}$. The diameter of a ball is $d(B) := \sup\{d(x, y) : x, y \in B\}$. Given two sets $A, B \subset X$ the distance between them is defined as $d(A, B) := \inf\{d(a, b) : a \in A, b \in B\}$. If $x \in X$, we further define $d(x, A) := d(\{x\}, A)$.

Proposition 2.3.6. *Each ball in X is both open and closed. Each point of it is a center. A ball may have infinitely many radii.*

Proof. (See [10, p. 47]). Let $a \in X$ and $0 < r < \infty$. We define an equivalence relation on X by $x \sim y :\Leftrightarrow d(x, y) < r$. Clearly $B_r(a)$ is the equivalence class containing a. But it is also the complement of the union of all other classes, hence closed. Now consider $\overline{B}_r(a)$. We show that for every $b \in \overline{B}_r(a)$ we have $B_r(b) \subset \overline{B}_r(a)$: If $x \in B_r(b)$, then $d(x, a) \leq \max d(a, b), d(x, b) \leq r$, so $x \in \overline{B}_r(a)$. This shows that $\overline{B}_r(a)$ is open and that every point of the ball $B_r(a)$ is a center for it. An example for a ball with infinitely many radii is the unit ball in $\mathbb{Q}_p$, which has all radii $1 < r < p$. □

Remark 2.3.7. In some literature one speaks of 'clopen' balls.

2.3.2 Spherical completeness

Definition 2.3.8. (See [10, p. 52]). The metric in X is called discrete, if for any two sequences $(x_1, x_2, \ldots), (y_1, y_2, \ldots) \in X$ such that if $d(x_1, y_1) > d(x_2, y_2) > \ldots$, we have $\lim_{n\to\infty} d(x_n, y_n) = 0$.

Definition 2.3.9. (See [10, p. 52]). An ultrametric space is spherically complete, if each nested sequence of non-empty balls $B_1 \supset B_2 \supset \ldots$ has a nonempty intersection.

Proposition 2.3.10. *If the metric on a complete ultrametric space X is discrete, then X is spherically complete.*

Proof. (See [10, p. 52]). Let $B_1 \supset B_2 \supset \ldots$ be a nested sequence of balls. We may suppose that $B_n \neq B_{n+1}$ for all n. The discreteness of the metric gives $\lim_{n\to\infty} d(B_n) = 0$. The completeness implies that the intersection of the B_n is non-empty. □

Corollary 2.3.11. *Discretely valued fields and finite dimensional normed spaces over discretely valued fields are spherically complete.*

Proof. See [10, p. 53]. □

Example 2.3.12.

- Let K be finite dimensional algebraic extension of $\mathbb{Q}_p$. Then K is spherically complete.

- Let V be a finite dimensional vector space over K. Then V is spherically complete. In particular, for any $x \in K \backslash \{0\}$ the ray Kx is spherically complete.

It turns out that $\mathbb{C}_p$ is not spherically complete (see [9, sec. 3.3.4]). It is possible to define an extension of $\mathbb{C}_p$ that is both spherical and algebraically complete (see [9, sec. 3.2.5]). In the literature this completion is denoted by Ω_p.

2.4 Hensel lifting

We want to recall the classical Hensel lifting. On the one hand this gives good insights on techniques used in the p-adic world. On the other hand, and far more importantly, the idea of the proof is used in the proof of theorem 4.4.1.

Theorem 2.4.1 (Hensel's lemma)**.** *Let $f \in \mathbb{Z}_p[X]$ and $x \in \mathbb{Z}_p$ such that $|f(x)| = p^{-n}, n \geq 0$, $|f'(x)| = p^{-k} \neq 0$ and $k < \frac{n}{2}$. Then $\hat{x} = x - \frac{f(x)}{f'(x)}$ satisfies*

1. $|f(\hat{x})| \leq p^{-n-1}$

2. $|\hat{x} - x| = p^{-n+k}$

3. $|f'(\hat{x})| = |f'(x)|$

Moreover, the sequence $(x_i)_{i \geq 0}$ with $x_0 = x$ and $x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$ is defined and converges to a zero $\zeta \in \mathbb{Z}_p$ of f.

Proof. (See [9, p. 47]). First observe that $|\hat{x}| \leq \max\{|x|, p^{-n+k}\} \leq 1$, so $\hat{x} \in \mathbb{Z}_p$. We also have

$$|\hat{x} - x| = |f(x)| \, |f'(x)|^{-1} = p^{-n+k}.$$

Consider the Taylor expansion of f at x: $f(y) = f(x) + f'(x)(y-x) + t(y)(y-x)^2$, where $t \in \mathbb{Z}_p[X]$. Inserting $\hat{x}$ for y gives

$$f(\hat{x}) = f(x) - f'(x)\frac{f(x)}{f'(x)} + t(\hat{x})(\hat{x}-x)^2 = t(\hat{x})(\hat{x}-x)^2.$$

We have $t(\hat{x}) \in \mathbb{Z}_p$. Recall that we have assumed $k < \frac{n}{2}$. Hence

$$|f(\hat{x})| = |t(\hat{x})| \left|(\hat{x}-x)^2\right| \leq |\hat{x}-x|^2 \leq p^{-2n+2k} < p^{-n}.$$

Using the Taylor expansion of f' around x we can conclude that $f'(\hat{x}) = f'(x) + s(\hat{x})(\hat{x}-x)$ for some $s \in \mathbb{Z}_p[X]$. We can estimate:

$$|s(\hat{x})(\hat{x}-x)| \leq |\hat{x}-x| = p^{-n+k} < p^{-k} = |f'(x)|.$$

Using proposition 2.3.3 we can conclude that $|f'(x)| = |f'(\hat{x})|$. Now $\hat{x}$ satisfies the same requirements x has before. We therefore may apply the same steps to $\hat{x}$. This defines a sequence $(x_i)_{i\geq 0}$ with $x_0 = x$ and $x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$. For every i we have

$$|f(x_i)| = p^{-n-i} \quad \text{and} \quad |x_i - x_{i+1}| = p^{n+i-k}.$$

This shows that the sequence $(x_i)_{i\geq 0}$ is a Cauchy sequence having a limit $\zeta \in \mathbb{Z}_p$ with $f(\zeta) = 0$. □

Remark 2.4.2. The Hensel lifting ensures us a linear rate of convergence.

The Hensel lifting reduces solving $f(x) = 0$ to finding a solution $f(x) = 0 \mod p$ that satisfies $f'(x) \neq 0 \mod p$. An algorithm for solving equations mod p can, for instance, be found at [**?**, p. 382]. Kuhlmann's Hensel lifting in in [7] has similar preconditions as theorem 2.4.1. However, we don't want to recall his proof. The main differences to Kuhlmann when compared to our method will be the following: First, we won't be restricted to integer polynomials only. Second, we take a polynomial system $f(x) = 0$ of which we know a solution ζ. Then we use ζ as starting value for the Newton algorithm with respect to polynomial systems that are 'close' to f.

2.5 Analytic functions

For a full introduction to multivariate calculus in the p-adics see [11]. Let K be a non-archimedean valued field and V, W be finite dimensional K-Banach spaces. The goal of this section is to introduce analytic functions on V and develop the Taylor expansion. As in standard analysis one can define differentiation. Unfortunately, in the p-adic world this definition is not as useful as in complex analysis. It is, for example, possible to define a function that is injective but has a derivative that is constant zero (see [9, p. 218]). Instead one defines the following:

Definition 2.5.1. (See [11, p. 20]). Let $U \subset V$ be open and $f : U \to W$. f is called strictly differentiable in $v_0 \in U$, if there exists a linear map

$$Df(v_0) : V \to W$$

such that for any $\epsilon > 0$ there is an open neighborhood $U_\epsilon \subset U$ of v_0 with

$$\|f(v_1) - f(v_2) - Df(v_0)(v_1 - v_2)\| \leq \epsilon \|v_1 - v_2\|, \text{ for any } v_1, v_2 \in U_\epsilon$$

We proceed with a useful proposition on sequences in V:

Proposition 2.5.2. *Let $a_1, a_2, \dots$ be a sequence in V.*

1. *If $\lim\limits_{n\to\infty} a_n = a \neq 0$, then there exists $N \in \mathbb{N}$ with $\|a\| = \|a_n\|$ for all $n \geq N$.*
2. *$\sum\limits_n a_n$ converges, if and only if $\lim\limits_{n\to\infty} \|a_n\| = 0$.*

Proof. (See [10, p. 61]). For 1: There exists $N \in \mathbb{N}$ such that $\|a_n - a\| < \|a\|$ for all $n \geq N$. From proposition 2.3.3 it follows that for all $n \geq N$: $\|a_n\| = \|a\|$. For 2: If $\sum\limits_n a_n$ converges, it is a well known fact that $\lim\limits_{n\to\infty} a_n = 0$. Suppose $\lim\limits_{n\to\infty} a_n = 0$. Then for $m \geq n$:

$$\left\| \sum_{j=n}^{m} a_j \right\| \leq \max\{\|a_j\| : n \leq j \leq m\}.$$

This shows that $\sum_{j=1}^{n} a_j$ is a Cauchy sequence. Hence converges. □

Remark 2.5.3. The strong triangle inequality also applies to convergent series: If $\sum\limits_n a_n$ is a convergent series we have for any k that $\left\| \sum\limits_{n=1}^{k} a_n \right\| \leq \max\limits_{1\leq n\leq k} \|a_n\|$.

Definition 2.5.4. Let $s \in \mathbb{N}$ and $a = (a_1, \ldots, a_s) \in \mathbb{N}_0^s$. We set $|a| := \sum a_i$. We further define an order on $\mathbb{N}_0^s$: Let $a = (a_1, \ldots, a_s), b = (b_1, \ldots, b_s) \in \mathbb{N}_0^s$. Then $a < b$, if and only if

1. $|a| \leq |b|$ and
2. $\exists\, m$ such that $a_m < b_m$ and $a_i = b_i$ for all $i < m$.

This way $\mathbb{N}_0^s$ becomes a totally ordered set. We say that a series (a_n) tends to infinity, if for every $b \in \mathbb{N}_0^s$ there is a $N \in \mathbb{N}$ such that $b < a_N$. We also write $a_n \to \infty$.

Definition 2.5.5. (See [11, p. 25]). By a power series $f(X)$ in s variables $X = (X_1, \ldots, X_s)$ over V we mean a formal power series

$$f(X) = \sum_{\alpha \in \mathbb{N}_0^s} \lambda_\alpha X^\alpha,$$

where $\alpha = (\alpha_1, \ldots, \alpha_s)$, $\lambda_\alpha \in V$ and $X^\alpha := \prod_{i=1}^{s} X_i^{\alpha_i}$. We define the terms in f to be ordered in the order of definition 2.5.4 (with respect to α).

The power series form a ring with addition and multiplication defined as usual.

Definition 2.5.6. We define a map

$$r: \text{ formal power series over } V \to \mathbb{R}_{\geq 0}.$$

If $f(X) = \sum \lambda_\alpha X^\alpha$ is a formal power series over V, then $r(f) := \frac{1}{\limsup\limits_{\alpha \to \infty}\{\sqrt[|\alpha|]{\|\lambda_\alpha\|}\}}$.

Lemma 2.5.7. *Let $f(X) = \sum \lambda_\alpha X^\alpha$ be a formal power series over V. Then $f(x) := \sum \lambda_\alpha x^\alpha$ converges for all $x \in B_{r(f)}(0)$.*

Proof. Suppose $\|x\| < r(f)$. There exists some t with $\|x\| < t < r(f)$. From $\frac{t}{r(f)} < 1$ we deduce that there exists an $b \in \mathbb{N}_0^s$ such that $\sup\limits_{\alpha \geq b} t \cdot \sqrt[|\alpha|]{\|\lambda_\alpha\|} < 1$. Consequently we have for all $\alpha \geq b$ that

$$\|\lambda_\alpha x^\alpha\| \leq \|\lambda_\alpha\| \, \|x\|^{|\alpha|} = \|\lambda_\alpha\| \, t^{|\alpha|} \left(\frac{\|x\|}{t}\right)^{|\alpha|} < \left(\frac{\|x\|}{t}\right)^{|\alpha|}.$$

The term on the far right tends to zero as $|\alpha|$ tends to infinity. Using proposition 2.5.2 we can conclude that $f(x)$ converges. □

Remark 2.5.8. For a univariate power series $f(X)$ ($s = 1$) it is not only true that $f(x)$ converges for all $x \in B_{r(f)}(0)$, but also that $f(x)$ diverges, if $\|x\| > r(f)$ (see [10, p. 60]). However, the second statement becomes false, if $s > 1$. For an example let $f(X_1, X_2) = \sum\limits_{i=0}^{\infty} X_1^i$. Then $r(f) = 1$. We have that $\|(0, p^{-1})\| = p > r(f)$, but $f(0, p^{-1})$ converges.

Lemma 2.5.9. *Let $f(X) = \sum \lambda_\alpha X^\alpha$ be a power series over V. Suppose $z \in B_{r(f)}(0)$. Then*

$$f(x - z) := \sum_{\alpha \in \mathbb{N}_0^s} \lambda_\alpha (x - z)^\alpha$$

converges for all $x \in B_{r(f)}(0)$. Further, there is a formal power series $g(X)$ such that $g(x) = f(x - z)$ for all $x \in B_{r(f)}(0)$.

Proof. The first part is a corollary of the strong triangle inequality: For all $x \in B_{r(f)}(0)$ we have $\|x - z\| \leq \max\{\|x\|, \|z\|\} \leq r(f)$. For the second part expand all terms $(X - z)^\alpha$, $\alpha \in \mathbb{N}_0^s$. □

Corollary 2.5.10 (Point of expansion)**.** *Let $f(X)$ be a power series over V. Suppose $z \in B_{r(f)}(0)$. Then there exists a formal power series $g(X)$ with $B_{r(g)}(0) = B_{r(f)}(0)$ and*

$$\forall x \in B_{r(f)}(0) : g(x - z) = f(x).$$

Proof. (See [11, p. 31]). Set $g(X) := f(X + z)$. By lemma 2.5.9 $g(X)$ converges for all $x \in B_{r(f)}(0)$. Hence $B_{r(f)}(0) \subset B_{r(g)}(0)$. This implies $z \in B_{r(g)}(0)$. By symmetry we also have $B_{r(g)}(0) \subset B_{r(f)}(0)$, hence $B_{r(g)}(0) = B_{r(f)}(0)$. □

Remark 2.5.11. If $r(g) = r(f)$, then clearly $B_{r(g)}(0) = B_{r(f)}(0)$. However, the opposite is not always true. From proposition 2.3.6 we know that a ball may have infinitely many radii.

Definition 2.5.12. Let $f(X) := \sum_\alpha \lambda_\alpha X^\alpha$ be a formal power series and let $\underline{i} := (\delta_{ij})_{j=1}^s \in \mathbb{N}_0^s = (0, \ldots, 0, 1, 0, \ldots, 0)$ (the i-th entry is 1). We define the i-th formal derivative as

$$\frac{\partial f}{\partial X_i} := \sum_\alpha \alpha_i \lambda_\alpha X^{\alpha - \underline{i}}$$

More general, for $\alpha = (\alpha_1, \ldots, \alpha_s) \in \mathbb{N}_0^s$ we define

$$\frac{\partial^\alpha f}{\partial X^\alpha} := \frac{\partial^{\alpha_1}}{\partial X_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_s}}{\partial X_s^{\alpha_s}}(f).$$

As in standard analysis one checks that taking derivatives is linear and that

$$\frac{\partial}{\partial X_i}\frac{\partial}{\partial X_j}(f) = \frac{\partial}{\partial X_j}\frac{\partial}{\partial X_i}(f)$$

for all $1 \leq i, j \leq s$. For the derivative we have the following:

Lemma 2.5.13. *Let $f(X)$ be a formal power series over V with $r(f)$ as defined above. Let $1 \leq i \leq s$. Then $r(f) \leq r(\frac{\partial f}{\partial X_i})$.*

Proof. (See [11, p.31]). Let $f(X) = \sum_\alpha \lambda_\alpha X^\alpha$ such that and $\frac{\partial f}{\partial X_i} := \sum_\alpha \alpha_i \lambda_\alpha X^{\alpha - \underline{i}}$. Recall that for all $n \in \mathbb{N}$: $|n| \leq 1$, hence for all $\alpha \in \mathbb{N}_0^s$: $\|\alpha_i \lambda_\alpha\| \leq \|\lambda_\alpha\|$. Consequently $\limsup_{\alpha \to \infty}\{\sqrt[|\alpha|]{\|\lambda_\alpha\|}\} \geq \limsup_{\alpha \to \infty}\{\sqrt[|\alpha|]{\|\alpha_i \lambda_\alpha\|}\}$. The claim follows then. □

Corollary 2.5.14. *Let $f(X)$ be a formal power series over V. Let $\alpha \in \mathbb{N}_0^s$. Then $r(f) \leq r(\frac{\partial^\alpha f}{\partial X^\alpha})$.*

Proposition 2.5.15 (Taylor expansion)**.** *Let $f(X)$ be a formal power series. Suppose $z \in B_{r(f)}(0)$. Then*

$$f(X - z) = \sum_{\alpha \in \mathbb{N}_0^s} \frac{1}{\alpha_1! \cdot \ldots \cdot \alpha_s!} \frac{\partial^\alpha f}{\partial X^\alpha}(z)(X - z)^\alpha.$$

Proof. (See [11, p. 33]). Let $f(X - z) = \sum\limits_{\alpha} \lambda_\alpha (X - z)^\alpha$. Using the definition of the formal derivative by induction on α one checks that $\left(\prod\limits_{i=1}^{s} \alpha_i!\right) \lambda_\alpha = \frac{\partial^\alpha f}{\partial X^\alpha}(z)$. □

Definition 2.5.16. Let $f(X)$ be a formal power series. Suppose $z \in B_{r(f)}(0)$ and $k \in \mathbb{N}_0$. We define the multilinear map $D^k f(z) \in \mathcal{L}_k(K^s, V)$ by its images on the standard basis $\mathcal{E}$:

$$D^k f(z)(e_{i_1}, \ldots, e_{i_k}) := \frac{\partial^\alpha f}{\partial X^\alpha}(z). \tag{2.5.17}$$

where $i_1, \ldots, i_k \in \{1, \ldots, s\}$ and $\alpha = (\alpha_1, \ldots, \alpha_k) = \sum\limits_{j=1}^{k} e_{i_j}$.

Example 2.5.18. Let $V = \mathbb{Q}_p^2$, $f(X, Y) = \begin{pmatrix} X^2 - XY + 1 \\ Y^2 - 1 \end{pmatrix}$ and $z = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

- For $k = 0$ we have that: $D^0 f(z) = f(z) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

- For $k = 1$ we have that:
    1. $D^1 f(z)(e_1) = \frac{\partial f}{\partial X}(z) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.
    2. $D^1 f(z)(e_2) = \frac{\partial f}{\partial Y}(z) = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$.

  Hence for $x = \lambda e_1 + \mu e_2 \in V$ we have $D^1 f(z)(x) = \lambda \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \mu \begin{pmatrix} -1 \\ 2 \end{pmatrix}$.

- For $k = 2$ we have that:
    1. $D^2 f(z)(e_1, e_1) = \frac{\partial^2 f}{\partial X^2}(z) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$.
    2. $D^2 f(z)(e_1, e_2) = D^2 f(z)(e_2, e_1) = \frac{\partial}{\partial X}\frac{\partial}{\partial Y}(f)(z) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$.
    3. $D^2 f(z)(e_2, e_2) = \frac{\partial^2 f}{\partial Y^2}(z) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$.

  Hence for $x_1 = \lambda_1 e_1 + \mu_1 e_2 \in V$ and $x_2 = \lambda_2 e_1 + \mu_2 e_2 \in V$ we have

  $$D^2 f(z)(x_1, x_2) = \lambda_1 \lambda_2 \begin{pmatrix} 2 \\ 0 \end{pmatrix} + (\lambda_1 \mu_2 + \lambda_2 \mu_1) \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \mu_1 \mu_2 \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

- For $k \geq 3$ we have that $D^k f(z) = 0$.

Recall from definition 2.2.9 that for a multilinear operator $A : V \times \ldots \times V \to V$ we have set $A(v^k) := A(v, \ldots, v)$. As in complex analysis we can rewrite the Taylor expansion of proposition 2.5.15.

Corollary 2.5.19. *Let $f(X)$ be a formal power series. Suppose $z \in B_{r(f)}(0)$. Then*

$$f(X - z) = \sum_{k=0}^{\infty} \frac{D^k f(z)}{k!}(X - z)^k. \tag{2.5.20}$$

Proof. Let $f_k(X-z) = \sum\limits_{|\alpha|=k} \frac{1}{\alpha_1! \cdot \ldots \cdot \alpha_s!} \frac{\partial^\alpha f}{\partial X^\alpha}(z)(X-z)^\alpha$. The goal is to compare (2.5.20) to (2.5.17) for every k. Clearly $f_0(X - z) = f(z) = D^0 f(z)$. Let $k \geq 1$, $\mathcal{E} = \{e_1, \ldots, e_r\}$ be the standard basis and $\{e_{i_1}, \ldots, e_{i_k}\} \subset \mathcal{E}$. Set $\alpha = \sum\limits_{j=1}^{k} e_{i_j}$. We know that for any permutation $\sigma : \{1, \dots, k\} \to \{1, \dots, k\}$ it is true that

$$D^k f(z)(e_{i_1}, \dots, e_{i_k}) = D^k f(z)(e_{\sigma(i_1)}, \dots, e_{\sigma(i_k)}) = \frac{\partial^\alpha f}{\partial X^\alpha}(z),$$

There are exactly $\binom{k}{\alpha} = \frac{k!}{\alpha_1! \cdot \ldots \alpha_s!}$ possibilites to order the e_{i_j} and obtain $\frac{\partial^\alpha f}{\partial X^\alpha}(z)$. Hence

$$D^k f(z)(X - z)^k = \sum_{|\alpha|=k} \binom{k}{\alpha} \frac{\partial^\alpha f}{\partial X^\alpha}(z)(X - z)^\alpha = k! \cdot f_k(X - z).$$

The claim follows then. □

Example 2.5.21. For example 2.5.18 this means

$$f(X-1, Y-1) = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + (X-1) \begin{pmatrix} 1 \\ 2 \end{pmatrix} + (Y-1) \begin{pmatrix} -1 \\ 2 \end{pmatrix} + \frac{1}{2}(X-1)^2 \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$
$$+ (X-1)(Y-1) \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \frac{1}{2}(Y-1)^2 \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

Proposition 2.5.22. *Let $f(X)$ be a non-zero power series. There is a point $z \in B_{r(f)}(0)$ such that $f(x) \neq 0$. In particular, if two power series $f(X)$, $g(X)$ coincide on $B_{r(f)}(0)$, then $f(X) = g(X)$.*

Proof. See [11, p. 39]. □

Definition 2.5.23. (See [11, p. 38]). Let $U \subset K^s$. $f : U \to V$ is called locally analytic, if for any point $x_0 \in U$ there is a ball $B_\varepsilon \subset U$ around x_0 and a power series $F(X) = \sum_{\alpha \in \mathbb{N}_0^s} \lambda_\alpha X^\alpha$ such that $B_\varepsilon \subset B_{r(F)}(x_0)$ and:

$$\forall x \in B_\varepsilon : f(x) = F(x - x_0).$$

Let $U \subset K^s$. $f : U \to V$ is called analytic, if there is $x_0 \in U$ and a power series $F(X) = \sum_{\alpha \in \mathbb{N}_0^s} \lambda_\alpha X^\alpha$ such that $U \subset B_{r(F)}(x_0)$:

$$\forall x \in U : f(x) = F(x - x_0).$$

Proposition 2.5.24. *Let $U \subset K^s$, $0 \in U$ and $f : U \to V$ be (locally) analytic. Then f is strictly differentiable in every point $x \in U$ and the function $z \mapsto Df(z)$ is analytic.*

Proof. See [11, p. 39]. □

Remark 2.5.25. The linear maps $Df(z)$ (as defined in definition 2.5.1) and $Df(z)$ (as defined in definition 2.5.16) are the same. Proposition 2.5.22 tells us that the Taylor expansion for analytic functions is unique.

2.6 Orthogonal sets

2.6.1 Best approximation

Throughout this section (X, d) denotes an ultrametric space.

Definition 2.6.1. (See [10, p. 55]). Let Y be a subset of X. Let $x \in X, y \in Y$. The point y is called a best approximation of x in Y, if

$$d(x,y) = d(x,Y) := \inf\{d(x,z) : z \in Y\}.$$

Proposition 2.6.2. *Let Y be a non-empty subset of X without isolated points. Let $x \in X \backslash Y$ and suppose that x has a best approximation in Y. Then it has infinitely many best approximations in Y.*

Proof. (See [10, p. 55]). Let y be a best approximation of x in Y. Then all points in $B_{d(x,y)}(y) \cap Y$ have the same distance to x, thus are best approximations of x in Y. □

Proposition 2.6.3. *Let Y be a nonempty subset of X. If Y is spherically complete as a metric space, then each point of X has a best approximation in Y.*

Proof. (See [10, p. 55]). Let $x \in X$. For each $n \in \mathbb{N}$ set $B_n := \{y \in Y : d(x,y) \leq d(x,Y) + \frac{1}{n}\}$. By construction none of the B_n is empty and $B_1 \supset B_2 \supset \dots$ is a nested sequence of balls. By assumptions Y is spherically complete, which means that every nested sequence of non-empty balls has a non-empty intersection (see definition 2.3.9). It follows that $\bigcap_n B_n$ is not empty and every element contained in it is a best approximation of x. □

2.6.2 Orthogonal bases in K-Banach spaces

Let V be a finite-dimensional K-Banach space over a non-archimedean valued field K. The norm on V does not come from an inner product. It is therefore impossible to define orthogonality in the common way. Nevertheless, in [10] one can find a concept of orthgonality in K-Banach spaces. It was originally developed in the context of the vector space of continuous functions from $\mathbb{Z}_p$ to K and the so called Mahler base. However, we will use this concept for the vector space K^n and give an explicit way to compute orthogonal bases in it.

Definition 2.6.4. (See [10, p. 145]). Suppose $x, y \in V$. We write $x \perp y$, if and only if 0 is a best approximation of x in Ky. That is $\|x\| = \inf\{\|x - \lambda y\| : \lambda \in K\}$.

Lemma 2.6.5. *Let $x, y \in V$. Suppose there is $c \in (0, 1]$ such that $\|x + y\| \geq c\,\|x\|$. Then also $\|x + y\| \geq c\,\|y\|$.*

Proof. (See [10, p. 32]). Using $c^{-1}\|x + y\| \geq \|x\|$ we can estimate that $\|y\| = \|x + y - x\| \leq \max\{\|x + y\|, \|x\|\} \leq c^{-1}\|x + y\|$. □

Lemma 2.6.6. *Let $x, y \in V$. If $x \perp y$, then also $y \perp x$.*

Proof. (See [10, p. 146]). Suppose $x \perp y$. Clearly $\|y - 0 \cdot x\| \geq \|y\|$. Let $\lambda \in K^*$. Then by assumption

$$\|y - \lambda x\| = |\lambda|\left\|x - \lambda^{-1} y\right\| \geq |\lambda|\,\|x\| = \|\lambda x\|\,.$$

Lemma 2.6.5 with $c = 1$ implies $\|y - \lambda x\| \geq \|y\|$. □

Definition 2.6.7. (See [10, p. 146]).

1. Let $x \in V$ and $D_1, D_2 \subset V$. We write $x \perp D_1$, if $x \perp d$ for all $d \in D_1$. We write $D_1 \perp D_2$, if $d_1 \perp d_2$ for all $d_1 \in D_1, d_2 \in D_2$.

2. $\{x_1, x_2, \ldots\} \subset V$ is called an orthogonal set, if for each $n \in \mathbb{N}$:

$$x_n \perp \operatorname{Span}\{x_1, \ldots, x_{n-1}, x_{n+1}, \ldots\}.$$

   $\{x_1, x_2, \ldots\}$ is called orthonormal, if $\|x_i\| = 1$ for all i. (Span{} denotes the K-linear span.)

Contrary to the complex case it may occur that in an n-dimensional vector space we can find more that n vectors that are pairwise orthogonal to each other. This is why in the definition of the orthogonal base the condition for each x_i of being orthogonal to the whole K-linear span of $\{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots\}$ is indispensable. The following example illustrates this.

Example 2.6.8. Let $V = K^2$ equipped with the standard norm and

$$x_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, x_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, x_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Then for all $\lambda \in K$: $\|x_1 - \lambda x_2\| = \max\{1, |\lambda|\} \geq 1 = \|x_1\|$ Hence $x_1 \perp x_2$. In the same way on verifies that $x_1 \perp x_3$ and $x_2 \perp x_3$.

Lemma 2.6.9. *Let $\{x_1, x_2, \ldots\}$ be an orthogonal set of non-zero elements of V. If $\|V\| = |K|$ (see notation 2.1.12) one can find $\lambda_1, \lambda_2, \ldots$ such that $\{\lambda_1 x_1, \lambda_2 x_2, \ldots\}$ is an orthonormal set.*

Proof. (See [10, p. 146]). Let $\{x_1, x_2, \ldots\}$ be an orthogonal set of non-zero elements of V. Choose $\lambda_1, \lambda_2, \ldots$ such that for all i: $|\lambda_i| = \|x_i\|^{-1}$. Then also $\{\lambda_1 x_1, \lambda_2 x_2, \ldots\}$ is an orthogonal set and $\|\lambda_i x_i\| = 1$. □

Proposition 2.6.10.

1. *$\{x_1, x_2, \ldots\}$ is orthogonal, if and only if $\{x_1, x_2, \ldots, x_n\}$ is orthogonal for all $n \in \mathbb{N}$.*

2. *$\{x_1, \ldots, x_n\}$ is orthogonal, if and only if for each $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$:*

$$\left\| \sum_{i=1}^{n} \lambda_i x_i \right\| = \max\{|\lambda_i| \, \|x_i\| : 1 \leq i \leq n\}. \tag{2.6.11}$$

3. *$\{x_1, \ldots, x_n\}$ is orthogonal, if and only if for each $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$ and for all $m \in \{2, \ldots, n\}$:*

$$\left\| \sum_{i=1}^{m} \lambda_i x_i \right\| \geq |\lambda_m| \, \|x_m\| . \tag{2.6.12}$$

Proof. (See [10, p. 146]).

1. is immediate.

2. Let $\{x_1, \ldots, x_n\}$ be an orthogonal set. Then for all $1 \leq j \leq n$ we have $\lambda_j x_j \perp \text{Span}\{x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n\}$. This yields $\left\|\lambda_j x_j + \sum\limits_{i \neq j} \lambda_j x_j\right\| \geq \|\lambda_j x_j\|$. Consequently $\left\|\sum\limits_i \lambda_i x_i\right\| \geq \max\limits_i \|\lambda_i x_i\|$. The opposite inequality follows from the strict triangle inequality. For the other direction suppose that the equality holds for any $\lambda_1, \ldots, \lambda_n \in K$. Let $1 \leq j \leq n$. For any $v = \sum\limits_{i \neq j} \lambda_i x_i$ we have $\|x_j - v\| \geq \|x_j\|$. This tells us $x_j \perp \text{Span}\{x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n\}$. Consequently $\{x_1, \ldots, x_n\}$ is an orthogonal set.

3. Observe that (2.6.11) implies (2.6.12). For (3) it suffices to show the opposite implication. By assumption we have $\left\|\sum\limits_{i=1}^{n} \lambda_i x_i\right\| \geq \|\lambda_n x_n\|$. Lemma 2.6.5 implies that also $\left\|\sum\limits_{i=1}^{n} \lambda_i x_i\right\| \geq \left\|\sum\limits_{i=1}^{n-1} \lambda_i x_i\right\|$. Downward induction shows that $\left\|\sum\limits_{i=1}^{n} \lambda_i x_i\right\| \geq \max\limits_i \|\lambda_i x_i\|$.

□

Lemma 2.6.13. *If $\{x_1, x_2, \ldots\}$ is an orthogonal set and does not contain 0, then $x_1, x_2, \ldots$ are linearly independent.*

Proof. Suppose that $\sum\limits_{j=1}^{n} \lambda_{i_j} x_{i_j} = 0$. Let $N = \max\{i_j : 1 \leq j \leq n\}$. From proposition 2.6.10 we know that $\{x_1, \ldots, x_N\}$ is an orthogonal set. Hence $0 = \left\|\sum\limits_{j=1}^{n} \lambda_{i_j} x_{i_j}\right\| = \max\{\left\|\lambda_{i_j} \left\|x_{i_j}\right\|\right\| : 1 \leq j \leq n\}$. It follows that $\left\|\lambda_{i_j}\right\| = 0, 1 \leq j \leq n$. □

Definition 2.6.14. (See [10, p. 147]). Let $x_1, x_2, \ldots, x_n$ be non-zero elements of V. $\{x_1, x_2, \ldots x_n\}$ is an orthogonal base, if

1. $\{x_1, x_2, \ldots, x_n\}$ is an orthogonal set of V.

2. for each $x \in V$, there are $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$ such that $x = \sum\limits_{i=1}^{n} \lambda_i e_i$.

We have made up a concept of orthogonality in (finite-dimensional) vector spaces over non-archimedean fields. Note that in [10] the theory is used for $\dim V = \infty$. However, for our purposes $\dim V < \infty$ is sufficient.

Definition 2.6.15. If $\{x_1, x_2, \dots, x_n\}$ is an orthogonal base, we call the set

$$T = \mathrm{Span}\{x_2, \dots, x_n\}$$

an orthogonal complement of Kx in V.

Remark 2.6.16.

- Orthogonal complements need not be unique as the example 2.6.8 shows.
- Orthogonal complements are dependent of the choice of a base.
- Let $x \in V$ and T be an orthogonal complement of Kx. Then $V = Kx \bigoplus T$.
- Suppose V has dimension $n < \infty$ over K. Then lemma 2.6.13 implies that any orthogonal set $\{x_1, \dots, x_n\}$ of non-zero elements is a basis for V.

Lemma 2.6.17. *Let $x, y, z \in V$ such that $x \perp y$, $x \perp z$ and $y - z \in Kx$. Then $\|y\| = \|z\|$.*

Proof. By replacing x with a multiple we may assume without loss of generality that $y - z = x$. The strict triangle inequality gives $\|x\| \leq \max\{\|y\|, \|z\|\}$. Without restriction we may assume that $\|x\| \leq \|z\|$. Since $x \perp z$, we have by Proposition 2.6.10 that $\|y\| = \max\{\|x\|, \|z\|\} = \|z\|$. □

Lemma 2.6.18. *Let $y \in V$, $y \neq 0$, and T_y be an orthogonal complement of Ky. If $x \in y + T_y$, then y is a best approximation of x in Ky.*

Proof. Let $x = y + y'$, where $y' \in T_y$. Then for any $\lambda \in K$ we have that

$$\|x - \lambda y\| = \|(1 - \lambda)y + y'\| = \max\{\|(1 - \lambda)y\|, \|y'\|\} \geq \|y'\|.$$

This shows that $\lambda = 1$ minimizes the term. □

Lemma 2.6.19. *Let $x, y \in V \backslash \{0\}$ and T_y be an orthogonal complement of Ky. Let $x = \lambda y + y'$, with $\lambda \in K$ and $y' \in T_y$. If $\|x - y\| < \|y\|$, then $|\lambda| = 1$.*

Proof. We have that $\|x - y\| = \|(\lambda - 1)y + y'\| = \max\{|\lambda - 1| \|y\|, \|y'\|\}$. In particular, $|\lambda - 1| \|y\| < \|y\|$. Hence $|\lambda - 1| < 1$, from which follows that $|\lambda| = 1$. □

In [10, p. 148] one can find something like an p-adic Gram-Schmidt-algorithm. However, we now give a simpler way to compute orthogonal bases.

Lemma 2.6.20. *Let $\mathcal{E} = \{e_1, \dots, e_n\}$ be the standard basis of V. Let $V = K^n$ and $x = \sum\limits_{i=1}^{n} x_i e_i \in V$. Suppose that $\|x\| = |x_k|$. Then $\{x\} \cup (\mathcal{E} \backslash \{e_k\})$ is an orthogonal base for V. Or, equivalently, $\mathcal{E} \backslash \{e_k\}$ is an orthogonal complement for Kx in V.*

Proof. Without restriction we may suppose that $\|x\| = |x_1|$. Let $m \in \{2, \dots, n\}$ and $\lambda_1, \dots, \lambda_m \in K$. According to proposition 2.6.10 we have to show the inequality $\|\lambda_1 x + \sum_{i=2}^{m} \lambda_i e_i\| \geq \|\lambda_m\|$. We can estimate

$$\left\| \lambda_1 x + \sum_{i=2}^{m} \lambda_i e_i \right\| = \max\{|\lambda_1 x_1|, |\lambda_1 x_2 + \lambda_2|, \dots |\lambda_1 x_m + \lambda_m|\}$$
$$\geq \max\{|\lambda_1 x_1|, |\lambda_1 x_m + \lambda_m|\}.$$

Observe that by assumption $|\lambda_1 x_1| \geq |\lambda_1 x_m|$. There are two cases: If $|\lambda_m| > |\lambda_1 x_1|$, then $|\lambda_1 x_m + \lambda_m| = |\lambda_m|$. It follows that

$$\max\{|\lambda_1 x_1|, |\lambda_1 x_m + \lambda_m|\} = |\lambda_m|.$$

If otherwise $|\lambda_m| \leq |\lambda_1 x_1|$, then $|\lambda_1 x_m + \lambda_m| \leq |\lambda_1 x_1|$. This yields

$$\max\{|\lambda_1 x_1|, |\lambda_1 x_m + \lambda_m|\} = |\lambda_1 x_1| \geq |\lambda_m|.$$

In both cases $\|\lambda_1 x + \sum_{i=2}^{m} \lambda_i e_i\| \geq |\lambda_m|$. □

2.7 The p-adic projective space

Let K be a non-archimedean valued field and V a finite dimensional K-Banach space.

Definition 2.7.1. We define an equivalence relation on $V \backslash \{0\}$ by:

$$x \sim y \quad \Leftrightarrow \quad \exists \lambda \in K^* : x = \lambda y.$$

The set $\mathbb{P}(V) = V/\sim$ is called the projective space over V.

In [11, p. 51] a metric on p-adic manifolds is defined using charts. However, it would be nice to have a metric on $\mathbb{P}(V)$ that is independent of charts. If K is a finite algebraic extension of $\mathbb{Q}_p$ and V is a finite dimensional vector space over K

with $\|V\| = |K|$ (see notation 2.1.12), we can define such a metric on $\mathbb{P}(V)$. As before, $\mathcal{O}_K$ denotes the ring of integers in K, π is an uniformizer and P the maximal ideal in $\mathcal{O}_K$.

Definition 2.7.2. Let K be a finite algebraic extension of $\mathbb{Q}_p$ and V be a finite dimensional vector space over K with $\|V\| = |K|$. Let a, b in $\mathbb{P}(V)$. Then for $x \in a$ and $y \in b$ we define

$$d_{\mathbb{P}}(a,b) := \min_{\lambda \in K} \frac{\|x - \lambda y\|}{\|x\|}. \tag{2.7.3}$$

Proposition 2.7.4. *$d_{\mathbb{P}}$ satisfies the properties of an ultrametric on $\mathbb{P}(V)$.*

Proof. We have to show the following:

1. d is well-defined.

2. $d \geq 0$ and $d(a,b) = 0 \Leftrightarrow a = b$.

3. $d(a,b) = d(b,a)$ for all $a, b \in \mathbb{P}(K^n)$.

4. $d(a,b) \leq \max\{d(a,c), d(c,b)\}$ for all $a, b, c \in \mathbb{P}(K^n)$.

For this purpose let $x \in a$, $y \in b$, $z \in c$ be representatives.

1. Since K is a finite algebraic extension of $\mathbb{Q}_p$, corollary 2.3.11 says that the ray Ky is spherically complete. Proposition 2.6.3 implies that x has a best approximation in Ky, so the minimum in (2.7.3) exists. For the independence of representatives let $\mu_1, \mu_2 \in K^*$:

$$\min_{\lambda \in K} \frac{\|\mu_1 x - \lambda \mu_2 y\|}{\|\mu_1 x\|} = \min_{\lambda \in K} \frac{\left\|x - \lambda \frac{\mu_2}{\mu_1} y\right\|}{\|x\|} = \min_{\lambda \in K} \frac{\|x - \lambda y\|}{\|x\|}.$$

2. $d \geq 0$ is inherited from $\| \ \|$. Further, $d(a,b) = 0$, if and only if $\|x - \lambda y\| = 0$ for some $\lambda \in K$. But then $x = \lambda y$ and $a = b$.

3. By (1) and the assumption $\|V\| = |K|$ we may assume without restriction that $\|x\| = \|y\| = 1$. If $|\lambda| > 1$, then $\|x - \lambda y\| = \|\lambda y\| = |\lambda| > 1 \geq \|x - y\|$. If $|\lambda| < 1$, then $\|x - \lambda y\| = \|x\| = 1 \geq \|x - y\|$. This shows that $\min\limits_{\lambda \in K} \|x - \lambda y\| = \min\limits_{|\lambda|=1} \|x - \lambda y\|$. But then $\min\limits_{|\lambda|=1} \|x - \lambda y\| = \min\limits_{|\lambda|=1} \|\lambda^{-1} x - y\|$, which tells us that $d(a,b) = d(b,a)$.

4. Again we may assume $\|x\| = \|y\| = \|z\| = 1$. Let $\mu, \rho \in K$ with

$$\min_{|\lambda|=1} \|y - \lambda z\| = \|y - \mu z\| = |\pi|^m \text{ and } \min_{|\lambda|=1} \|z - \lambda x\| = \|z - \rho x\| = |\pi|^r .$$

Observe that $m, r \geq 0$. Recall that with P we have denoted the maximal ideal in $\mathcal{O}_K$. For all i we have $y_i - \mu z_i \equiv 0 \mod \mathsf{P}^m$ and $z_i - \rho x_i \equiv 0 \mod \mathsf{P}^r$. This yields $y_i - \mu\rho x_i \equiv 0 \mod \mathsf{P}^{\min(r,m)}$. It follows that $\min_\lambda \|y - \lambda x\| \leq |\pi|^{\min(r,m)}$ and consequently $d(a,b) \leq \max\{d(a,c), d(c,b)\}$.

□

Remark 2.7.5.

- For all $a, b \in \mathbb{P}(V)$: $d_{\mathbb{P}}(a,b) \leq 1$.
- For the first three conditions it suffices for the field K to be spherically complete. The requirement of being a finite algebraic extension of $\mathbb{Q}_p$ is only needed for the strong triangle inequality.

The following lemma relates $d_{\mathbb{P}}$ to the concept of orthogonality of section 2.6.

Lemma 2.7.6. *Let K be a finite algebraic extension of $\mathbb{Q}_p$ and V a finite dimensional vector space over K. For any $x, y \in V \backslash \{0\}$ we have*

$$x \perp y \quad \Leftrightarrow \quad d_{\mathbb{P}}(Kx, Ky) = 1.$$

Proof. Suppose $x \perp y$. Then for all $\lambda \in K$: $\|x - \lambda y\| \geq \|x\|$ and consequently $\min_{\lambda \in K} \frac{\|x - \lambda y\|}{\|x\|} = \frac{\|x\|}{\|x\|} = 1$. If on the other hand $d_{\mathbb{P}}(Kx, Ky) = 1$, we have for all $\lambda \in K$ that $\frac{\|x-\lambda y\|}{\|x\|} \geq 1$. Hence $\|x - \lambda y\| \geq \|x\|$. □

One can use lemma 2.6.18 to compute $d_{\mathbb{P}}$:

Lemma 2.7.7. *Let $x, y \in V\backslash\{0\}$ and T_y be an orthogonal complement of Ky. Compute μ such that $\mu x \in y + T_y$. If no such μ exists, then $d_{\mathbb{P}}(Kx, Ky) = 1$. Otherwise $d_{\mathbb{P}}(Kx, Ky) = \frac{\|\mu x - 1 \cdot y\|}{\|\mu x\|}$.*

Proof. If there exists no such μ, then $x \in T_y$ and by lemma 2.7.6 $d_{\mathbb{P}}(Kx, Ky) = 1$. Now suppose that $\mu x \in y + T_y$. Lemma 2.6.18 states that μx is a best approximation of y in Kx. Hence $d_{\mathbb{P}}(Kx, Ky) = \min_\lambda \frac{\|x - \lambda y\|}{\|x\|} = \min_\lambda \frac{\|\mu x - \lambda y\|}{\|\mu x\|} = \frac{\|\mu x - 1 \cdot y\|}{\|\mu x\|}$. □

Example 2.7.8. Let $K = \mathbb{Q}_p$, $V = K^3$ and $x = \begin{pmatrix} 1 \\ p \\ p^2 \end{pmatrix}, \quad y = \begin{pmatrix} 1+p \\ 1 \\ p \end{pmatrix}$.

As $\|y\| = |y_1| = 1$ we can choose $T = \mathrm{Span}\{e_2, e_3\}$ as an orthogonal complement for Ky. By doing so we have $(1+p)x \in y + T$. Therefore by lemma 2.7.7

$$d_{\mathbb{P}}(Kx, Ky) = \frac{\|(1+p)x - y\|}{\|(1+p)x\|} = \left\| \begin{pmatrix} 0 \\ -1+p+p^2 \\ -p+p^2+p^3 \end{pmatrix} \right\| = 1.$$

It follows that $x \perp y$. Since $\|y\| = |y_2| = 1$ we may also choose $T = \mathrm{Span}\{e_1, e_3\}$ as the orthogonal complement for Ky. Then $p^{-1}x \in y + T$ and by lemma 2.7.7

$$d_{\mathbb{P}}(Kx, Ky) = \frac{\|p^{-1}x - y\|}{\|p^{-1}x\|} = p^{-1} \left\| \begin{pmatrix} p^{-1} - 1 - p \\ 0 \\ 0 \end{pmatrix} \right\| = 1.$$

This again shows that $x \perp y$.

Example 2.7.9. Let $K = \mathbb{Q}_p$, $V = K^3$ and $x = \begin{pmatrix} 1 \\ 1+p \\ p^2 \end{pmatrix}, \quad y = \begin{pmatrix} 1+p \\ 1 \\ p \end{pmatrix}$.

Again we can choose $T = \mathrm{Span}\{e_2, e_3\}$ as an orthogonal complement for Ky. We have that $(1+p)x \in y + T$. Therefore by lemma 2.7.7

$$d_{\mathbb{P}}(Kx, Ky) = \frac{\|(1+p)x - y\|}{\|(1+p)x\|} = \left\| \begin{pmatrix} 0 \\ 2p+p^2 \\ -p+p^2+p^3 \end{pmatrix} \right\| = p^{-1}.$$

In particular, $x \not\perp y$. If we choose $T = \mathrm{Span}\{e_1, e_3\}$ as the orthogonal complement for Ky, then $(1+p)^{-1}x \in y + T$:

$$d_{\mathbb{P}}(Kx, Ky) = \frac{\|(1+p)^{-1}x - y\|}{\|(1+p)^{-1}x\|} = \left\| (1+p)^{-1} \cdot \begin{pmatrix} -2p - p^2 \\ 0 \\ -p \end{pmatrix} \right\| = p^{-1}.$$

This again shows that $x \perp y$.

3 The p-adic Newton method

The whole chapter follows [4, sec. 15] and [2, chap. 8]. We will convert the results of [4] and [2] to the p-adic world. The p-adic γ-constant and a p-adic version of Smales γ-theorem will be introduced. The original γ-theorem in [2] states that for an analytic function f and a zero ζ, there is an explicit constant $\gamma(f, \zeta)$ such that the Newton algorithm converges for any starting value z with $\|z - \zeta\| \, \gamma(f, \zeta) < \frac{3-\sqrt{7}}{2}$ quadratically fast.

If not explicitly stated otherwise, K denotes a complete and algebraically closed extension of $\mathbb{Q}_p$ (i.e. $K = \mathbb{C}_p$ or $K = \Omega_p$) and $V = K^n$, $n \in \mathbb{N}$. We will begin with a multidimensional version of Newton's algorithm. The Hensel lifting in theorem 2.4.1 measured the quality of root approximation with the norm of $f(z)$. In this chapter we will take another approach and try to measure the distance of an approximation to a zero itself. This will give a better rate of convergence on the one hand. On the other hand possible starting values will face more rigid restrictions.

3.1 Multivariate Newton method

Definition 3.1.1 (Newton Method: p-adic version)**.** Let $f : V \to V$ be analytic. Let $V \supset U = \{z \in V : Df(z) \text{ is invertible}\}$. The Newton operator with respect to f is defined as

$$N_f : U \to V, \quad v \mapsto N_f(v) = v - Df(v)^{-1} f(v).$$

For some $z = z_0 \in K^n$ such that

- $Df(z)$ is invertible,
- for all $i \geq 1$: $z_i = N_f(z_{i-1})$ and $Df(z_i)$ is invertible,

we call $(z_i)_{i \geq 0}$ the (affine) newton sequence with respect to f and starting value z_0. When constructing the sequence (z_i) we say that we apply the Newton algorithm with respect to f to the starting value z_0. $N_f^i(z)$ denotes the i-th iterate of N_f applied to z.

Recall from definition 2.2.11 that we have defined the operator norm for a multilinear map $\varphi : \underbrace{V \times \ldots \times V}_{k \text{ times}} \to V$ as

$$\|\varphi\| := \max_{\substack{v_i \in \mathbb{S}(V) \\ 1 \le i \le k}} \|\varphi(v_1, \ldots, v_k)\| .$$

Given a zero $\zeta \in K$ of f we want to investigate possible radii of convergence, where we can start the Newton algorithm and be sure that the sequence converges to ζ. We now define the p-adic γ-constant.

Definition 3.1.2 (p-adic γ-constant)**.** Let $f : V \to V$ be analytic and $z \in V$. If $Df(z)$ is invertible, we set

$$\gamma(f, z) := \sup_{k \ge 2} \left\| \frac{Df(z)^{-1} D^k f(z)}{k!} \right\|^{\frac{1}{k-1}} .$$

Else $\gamma(f, z) := \infty$.

Up to the fact that we are working with p-adic numbers, this definition is basically the same as the one in [2].

Proposition 3.1.3. *If $Df(z)$ is invertible, then $\gamma(f, z)$ exists and it is finite.*

Proof. Because f is analytic, the expression

$$C = \sup_{k} \left\| \frac{D^k f(z)}{k!} \right\|^{\frac{1}{k-1}}$$

exists and is finite. Lemma 2.2.13 for any k gives the inequality

$$\left\| \frac{Df(z)^{-1} D^k f(z)}{k!} \right\|^{\frac{1}{k-1}} \le \left\| Df(z)^{-1} \right\|^{\frac{1}{k-1}} \left\| \frac{D^k f(z)}{k!} \right\|^{\frac{1}{k-1}} \le \left(\left\| Df(z)^{-1} \right\| C^k \right)^{\frac{1}{k-1}} .$$

For $k \to \infty$ the right converges towards C. Hence the left side is bounded and $\gamma(f, z)$ is finite. □

Proposition 3.1.4. *Suppose that $Df(\zeta)$ is invertible. Let $z \in V$ such that $u := \gamma(f, \zeta) \|z - \zeta\| < 1$. Then $Df(z)$ is invertible and $\|Df(z)^{-1} Df(\zeta)\| = 1$.*

Proof. Since f is analytic and by proposition 2.5.24 we have that $Df : V \to \mathcal{L}(V, V)$ is analytic. We expand Df around ζ:

$$Df(z) = \sum_{k=0}^{\infty} \frac{1}{k!} D^k Df(\zeta)(z-\zeta)^k. \tag{3.1.5}$$

Multiplying with $Df(\zeta)^{-1}$ from the left yields $Df(\zeta)^{-1}Df(z) = \mathbf{1} + \Delta$, where we have set $\Delta = \sum\limits_{k=1}^{\infty} \frac{1}{k!} Df(\zeta)^{-1} D^k Df(\zeta)(z-\zeta)^k$. Using Proposition 2.2.15 we see that $\left\|Df(\zeta)^{-1}D^k Df(\zeta)\right\| = \left\|Df(\zeta)^{-1}D^{k+1}f(\zeta)\right\|$. Recall that $|k| \leq 1$ for all $k \in \mathbb{Z}$. We can estimate:

$$\begin{aligned}
\|\Delta\| &\leq \max\left\{\left\|1/k!\, Df(\zeta)^{-1} D^k Df(\zeta)(z-\zeta)^k\right\| : k \geq 1\right\} \\
&\leq \max\left\{|k|\, \gamma(f,\zeta)^{k-1} \|z-\zeta\|^{k-1} : k \geq 2\right\} \\
&= \max\left\{|k|\, u^{k-1} : k \geq 2\right\} \\
&< 1.
\end{aligned}$$

In particular for $k \geq 1$: $\left\|\Delta^k\right\| \leq \|\Delta\|^k$. This yields $\lim\limits_{k\to\infty} \Delta^k = 0$. By theorem 2.2.3 the K-vector space $\mathcal{L}(V, V)$ is complete. We can apply proposition 2.5.2: The series $\sum\limits_{k=0}^{\infty} (-\Delta)^k$ converges to $(\mathbf{1} + \Delta)^{-1}$. This shows that $\mathbf{1} + \Delta$ is invertible. Since $(\mathbf{1} + \Delta)^{-1} Df(\zeta)^{-1} = Df(z)^{-1}$, also $Df(z)$ is invertible. Furthermore,

$$\left\|(\mathbf{1} + \Delta)^{-1}\right\| = \max_{k \geq 0} \left\|(-\Delta)^k\right\| = \max\{1, \|\Delta\|, \left\|\Delta^2\right\|, \ldots\} = 1,$$

where we have used proposition 2.3.3. □

The analogue in [4] of the preceding proposition is the central step towards the γ-theorem. In [4] the condition $u < 1 - \frac{\sqrt{2}}{2}$ is required to conclude the inequality $\|Df(z)^{-1}Df(\zeta)\| < 1$. Here we only need $u < 1$ and obtain $\|Df(z)^{-1}Df(\zeta)\| = 1$.

3.2 Smale's γ-theorem: p-adic version

The γ-theorem was first introduced by Smale (see [2, p. 157]) for the complex numbers. We can now state it for the p-adics. Our version will require a $z \in K^n$ such that $\|z - \zeta\| \gamma(f, \zeta) < 1$. While the constant $\gamma(f, \zeta)$ is defined similarly to the one in [2], the upper bound 1 is actually larger than $\frac{3-\sqrt{7}}{2} \approx 0.177$.

Theorem 3.2.1 (Smales γ-theorem: p-adic version)**.** *Let $f : V \to V$ be analytic. Let $\zeta \in V$ with $f(\zeta) = 0$ and assume $Df(\zeta)$ to be invertible. Then if $z \in V$ is such that*

$$u = \gamma(f, \zeta) \|z - \zeta\| < 1,$$

the sequence $(z_i)_{i \geq 0}$ defined by the Newton Algorithm with respect to f and starting value z is well-defined and

$$\forall i \in \mathbb{N} : \|z_i - \zeta\| \leq u^{2^i - 1} \|z - \zeta\| .$$

Proof. Observe that by proposition 3.1.4 $Df(z)$ is invertible, so $N_f(z)$ exists. The difference between $N_f(z)$ and ζ is:

$$\begin{aligned} N_f(z) - \zeta &= z - \zeta - Df(z)^{-1} f(z) \\ &= Df(z)^{-1}(Df(z)(z - \zeta) - f(z)). \end{aligned} \tag{3.2.2}$$

If we expand f around ζ: $f(z) = \sum_{k=1}^{\infty} \frac{1}{k!} D^k f(\zeta)(z - \zeta)^k$ and use equation (3.1.5), the right-hand side of (3.2.2) becomes

$$Df(z)^{-1} \sum_{k=1}^{\infty} \frac{1}{(k-1)!} D^{k-1} Df(\zeta)(z - \zeta)^{k-1}(z - \zeta) - \frac{1}{k!} D^k f(\zeta)(z - \zeta)^k. \tag{3.2.3}$$

By proposition 2.2.15 we know that $D^{k-1} Df(\zeta)(z - \zeta)^{k-1}(z - \zeta) = D^k f(\zeta)(z - \zeta)^k$. Also $\frac{1}{(k-1)!} - \frac{1}{k!} = \frac{k-1}{k!}$. Inserting $\mathbf{1} = Df(\zeta) Df(\zeta)^{-1}$ into (3.2.3), we obtain

$$\begin{aligned} &Df(z)^{-1}(Df(z)(z - \zeta) - f(z)) \\ =&Df(z)^{-1} Df(\zeta) \sum_{k=1}^{\infty} (k-1) \frac{Df(\zeta)^{-1} D^k f(\zeta)}{k!} (z - \zeta)^k. \end{aligned} \tag{3.2.4}$$

Taking norms on both sides of (3.2.4) yields

$$\begin{aligned} \|N_f(z) - \zeta\| &\leq \left\|Df(z)^{-1} Df(\zeta)\right\| \|z - \zeta\| \max_k \; |k - 1| \, u^{k-1} \\ &= \left\|Df(z)^{-1} Df(\zeta)\right\| \|z - \zeta\| \, u, \end{aligned}$$

because $|k-1| \leq 1$ for $k \geq 2$. Using proposition 3.1.4 yields

$$\|N_f(z) - \zeta\| \leq \|z - \zeta\| \, u = \|z - \zeta\|^2 \, \gamma(f, \zeta). \tag{3.2.5}$$

This shows that $\gamma(f,\zeta) \, \|N_f(z) - \zeta\| < 1$. We may apply the Newton operator to $N_f(z)$. By induction we see that the sequence $(z_i)_{i \geq 0}$ defined by the Newton algorithm with respect to f and starting value z exists. For all $i \geq 1$ when replacing z by z_i in (3.2.5) we get

$$\|z_i - \zeta\| \leq \|z_{i-1} - \zeta\|^2 \, \gamma(f, \zeta).$$

Again by induction we conclude that $\|z_i - \zeta\| \leq \|z - \zeta\|^{2^i} \gamma(f,\zeta)^{2^i - 1} = u^{2^i - 1} \|z - \zeta\|$.
□

Remark 3.2.6. While the rate of convergence for Hensel lifting (theorem 2.4.1) was linear, here it is quadratic. Further, the Hensel lifting can only be applied to polynomials over $\mathbb{Z}_p$ (or $\mathcal{O}_K$ respectively), while here we deal with arbitrary analytic functions. If we want to apply the p-adic γ-theorem to polynomials $f \in \mathbb{Z}_p[X]$ no multiple zeros are allowed ($f'(\zeta)$ needs to be invertible). Proposition 3.1.4 then tells us that for points $z \in K^n$ with $\|z - \zeta\| \, \gamma(f, \zeta) < 1$ also $f'(z)$ is invertible. Contrary to that the Hensel lifting only requires $f'(z)$ to be invertible and $\frac{|f(z)|}{|f'(z)|^2} < 1$, but has no restrictions on ζ.

Example 3.2.7. Let $p = 3$, $K = \mathbb{Q}_p$ and $f : K \to K$ with

$$f(X) = X^2 - 1.$$

$\zeta = 1$ is a zero for f and $Df(\zeta) = 2$. We want to compute $\gamma(f, \zeta)$. The polynomial f has degree 2, so $D^k f(\zeta) = 0, k \geq 3$. This yields $\gamma(f,\zeta) = \left\| Df(\zeta)^{-1} \frac{D^2 f(\zeta)}{2} \right\| = \left| \frac{1}{2} \cdot \frac{2}{2} \right| = 1$. Let $z = -5$. Then $|\zeta - z| \, \gamma(f, \zeta) = |6| \cdot 1 = \frac{1}{3} < 1$. The first steps of the Newton sequence with respect to f and starting value z (displayed as fractions) are:

$$
\begin{array}{lcll}
z_0 & = & -5 & |z_0 - \zeta| = |6| = 3^{-1} \\[2ex]
z_1 & = & \frac{-26}{10} & |z_1 - \zeta| = \left|\frac{-36}{10}\right| = 3^{-2} \\[2ex]
z_2 & = & \frac{-388}{260} & |z_2 - \zeta| = \left|\frac{-648}{260}\right| = 3^{-4} \\[2ex]
z_3 & = & \frac{-6817}{6305} & |z_3 - \zeta| = \left|\frac{-52488}{25220}\right| = 3^{-8} \\[2ex]
z_4 & = & \frac{-43112257}{42981185} & |z_4 - \zeta| = \left|\frac{-86093442}{42981185}\right| = 3^{-16} \\[2ex]
z_5 & = & \frac{-1853024483819137}{1853015893884545} & |z_5 - \zeta| = \left|\frac{-3706040377703682}{1853015893884545}\right| = 3^{-32}
\end{array}
$$

(All calculations were made with Maxima 5.28.0.)

Example 3.2.8. Let $p = 2$, $V = \mathbb{Q}_p^2$ and $f : V \to V$ with

$$
f(X, Y) = \begin{pmatrix} X^2 - Y \\ X + Y - 2XY \end{pmatrix}.
$$

$\zeta = (1, 1)$ is a zero for f and

$$
Df(\zeta) = \begin{pmatrix} 2 & -1 \\ -1 & -1 \end{pmatrix}, \quad Df(\zeta)^{-1} = \frac{-1}{3}\begin{pmatrix} -1 & 1 \\ 1 & 2 \end{pmatrix}.
$$

Again we want to compute $\gamma(f, \zeta) = \left\| Df(\zeta)^{-1} \frac{D^2 f(\zeta)}{2} \right\|$. To compute $\gamma(f, \zeta)$ we use lemma 2.2.10: For a bilinear operator $A : V \times V \to V$ we have that $\|A\| = \max\limits_{i,j=1,2} \|A(e_i, e_j)\|$. This yields

$$
\begin{aligned}
\gamma(f, \zeta) &= \max_{i,j=1,2} \left\| Df(\zeta)^{-1} \frac{D^2 f(\zeta)}{2}(e_i, e_j) \right\| \\
&= \max_{i,j=1,2} \left\| Df(\zeta)^{-1} \frac{1}{2} \frac{\partial^{e_i + e_j} f}{\partial X^{e_i + e_j}}(\zeta) \right\|.
\end{aligned}
$$

We therefore must compute the second order derivatives of f:

$$
\frac{\partial^{e_1 + e_1} f}{\partial X^{e_1 + e_1}}(\zeta) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad \frac{\partial^{e_1 + e_2} f}{\partial X^{e_1 + e_2}}(\zeta) = \begin{pmatrix} 0 \\ -2 \end{pmatrix}, \quad \frac{\partial^{e_2 + e_2} f}{\partial X^{e_2 + e_2}}(\zeta) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.
$$

This yields

$$\begin{aligned}\gamma(f,\zeta) &= \max\left\{\left\|Df(\zeta)^{-1}\begin{pmatrix}1\\0\end{pmatrix}\right\|, \left\|Df(\zeta)^{-1}\begin{pmatrix}0\\-1\end{pmatrix}\right\|, \left\|Df(\zeta)^{-1}\begin{pmatrix}0\\0\end{pmatrix}\right\|\right\}\\ &= \max\left\{\left\|\frac{-1}{3}\begin{pmatrix}-1\\1\end{pmatrix}\right\|, \left\|\frac{-1}{3}\begin{pmatrix}-1\\-2\end{pmatrix}\right\|, \left\|\begin{pmatrix}0\\0\end{pmatrix}\right\|\right\}\\ &= 1.\end{aligned}$$

Let $z = (3,-1)$. Then $\|\zeta - z\|\,\gamma(f,\zeta) = 2^{-1}\cdot 1 < 1$. The first steps of the Newton sequence with respect to f and starting value z (displayed as fractions) are:

$$\begin{aligned} z_0 &= \begin{pmatrix}3\\-1\end{pmatrix} & \|z_0-\zeta\| &= \left\|\begin{pmatrix}-2\\-2\end{pmatrix}\right\| = 2^{-1}\\ z_1 &= \begin{pmatrix}\frac{13}{9}\\ \frac{-1}{3}\end{pmatrix} & \|z_1-\zeta\| &= \left\|\begin{pmatrix}\frac{4}{9}\\ \frac{-12}{9}\end{pmatrix}\right\| = 2^{-2}\\ z_2 &= \begin{pmatrix}\frac{2171}{2763}\\ \frac{169}{921}\end{pmatrix} & \|z_2-\zeta\| &= \left\|\begin{pmatrix}\frac{-592}{2763}\\ \frac{-752}{921}\end{pmatrix}\right\| = 2^{-4}\\ z_3 &= \begin{pmatrix}\frac{13524659161}{5591016153}\\ \frac{5933970419}{1863672051}\end{pmatrix} & \|z_3-\zeta\| &= \left\|\begin{pmatrix}\frac{7933643008}{5591016153}\\ \frac{4070298368}{1863672051}\end{pmatrix}\right\| = 2^{-8}\end{aligned}$$

(All calculations were made with Maxima 5.28.0.)

In order to end this chapter we give a corollary from the γ-theorem.

Corollary 3.2.9. *Let $f : V \to V$ be analytic. Let $\zeta \neq \xi$ be two zeros of f such that both $Df(\zeta), Df(\xi)$ are invertible. Then*

$$\|\zeta - \xi\| \min\{\gamma(f,\zeta), \gamma(f,\xi)\} \geq 1.$$

Proof. Without loss of generality let $\min\{\gamma(f,\zeta),\gamma(f,\xi)\} = \gamma(f,\zeta)$. Suppose that $\|\zeta-\xi\|\,\gamma(f,\zeta) < 1$. By theorem 3.2.1 we have that the Newton sequence with respect to f and starting value ξ converges to ζ. However, $N_f(\xi) = \xi$ and consequently $\xi = \zeta$. □

4 The p-adic Newton method: projective version

The whole chapter follows [4, sec. 16] and [2, chap 14], where a projective version of the Newton algorithm is introduced. Again we want to convert the results to the p-adic world. The main issue that has to be solved is that in [4] orthogonality in $\mathbb{C}^n$ plays a crucial role. This concept can not be translated one-to-one as the p-adic norm does not come from an inner product. However, we can use the definition of orthogonality of section 2.6 and surprisingly receive results similar to the ones in [4]. If not explicitly stated otherwise, let K be a complete and algebraically closed extension of $\mathbb{Q}_p$ (i.e. $K = \mathbb{C}_p$ or $K = \Omega_p$). In this chapter we turn away from analytic functions in general and focus on polynomials again. Given a system of n polynomials in n variables, we will homogenize it by introducing one further variable. By doing so, we can make some nice estimates on the norms appearing.

4.1 The projective Newton operator

Definition 4.1.1. Let $d \in \mathbb{N}$. Define $\mathcal{H}_d$ as the set of homogeneous polynomials over $\mathbb{Q}_p$ in $n+1$ variables of degree d. Let $\underline{d} = (d_1, \ldots, d_n)$. Then we define

$$\mathcal{H}_{\underline{d}} := \mathcal{H}_{d_1} \times \ldots \times \mathcal{H}_{d_n}.$$

In the following let $\underline{d} \in \mathbb{N}^n$ be fixed. Every $f \in \mathcal{H}_{\underline{d}}$ defines an analytic function

$$f : K^{n+1} \to K^n$$

and for every $z \in K^{n+1}$ the derivative of f at z is a linear map $Df(z) : K^{n+1} \to K^n$.

Notation 4.1.2. For a homogeneous polynomial $f \in \mathcal{H}_d$ we use the multiindex notation.

$$f = \sum_{\substack{\alpha \in \mathbb{N}^{n+1} \\ |\alpha| = d}} \lambda_\alpha X^\alpha, \lambda_\alpha \in \mathbb{Q}_p,$$

where again $X^\alpha := X_0^{\alpha_0} X_1^{\alpha_1} \dots X_n^{\alpha_n}$.

Using the results of section 2.6, we can define a projective newton operator.

Definition 4.1.3. Let $z \in K^{n+1}$ and T_z be an orthogonal complement of Kz. Suppose that $Df(z)|_{T_z}$ is invertible. The projective newton operator associated to f and T_z is defined as

$$N_f(z, T_z) = z - Df(z)|_{T_z}^{-1} f(z).$$

A sequence $(z_i, T_i)_{i \geq 0}$ is called a (projective) newton sequence with respect to f if

- $z_0 \in K^{n+1}$, T_0 is an orthogonal complement of Kz_0 and $Df(z_0)|_{T_0}$ is invertible.
- For all $i \geq 1$: $z_i = N_f(z_{i-1}, T_{i-1})$ and T_i is an orthogonal complement of Kz_i such that $Df(z_i)_{T_i}$ is invertible.

Note that for any $\lambda \in K^*$ $N_f(\lambda z) = \lambda N_f(z)$. Let $\mathcal{U} \subset \mathbb{P}(K^{n+1})$ be the subset with $y \in \mathcal{U}$, if and only if $Df(z)$ is invertible for all $z \in y$. Then $N_f(z)$ can be interpreted as a function from $\mathcal{U}$ to $\mathbb{P}(K^{n+1})$. Depending on the choice of the orthogonal complements for the same starting value we can have different newton sequences. Nevertheless, at the end of this chapter we will see that under certain circumstances newton sequences converge independently of the choice of the orthogonal complements.

Definition 4.1.4. Let $f \in \mathcal{H}_{\underline{d}}$, $z \in K^{n+1}$ and T_z be an orthogonal complement of Kz. Suppose that $Df(z)|_{T_z}$ is invertible. We define

$$\gamma(f, z, T_z) := \|z\| \max_{k \geq 2} \left\| Df(z)|_{T_z}^{-1} \frac{D^k f(z)}{k!} \right\|^{\frac{1}{k-1}}.$$

And $\gamma(f, z, T_z) = \infty$, if $Df(z)|_{T_z}$ is not invertible.

Note that due to the homogeneity of f for all $\lambda \in K^*$:

$$\gamma(f, \lambda z, T_z) = \gamma(f, z, T_z) \text{ and } \gamma(\lambda f, z, T_z) = \gamma(f, z, T_z)$$

We can view γ as function on $\mathcal{W} \subset \mathbb{P}(\mathcal{H}_{\underline{d}}) \times \mathbb{P}(K^{n+1}) \times \mathcal{T}$, where $\mathcal{T}$ is the set of n-dimensional linear subspaces of K^{n+1} and $\mathcal{W}$ is the subset with $(f, z, T) \in \mathcal{W}$, if and only if T is an orthogonal complement of Kz.

Lemma 4.1.5. *Let $f \in \mathcal{H}_{\underline{d}}$ and $f(\zeta) = 0$. Then $Df(\zeta)\zeta = 0$.*

Proof. Let $f = \sum\limits_{\substack{\alpha \in \mathbb{N}^{n+1} \\ |\alpha| = d}} \lambda_\alpha X^\alpha \in \mathcal{H}_d$. Then Eulers formula states

$$\sum_{i=0}^{n} X_i \frac{\partial f}{\partial X_i} = d \cdot f.$$

For $f = (f_1, \ldots, f_n) \in \mathcal{H}_{\underline{d}}$ we therefore have $Df(\zeta)\zeta = d \cdot f(\zeta) = 0$. □

The following proposition deals with the independence of γ of the orthogonal complement.

Proposition 4.1.6. *Let $f \in \mathcal{H}_{\underline{d}}$, $\zeta \in K^{n+1}$ be a zero of f and T_1, T_2 be two orthogonal complements of $K\zeta$. Assume that $Df(\zeta)$ has full rank. Then $\gamma(f, \zeta, T_1) = \gamma(f, \zeta, T_2) < \infty$.*

Proof. Since $Df(\zeta)$ has full rank and $Df(\zeta)\zeta = 0$, we conclude that both $Df(\zeta)|_{T_1}$ and $Df(\zeta)|_{T_2}$ are invertible. It is therefore sufficient to show that for any $x \in K^n$ $\left\|Df(\zeta)|_{T_1}^{-1} x\right\| = \left\|Df(\zeta)|_{T_1}^{-1} x\right\|$. Let x_1, x_2 be preimages of x in T_1, T_2 respectively. Then $Df(\zeta)(x_1 - x_2) = 0$ and consequently $x_1 - x_2 \in K\zeta$. We can apply lemma 2.6.17 to conclude that $\|x_1\| = \|x_2\|$. □

The preceding proposition now allows us to justify the notation $\gamma(f, \zeta)$ instead of $\gamma(f, \zeta, T_\zeta)$.

We now want to state a result that is similar to that in proposition 3.1.4. It will lead us to a projective version of the p-adic γ-theorem (theorem 3.1.2). If in the following we write T_x in the context of an $x \in K^{n+1}$, we mean an arbitrary but fixed orthogonal complement of Kx. Let $\zeta \neq 0$ be a zero of f. Throughout the whole chapter we will keep this notion. Due to the homogeneity of f we can assume $\zeta \in \mathbb{S}(K^{n+1})$. We will always assume $Df(\zeta)$ to have full rank. Take the Taylor expansion of Df at ζ:

$$Df(z) = \sum_{k=1}^{\infty} \frac{1}{(k-1)!} D^k f(\zeta)(z - \zeta)^{k-1}$$

(Observe that, once again, we have used proposition 2.2.15.) If we restrict $Df(z)$ to T_z and compose with $Df(\zeta)|_{T_\zeta}^{-1}$ from the left, we obtain

$$Df(\zeta)|_{T_\zeta}^{-1} Df(z)|_{T_z} = Df(\zeta)|_{T_\zeta}^{-1} Df(\zeta)|_{T_z} + \sum_{k=2}^{\infty} \frac{Df(\zeta)|_{T_\zeta}^{-1} D^k f(\zeta)(z-\zeta)^{k-1}|_{T_z}}{(k-1)!}.$$

We write $Df(\zeta)|_{T_\zeta}^{-1} Df(z)|_{T_z} = P + \Delta$, where

$$\begin{aligned} P &= Df(\zeta)|_{T_\zeta}^{-1} Df(\zeta)|_{T_z} \\ \Delta &= \sum_{k=2}^{\infty} \frac{Df(\zeta)|_{T_\zeta}^{-1} D^k f(\zeta)(z-\zeta)^{k-1}|_{T_z}}{(k-1)!}. \end{aligned}$$

The following lemmata deal with $\|P\|$ and $\|\Delta\|$.

Lemma 4.1.7. $\|P\| \leq 1$

Proof. If $x \in T_z$, $\|x\| = 1$, write it as $x = \lambda\zeta + \zeta'$, where $\zeta' \in T_\zeta, \lambda \in K$. Observe that $\max\{|\lambda\zeta|, |\zeta'|\} = \|x\| = 1$. Then

$$Px = Df(\zeta)|_{T_\zeta}^{-1} Df(\zeta)(\lambda\zeta + \zeta') = \zeta'.$$

It follows that $\|Px\| \leq 1$ and consequently $\|P\| \leq 1$. □

Lemma 4.1.8. *If $\|z - \zeta\| < 1$, then $P : T_z \to T_\zeta$ is injective, hence invertible.*

Proof. Since $\|z - \zeta\| < 1$, it must be $z \not\perp \zeta$ (see lemma 2.7.6). Hence $K\zeta \notin T_z$. Consequently $\ker D(\zeta) \cap T_z = 0$ and P is injective. □

Lemma 4.1.9. *If $\|z - \zeta\| < 1$, then $\|P\| = \|P^{-1}\| = 1$.*

Remark 4.1.10. In the corresponding lemma in [4] one had $\|P^{-1}\| = (\cos\delta)^{-1} \geq 1$, where δ was the spherical distance between z and ζ.

Proof of lemma 4.1.9. The equation $1 = \|\mathbf{1}\| \leq \|P\| \|P^{-1}\|$ combined with lemma 4.1.7 shows that it suffices to prove $\|P^{-1}\| \leq 1$. We have that $\|z - \zeta\| < 1$ and $\|\zeta\| = 1$. Using corollary 2.3.4 we can deduce that also $\|z\| = 1$. Write $\zeta = \mu z + z'$, where $z' \in T_z$. Lemma 2.6.19 tells us that $|\mu| = 1$. The equation $Df(\zeta) = 0$ then yields $Df(\zeta)z = -\mu^{-1} Df(\zeta)z'$. Finally, let $x \in T_\zeta$, $\|x\| = 1$ and write it as $x = \lambda z + z''$, where $z'' \in T_z$. Observe that $\max\{|\lambda|, \|z''\|\} = 1$. Then we apply P^{-1} to

x and obtain $P^{-1}x = -\mu^{-1}\lambda z' + z''$. This yields $\|P^{-1}x\| \leq \max\{\|\mu^{-1}\lambda z'\|, \|z''\|\} \leq 1$, and consequently $\|P^{-1}\| \leq 1$. □

From now on we consider $z \in K^{n+1}$ with $u := \gamma(f,\zeta)\,\|z-\zeta\| < 1$.

Lemma 4.1.11. *If $u = \gamma(f,\zeta)\,\|z-\zeta\| < 1$, then $\|\Delta\| \leq u$.*

Proof.

$$\begin{aligned}
\|\Delta\| &\leq \max_{k\geq 2} \left\| \frac{Df(\zeta)|_{T_\zeta}^{-1} D^k f(\zeta)(z-\zeta)^{k-1}|_{T_z}}{(k-1)!} \right\| \\
&\leq \max_{k\geq 2} \left\| \frac{Df(\zeta)|_{T_\zeta}^{-1} D^k f(\zeta)(z-\zeta)^{k-1}}{(k-1)!} \right\| \\
&\leq \max_{k\geq 2} \gamma(f,\zeta)^{k-1}\, \|z-\zeta\|^{k-1}\, |k| \\
&= u,
\end{aligned}$$

where we again have used that $|k| \leq 1$ for all $k \in \mathbb{Z}$. □

The following proposition can be viewed as the projective version of proposition 3.1.4. A corresponding proposition can be found in [4, sec. 16.6].

Proposition 4.1.12. *Let $z \in K^{n+1}\backslash\{0\}$ such that both $u = \gamma(f,\zeta)\,\|z-\zeta\| < 1$ and $\|z-\zeta\| < 1$. Then $Df(z)$ restricted to any orthogonal complement of Kz is invertible. Furthermore, if we fix an orthogonal complement T_z, we have*

$$\left\|Df(z)|_{T_z}^{-1} Df(\zeta)|_{T_\zeta}\right\| = \left\|Df(\zeta)|_{T_\zeta}^{-1} Df(z)|_{T_z}\right\| = 1.$$

Proof. Let T_z be a arbitrary but fixed orthogonal complement of Kz. We use the notation above:

$$Df(\zeta)|_{T_\zeta}^{-1} Df(z)|_{T_z} = P + \Delta.$$

Using lemma 4.1.9 and lemma 4.1.11 we see that $\|P\| = 1 > u \geq \|\Delta\|$, from which follows that $\left\|Df(\zeta)|_{T_\zeta}^{-1} Df(z)|_{T_z}\right\| = \|P+\Delta\| = 1$. Furthermore,

$$\left\|P^{-1}\Delta\right\| \leq \left\|P^{-1}\right\| \|\Delta\| \leq u < 1.$$

Thus the series $\sum_{i=0}^{\infty} -(P^{-1}\Delta)^i$ converges to $(1+P^{-1}\Delta)^{-1}$. Multiplying with P^{-1} from the right yields $(\mathbf{1}+P^{-1}\Delta)^{-1}P^{-1} = (P+\Delta)^{-1}$. Therefore $Df(z)|_{T_z}$ is invertible and

$$\begin{aligned}\left\|Df(z)|_{T_z}^{-1}Df(\zeta)\right\| &= \left\|(P+\Delta)^{-1}\right\| \\ &\le \left\|(\mathbf{1}+P^{-1}\Delta)^{-1}\right\| \left\|P^{-1}\right\| \\ &\le \left\|P^{-1}\right\| \max_i \left\|P^{-1}\Delta\right\|^i \\ &= 1.\end{aligned}$$

On the other hand we have that

$$1 = \|\mathbf{1}\| \le \left\|Df(z)|_{T_z}^{-1}Df(\zeta)|_{T_\zeta}\right\| \left\|Df(\zeta)|_{T_\zeta}^{-1}Df(z)|_{T_z}\right\| \le \left\|Df(z)|_{T_z}^{-1}Df(\zeta)|_{T_\zeta}\right\|,$$

which proves $\left\|Df(z)|_{T_z}^{-1}Df(\zeta)|_{T_\zeta}\right\| = 1$. □

Corollary 4.1.13. *Let $z \in K^{n+1}\backslash\{0\}$ such that both $u = \gamma(f,\zeta)\|z-\zeta\| < 1$ and $\|z-\zeta\| < 1$. Then for any two orthogonal complements T_1, T_2 of Kz*

$$\left\|Df(z)|_{T_1}^{-1}Df(z)|_{T_2}\right\| \le 1.$$

Proof. Choose an orthogonal complement T_ζ of $K\zeta$ and write

$$Df(z)|_{T_1}^{-1}Df(z)|_{T_2} = Df(z)|_{T_1}^{-1}Df(\zeta)|_{T_\zeta}Df(\zeta)|_{T_\zeta}^{-1}Df(z)|_{T_2}.$$

Then use lemma 2.2.14 and proposition 4.1.12. □

The next corollary extends proposition 4.1.6 to a neighborhood of ζ.

Corollary 4.1.14. *Let $z \in K^{n+1}\backslash\{0\}$ such that both $u = \gamma(f,\zeta)\|z-\zeta\| < 1$ and $\|z-\zeta\| < 1$. Then for any two orthogonal complements T_1, T_2 of Kz*

$$\gamma(f,z,T_1) = \gamma(f,z,T_2).$$

Proof. Proposition 4.1.12 tells us that $Df(z)$ restricted to either T_1 or T_2 is invertible. Consider

$$\gamma(f,z,T_1) = \max_{k\ge 2}\left\|Df(z)|_{T_1}^{-1}\frac{D^kf(z)}{k!}\right\|^{\frac{1}{k-1}}.$$

Inserting $Df(z)|_{T_2}Df(z)|_{T_2}^{-1} = \mathbf{1}$ and using lemma 2.2.14 for any $k \ge 2$ we can estimate

$$\left\| Df(z)|_{T_1} \frac{D^k f(z)}{k!} \right\| \leq \left\| Df(z)|_{T_2}^{-1} \frac{D^k f(z)}{k!} \right\|.$$

Interchanging the roles of T_1 and T_2 yields the opposite inequality. □

Now we let K be a finite algebraic extension of $\mathbb{Q}_p$. Then we can use the metric defined in 2.7.2 and state a theorem that is similar to theorem 3.2.1 and is a p-adic version to the projective γ-theorem in [4, sec. 16.6].

Theorem 4.1.15. *Let K be a finite algebraic extension of $\mathbb{Q}_p$. Let $f \in \mathcal{H}_{\underline{d}}$ and $\zeta \in \mathbb{S}(K^{n+1})$ be a zero of f such that $Df(\zeta)$ has full rank. Let $d_{\mathbb{P}}$ the metric on $\mathbb{P}(K^{n+1})$ as defined in 2.7.2. Suppose for some $z \in K^{n+1} \backslash \{0\}$ that $d_{\mathbb{P}}(Kz, K\zeta) < 1$ and $u := d_{\mathbb{P}}(Kz, K\zeta)\gamma(f, \zeta) < 1$. If $z_0 = z$, any newton sequence $(z_i, T_i)_{i \geq 0}$ is defined and*

$$d_{\mathbb{P}}(Kz_i, K\zeta) \leq u^{2^i - 1} d_{\mathbb{P}}(Kz, K\zeta).$$

Proof. First observe that by proposition 4.1.12, $Df(z)$ restricted to any orthogonal complement of Kz is invertible. Fix an orthogonal complement T_z of Kz to use for the p-adic Newton operator. After scaling we can assume that $z \in \mathbb{S}(K^{n+1})$. From the proof of proposition 2.7.4 we know that

$$d_{\mathbb{P}}(Kz, K\zeta) = \min_{\lambda \in K} \frac{\|z - \lambda\zeta\|}{\|z\|} = \min_{\lambda \in \mathbb{S}(K)} \|z - \lambda\zeta\|.$$

It follows that there exists some $\mu \in \mathbb{S}(K)$ with $d_{\mathbb{P}}(Kz, K\zeta) = \|z - \mu\zeta\| < 1$. We write $\mu\zeta = \rho z + z'$ with $z' \in T_z$. Lemma 2.6.19 tells us that $|\rho| = 1$. If we set $\xi = \rho^{-1}\mu\zeta$, we obtain a representative $\xi \in K\zeta \cap (z + T_z)$ with $|\xi| = 1$. To measure the distance between $N_f(z)$ and ξ we proceed as for equation (3.2.4):

$$\begin{aligned} N_f(z) - \xi &= z - \xi - Df(z)|_{T_z}^{-1} f(z) \\ &= Df(z)|_{T_z}^{-1} \left(Df(z)(z - \xi) - f(z) \right) \\ &= Df(z)|_{T_z}^{-1} Df(\xi) \left(\sum_{k=2}^{\infty} (k-1) Df(\xi)|_{T_\xi}^{-1} \frac{D^k(\xi)}{k!} (z - \xi)^k \right). \end{aligned}$$

By Lemma 2.6.18 we have $d_{\mathbb{P}}(Kz, K\xi) = \|z - \xi\|$. This yields, if we set $\hat{z} = N_f(z, T_z)$ and use $\gamma(f, \zeta) = \gamma(f, \xi)$ as well as proposition 4.1.12:

$$d_{\mathbb{P}}(K\hat{z}, K\xi) \leq \|N_f(z) - \xi\|$$

$$
\begin{aligned}
&\leq \left\| Df(z)|_{T_z}^{-1} Df(\xi)|_{T_\xi} \right\| \max_{k\geq 2} \left\| (k-1) Df(\xi)|_{T_\zeta}^{-1} \frac{D^k(\xi)}{k!} (z-\zeta)^k \right\| \\
&\leq \max_{k\geq 2} |k-1|\, u^{k-1} d_{\mathbb{P}}(Kz, K\xi) \\
&\leq u\ d_{\mathbb{P}}(Kz, K\xi). \qquad (4.1.16)
\end{aligned}
$$

The assumptions for $N_f(z)$ to apply proposition 4.1.12 are fulfilled. We can deduce that $Df(\hat{z})$ restricted to any orthogonal complement $T_{\hat{z}}$ of $K\hat{z}$ is invertible. Hence $N_f(\hat{z}, T_{\hat{z}})$ is defined. Consequently, any Newton sequence $(z_i, T_i)_{i\geq 0}$ is defined. Replacing z by z_i in 4.1.16 yields

$$
d_{\mathbb{P}}(Kz_{i+1}, K\xi) \leq \gamma(f, \xi) d_{\mathbb{P}}(Kz_i, K\xi)^2.
$$

As in the proof for theorem 3.2.1 we deduce for any $i \geq 1$:

$$
d_{\mathbb{P}}(Kz_i, K\xi) \leq u^{2^i-1} d_{\mathbb{P}}(Kz, K\xi).
$$

Using $K\zeta = K\xi$ completes the proof. □

Example 4.1.17. Let $p = 3$, $K = \mathbb{Q}_p$ and $f : K^2 \to K$ with

$$
f(X, Y) = X^2 - Y^2.
$$

(This is the homogenized version of example 3.2.7). $\zeta = (1, 1)$ is a zero for f and $Df(\zeta) = \begin{pmatrix} 2 & -2 \end{pmatrix}$ has full rank. As an orthogonal complement for $K\zeta$ we choose $T = \operatorname{Span}\{e_1\}$. This way $Df(\zeta)|_T^{-1} = \left(\lambda \mapsto \frac{\lambda}{2} e_1\right)$. For computing $\gamma(f, \zeta) = \gamma(f, \zeta, T)$ again we only have to consider the second order derivative. This gives

$$
\begin{aligned}
\gamma(f, \zeta) &= \|\zeta\| \left\| Df(\zeta)|_T^{-1} \frac{D^2 f(\zeta)}{k!} \right\| \\
&= 1 \cdot \max_{i,j=1,2} \left\| Df(\zeta)|_T^{-1} \frac{D^2 f(\zeta)}{k!} (e_i, e_j) \right\| \\
&= \max \left\| \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix} \right\|, \left\| \begin{pmatrix} \frac{-1}{2} \\ 0 \end{pmatrix} \right\|, \left\| \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\| \\
&= 1.
\end{aligned}
$$

Let $z = (-1, 2)$. For computing $d_{\mathbb{P}}(K\zeta, Kz)$, we use lemma 2.6.18: We search for a

$\lambda \in K$ such that $\lambda(-1,2) \in \zeta + K(1,0) = \{(1,1) + \mu(1,0) : \mu \in K\}$. Clearly $\lambda = \frac{1}{2}$ satisfies the property. This yields

$$d_{\mathbb{P}}(K\zeta, Kz) = \frac{\left\|\frac{1}{2}z - \zeta\right\|}{\left\|\frac{1}{2}z\right\|} = \left\|\left(\frac{3}{2}, 0\right)\right\| = 3^{-1}.$$

It follows that $d_{\mathbb{P}}(K\zeta, Kz)\gamma(f,\zeta) = 3^{-1} \cdot 1 < 1$. The first steps of the Newton sequence with respect to f and starting value z (displayed as fractions) are:

$$\begin{array}{llll}
z_0 = \begin{pmatrix} -1 \\ 2 \end{pmatrix} & T_0 = \mathrm{Span}\{e_1\} & d_{\mathbb{P}}(Kz_0, K\zeta) = \left\| \frac{1}{2} \begin{pmatrix} -1 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\| & = 3^{-1} \\
z_1 = \begin{pmatrix} \frac{-5}{2} \\ 2 \end{pmatrix} & T_1 = \mathrm{Span}\{e_1\} & d_{\mathbb{P}}(Kz_1, K\zeta) = \left\| \frac{1}{2} \begin{pmatrix} \frac{-5}{2} \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\| & = 3^{-2} \\
z_2 = \begin{pmatrix} \frac{-41}{20} \\ 2 \end{pmatrix} & T_2 = \mathrm{Span}\{e_2\} & d_{\mathbb{P}}(Kz_2, K\zeta) = \left\| \frac{1}{2} \begin{pmatrix} \frac{-41}{20} \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\| & = 3^{-4} \\
z_3 = \begin{pmatrix} \frac{-41}{20} \\ \frac{3281}{1600} \end{pmatrix} & T_3 = \mathrm{Span}\{e_2\} & d_{\mathbb{P}}(Kz_3, K\zeta) = \left\| \frac{-20}{41} \begin{pmatrix} \frac{-41}{20} \\ \frac{3281}{1600} \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\| & = 3^{-8}
\end{array}.$$

(All calculations were made with Maxima 5.28.0.)

Finally we give a corollary similar to corollary 3.2.9.

Corollary 4.1.18. *Let K be a finite algebraic extension of $\mathbb{Q}_p$. Let $f \in \mathcal{H}_{\underline{d}}$ and $d_{\mathbb{P}}$ be the projective metric on $\mathbb{P}(K^{n+1})$. Let $\zeta, \xi \in \mathbb{S}(K^{n+1}), \zeta \neq \xi$ be two zeros of f. Suppose $d_{\mathbb{P}}(K\xi, K\zeta) < 1$. Then $d_{\mathbb{P}}(Kz, K\zeta)\min\{\gamma(f,\zeta), \gamma(f,\xi)\} \geq 1$.*

4.2 A higher derivative estimate

In the preceding section we learned about the number γ and its relation to radii of convergence for the projective newton algorithm. For computing γ, computing all higher derivatives is required. The goal of this section is to find an upper bound for γ that can be computed an easier way. We will follow the steps in [4, sec. 16.7] and translate them to the p-adics. The proofs can almost be adopted one-to-one.

First, return to K as a complete and algebraically closed extension of $\mathbb{Q}_p$. $\mathcal{H}_d$ is a vector space of dimension $\binom{n+d}{d}$. A basis is, for instance, the system of monomials

$$\{X^\alpha : |\alpha| = d\}, \text{ where } X^\alpha := X_0^{\alpha_0} \cdot \ldots \cdot X_n^{\alpha_n}.$$

We will scale this basis. Let

$$e_\alpha := \binom{d}{\alpha} X^\alpha, \text{ where } \binom{d}{\alpha} := \frac{d!}{\alpha_0! \cdot \ldots \cdot \alpha_n!}.$$

Then, as for K^{n+1}, one defines a norm on $\mathcal{H}_d$ via

$$\left\| \sum_\alpha \lambda_\alpha e_\alpha \right\| = \max_\alpha |\lambda_\alpha|.$$

Recall $\mathcal{H}_{\underline{d}} = \mathcal{H}_{d_1} \times \ldots \times \mathcal{H}_{d_n}$. Every $\mathcal{H}_{d_i}$ is endowed with the basis $\mathcal{B}_i = \{e_\alpha : |\alpha| = d_i\}$. We give $\mathcal{H}_{\underline{d}}$ the basis $\mathcal{B} = \bigcup\limits_{1 \le i \le n} \mathcal{B}_i$. The union is meant to be disjoint. As above we can define the maximum-norm with respect to $\mathcal{B}$ on $\mathcal{H}_{\underline{d}}$.

Lemma 4.2.1. *If $f = (f_1, \ldots, f_n) \in \mathcal{H}_{\underline{d}}$, then $\|f\| = \max\limits_i \|f_i\|$.*

Proof. Write $f_i = \sum\limits_{|\alpha|=d_i} \lambda_{i,\alpha} e_\alpha$. Then $f = \sum\limits_i \sum\limits_{|\alpha|=d_i} \lambda_{i,\alpha} e_\alpha$ and $\|f\| = \max\limits_i \max\limits_{|\alpha|=d_i} \|\lambda_{i,\alpha}\| = \max\limits_i \|f_i\|$. □

An important property of $\| \ \|$ is the following:

Lemma 4.2.2. *Let $f \in \mathcal{H}_{\underline{d}}$ and $z \in K^{n+1}$. Then $\|f(z)\| \le \|f\| \, \|z\|$.*

Proof. Suppose $f = \sum\limits_{|\alpha|=d} \lambda_\alpha e_\alpha \in \mathcal{H}_d, d \in \mathbb{N}$. Then $\|f(z)\| \le \max\limits_{|\alpha|=d} |\lambda_\alpha| \left|\binom{d}{\alpha}\right| \|z\|^d$. Since $\binom{d}{\alpha} \in \mathbb{N}$, we have $\left|\binom{d}{\alpha}\right| \le 1$ and hence $\|f(z)\| \le \|f\| \, \|z\|^d$. Suppose $f = (f_1, \ldots, f_n) \in \mathcal{H}_{\underline{d}}$. Then $\|f(z)\| \le \max\limits_i \|f_i(z)\| \le \max\limits_i \|f_i\| \, \|z\|^d = \|f\| \, \|z\|^d$. □

We will now define a quantity that plays a crucial part in the rest of the work. A similar definition is also made in [4].

Definition 4.2.3. Let $f \in \mathcal{H}_d, z \in K^{n+1} \backslash \{0\}$ and T_z be an orthogonal complement of Kz. If $Df(z)|_{T_z}$ is invertible, we define

$$\mu(f, z, T_z) := \|f\| \left\| Df(z)|_{T_z}^{-1} \operatorname{diag}\left(\|z\|^{d_i - 1} \right) \right\|.$$

Otherwise, we set $\mu(f, z, T_z) = \infty$.

Observe that for any $\lambda \in K^*$: $\mu(\lambda f, z, T_z) = \mu(f, z, T_z)$ and $\mu(f, \lambda z, T_z) = \mu(f, z, T_z)$. As for γ we can view μ as a function on $\mathcal{U}$ (see page 41). Further, we can adopt the proof for proposition 4.1.14 one-to-one to μ to show the following: For $\zeta \in \mathbb{S}(K^{n+1})$ a zero of f such that $\|z - \zeta\| < 1$, $\|z - \zeta\| \gamma(f, \zeta) < 1$, we have for any two orthogonal complements T_1, T_2 of Kz:

$$\mu(f, z, T_1) = \mu(f, z, T_2).$$

Therefore, the notation $\mu(f, z)$ will be convenient. We now can state the main theorem of this section.

Theorem 4.2.4. *For $f \in \mathbb{P}(\mathcal{H}_d), z \in K^{n+1}$ and T_z an orthogonal complement of Kz we have*

$$\gamma(f, z, T_z) \leq \mu(f, z, T_z)$$

The theorem is the p-adic analogue of the main theorem in [4, sec. 16.7]. While in [4] one has the estimate $\gamma \leq \frac{1}{2} D^{\frac{3}{2}} \mu$, where D is the highest homogeneous degree in f, we now have $\gamma \leq 1 \cdot \mu$. For the proof we first will need a couple of lemmas.

Lemma 4.2.5. *Let $f \in \mathcal{H}_d$ and $k \leq d$. All k-th derivatives of f of are homogeneous polynomials of degree $d - k$. Let $w_1, \ldots, w_k \in K^{n+1}$. If we apply the k-linear map $D^k f(X)$ to $(w_1, \ldots, w_k)$, we have that $D^k f(X) w_1, \ldots, w_k) \in \mathcal{H}_{d-k}$ and*

$$\left\| D^k f(X)(w_1, \ldots, w_k) \right\| \leq \|w_1\| \cdot \ldots \cdot \|w_k\| \left| \binom{d}{k} \right| \|f\| .$$

The norms belong to the respective vector spaces.

Proof. Let $f = \sum_\alpha \lambda_\alpha e_\alpha$. The fact that $D^k f(X) w_1, \ldots, w_k) \in \mathcal{H}_{d-k}$ is trivial. We investigate the first derivatives. For a fixed $w \in K^{n+1}$ we have $Df(X)w = \sum_i w_i \frac{\partial f}{\partial X_i}$, so $\|Df(X)w\| \leq \max_i \left\| w_i \frac{\partial f}{\partial X_i} \right\| \leq \|w\| \max_i \left\| \frac{\partial f}{\partial X_i} \right\|$. For a fixed i we can estimate

$$\begin{aligned} \left\| \frac{\partial f}{\partial X_i} \right\| &= \left\| \sum_{\alpha : \alpha_i \neq 0} \lambda_\alpha \binom{d}{\alpha} \alpha_i X_i^{\alpha_i - 1} \prod_{j \neq i} X_j^{\alpha_j} \right\| \\ &= \left\| d \sum_\alpha \lambda_\alpha e_{\alpha - \underline{i}} \right\| \\ &= |d| \, \|f\| . \end{aligned}$$

It follows that $\|Df(X)w\| \leq \|w\| \, |d| \, \|f\|$. Now let $k \geq 1$. For the $(k-1)$-th derivative we have that $Df^{k-1}(X)(w_1, \ldots, w_k) \in \mathcal{H}_{d-k+1}$. Applying the procedure above we conclude that

$$\left\|D^k f(X)(w_1, \ldots, w_k)\right\| \leq \|w_k\| \, |d-k+1| \left\|D^{k-1} f(X)(w_1, \ldots, w_{k-1})\right\|.$$

And consequently $\left\|D^k f(X)(w_1, \ldots, w_k)\right\| \leq \|w_1\| \cdot \ldots \cdot \|w_k\| \left|\frac{d!}{(d-k)!}\right| \|f\|$. □

Lemma 4.2.6. *Let $f \in \mathcal{H}_d, z \in K^{n+1}$. Then $\left\|D^k f(z)\right\| \leq \|z\|^{d-k} \left|\frac{d!}{(d-k)!}\right| \|f\|$.*

Proof. Consider the linear map $\mathrm{eval}_z : \mathcal{H}_d \to K, f \mapsto f(z)$. For any $f \in \mathcal{H}_d$:

$$\|\mathrm{eval}_z(f)\| = \left\|\sum_\alpha \lambda_\alpha z^\alpha\right\| \leq \|z\|^d \max_\alpha |\lambda_\alpha| = \|z\|^d \|f\|.$$

Using eval_z on $D^k f(X)(w_1, \ldots, w_k)$ we obtain

$$\left\|D^k f(z)(w_1, \ldots, w_k)\right\| \leq \|z\|^{d-k} \left\|D^k f(X)(w_1, \ldots, w_k)\right\|.$$

Lemma 4.2.5 tells us that $\left\|D^k f(X)(w_1, \ldots, w_k)\right\| \leq \left|\frac{d!}{(d-k)!}\right| \|f\|$ for any $w_1, \ldots, w_k \in \mathbb{S}(K^{n+1})$. The assertion follows then. □

Lemma 4.2.7. *Let $f = (f_1, \ldots, f_n) \in \mathcal{H}_{\underline{d}}, z \in K^{n+1} \backslash \{0\}$. Then for any k:*

$$\left\|\mathrm{diag}\left(\|z\|^{d_i - k}\right)^{-1} \frac{D^k f(z)}{k!}\right\| \leq \|f\|.$$

Proof. Let $w_1, \ldots, w_k \in \mathbb{S}(K^{n+1})$. Write $w = (w_1, \ldots, w_n)$. Using lemma 4.2.1 and lemma 4.2.6 we obtain

$$\begin{aligned}\left\|\mathrm{diag}\left(\|z\|^{d_i - k}\right)^{-1} \frac{D^k f(z)w}{k!}\right\| &= \left\|\left((\|z\|^{d_i-k})^{-1} \frac{D^k f_i(z)w}{k!}\right)_{i=1}^n\right\| \\ &= \max_i \left\|(\|z\|^{d_i-k})^{-1} \frac{D^k f_i(z)w}{k!}\right\| \\ &\leq \max_i \|z\|^{k-d_i} |k!|^{-1} \left\|D^k f_i(z)w\right\| \\ &\leq \max_i \left|\binom{d_i}{k}\right| \|f_i\|.\end{aligned}$$

Observe that for any i: $\binom{d_i}{k} \in \mathbb{N}$, which means $\left|\binom{d_i}{k}\right| \leq 1$. The claim follows then. □

An immediate consequence is the following:

Corollary 4.2.8. *Let $f \in \mathcal{H}_{\underline{d}}, z \in \mathbb{S}(K^{n+1})$. Then $\|Df(z)\| \leq \|f\|$.*

Lemma 4.2.9. *Let $f \in \mathcal{H}_{\underline{d}}, z \in K^{n+1}$, T_z be an orthogonal complement of Kz. Then $\mu(f, z, T_z) \geq 1$.*

Proof. If $Df(z)|_{T_z}$ is not invertible, then the claim is trivially true. Assume that $Df(z)|_{T_z}$ is invertible. Using lemma 2.2.14 we can estimate

$$\begin{aligned}
1 = \|\mathbf{1}\| &\leq \|f\| \left\| Df(z)|_{T_z}^{-1} \operatorname{diag}\left(\|z\|^{d_i-1}\right) \right\| \|f\|^{-1} \left\| \operatorname{diag}\left(\|z\|^{d_i-1}\right)^{-1} Df(z)|_{T_z} \right\| \\
&\leq \mu(f, z, T_z) \|f\|^{-1} \left\| \operatorname{diag}\left(\|z\|^{d_i-1}\right)^{-1} Df(z) \right\| \\
&\leq \mu(f, z, T_z) \|f\|^{-1} \|f\| \\
&= \mu(f, z, T_z).
\end{aligned}$$

The last inequality comes from corollary 4.2.8. □

Now we can prove the main theorem.

Proof of theorem 4.2.4. Let $f \in \mathcal{H}_{\underline{d}}, z \in K^{n+1} \backslash \{0\}$, T_z be an orthogonal complement of Kz. If $Df(z)|_{T_z}$ is not invertible, then $\gamma(f, z, T_z) = \mu(f, z, T_z) = \infty$ and the claim is true. Otherwise recall that

$$\gamma(f, z, T_z) = \|z\| \max_{k \geq 2} \left\| Df(z)|_{T_z}^{-1} \frac{D^k f(z)}{k!} \right\|^{\frac{1}{k-1}}.$$

For any $k \geq 2$ we can use lemma 4.2.7 to estimate

$$\begin{aligned}
\left\| Df(z)|_{T_z}^{-1} \frac{D^k f(z)}{k!} \right\| &= \left\| Df(z)|_{T_z}^{-1} \operatorname{diag}\left(\|z\|^{d_i-k}\right) \operatorname{diag}\left(\|z\|^{d_i-k}\right)^{-1} \frac{D^k f(z)}{k!} \right\| \\
&\leq \|z\|^{1-k} \left\| Df(z)|_{T_z}^{-1} \operatorname{diag}\left(\|z\|^{d_i-1}\right) \operatorname{diag}\left(\|z\|^{d_i-k}\right)^{-1} \frac{D^k f(z)}{k!} \right\| \\
&\leq \|z\|^{1-k} \mu(f, z, T_z) \|f\|^{-1} \left\| \operatorname{diag}\left(\|z\|^{d_i-k}\right)^{-1} \frac{D^k f(z)}{k!} \right\| \\
&\leq \|z\|^{1-k} \mu(f, z, T_z).
\end{aligned}$$

This yields $\gamma(f, z, T_z) \leq \max\limits_{k \geq 2} \mu(f, z, T_z)^{\frac{1}{k-1}}$. We have $\mu(f, z, T_z)^{\frac{1}{k-1}} \leq \mu(f, z, T_z)$ by lemma 4.2.9. It follows that $\gamma(f, z, T_z) \leq \mu(f, z, T_z)$. □

4.3 A Lipschitz estimate

The corresponding section in [4] has this name, because one has some kind of Lipschitz property when perturbing the inputs of μ. We will follow [4, 16.8] and translate the results to the p-adics. It will turn out that in our case, $\mu(f, z, T_z)$ is local constant in neighborhoods of Kf and of Kz. As usual, we assume that K is a complete algebraically closed extension of $\mathbb{Q}_p$.

Proposition 4.3.1. *Let $f, g \in \mathcal{H}_{\underline{d}}, z \in \mathbb{S}(K^{n+1})$ and T_z be an orthogonal complement of Kz. Suppose that $Df(z)|_{T_z}$ is invertible. Let $d_{\mathbb{P}}$ denote the ultrametric on $\mathbb{P}(\mathcal{H}_{\underline{d}})$ as a $\mathbb{Q}_p$-vector space. If $d_{\mathbb{P}}(Kf, Kg)\mu(f, z, T_z) < 1$, then $\mu(f, z, T_z) = \mu(g, z, T_z)$.*

Proof. Write $A = Df(z)|_{T_z}$ such that $\mu(f, z, T_z) = \|f\| \, \|A^{-1}\|$. Using lemma 4.2.9 we see that $d_{\mathbb{P}}(Kf, Kg) < 1$. Let $\tilde{g} \in Kg$ be a representative such that $\frac{\|f - \tilde{g}\|}{\|f\|} = d_{\mathbb{P}}(Kf, Kg)$. It follows that $\|f - \tilde{g}\| < \|f\|$ and by corollary 2.3.4 we have that $\|f\| = \|\tilde{g}\|$. Let $h = f - \tilde{g}$ and $\Delta = Dh(z)|_{T_z}$. Then

$$D\tilde{g}(z)|_{T_z} = A - \Delta.$$

Corollary 4.2.8 gives $\|\Delta\| \leq \|h\| = \|f - \tilde{g}\|$. Consequently

$$\|\Delta\| \left\|A^{-1}\right\| \leq \frac{\|f - \tilde{g}\|}{\|f\|} \mu(f, z, T_z) = d_{\mathbb{P}}(Kf, Kg)\mu(f, z, T_z) < 1.$$

This implies that $\sum\limits_{i \geq 0} (\Delta A^{-1})^i$ converges to $(\mathbf{1} - \Delta A^{-1})^{-1}$. Further, we have that $\|(\mathbf{1} - \Delta A^{-1})^{-1}\| \leq 1$, which implies

$$\left\|(A - \Delta)^{-1}\right\| \leq \left\|A^{-1}\right\|.$$

But $(A - \Delta)^{-1} = D\tilde{g}(z)|_{T_z}^{-1}$, so $D\tilde{g}(z)|_{T_z}$ must be invertible and

$$\mu(g, z, T_z) = \|\tilde{g}\| \left\|(A - \Delta)^{-1}\right\| \leq \|f\| \left\|A^{-1}\right\| = \mu(f, z, T_z).$$

We have used that $\|f\| = \|\tilde{g}\|$. The assumptions are symmetric in f and g. Hence $\mu(g, z, T_z) \geq \mu(f, z, T_z)$ and consequenlty $\mu(g, z, T_z) = \mu(f, z, T_z)$. □

Proposition 4.3.2. *Let $f \in \mathcal{H}_{\underline{d}}$ and $\zeta \in \mathbb{S}(K^{n+1})$ be a zero of f such that $Df(\zeta)$ has full rank. Then for all $z \in K^{n+1}$ with $\|z - \zeta\| \, \mu(f, \zeta) < 1$ we have $\mu(f, \zeta) = \mu(f, z)$.*

Proof. As above from lemma 4.2.9 and corollary 2.3.4 we deduce that $\|z\| = 1$. Let T_z be an orthogonal complement of Kz. Recall from theorem 4.2.4 that $\gamma(f, \zeta) \leq \mu(f, \zeta)$. Proposition 4.1.12 tells us that $Df(z)|_{T_z}$ is invertible and

$$\left\| Df(z)|_{T_z}^{-1} Df(\zeta)|_{T_\zeta} \right\| = \left\| Df(\zeta)|_{T_\zeta}^{-1} Df(z)|_{T_z} \right\| = 1.$$

Note that due to $\|z\| = 1$ we have that $\mu(f, z) = \|f\| \left\| Df(z)|_{T_z}^{-1} \right\|$. This yields

$$\begin{aligned} \mu(f, z) &= \|f\| \left\| Df(z)|_{T_z}^{-1} \right\| \\ &= \|f\| \left\| Df(z)|_{T_z}^{-1} Df(\zeta)|_{T_\zeta} Df(\zeta)|_{T_\zeta}^{-1} \right\| \\ &\leq \|f\| \underbrace{\left\| Df(z)|_{T_z}^{-1} Df(\zeta)|_{T_\zeta} \right\|}_{=1} \left\| Df(\zeta)|_{T_\zeta}^{-1} \right\| \\ &= \mu(f, \zeta). \end{aligned}$$

Interchanging the roles of z and ζ yields the opposite inequality. □

4.4 Multivariate Hensel lifting

We can now state a multivariate Hensel lifting.

Theorem 4.4.1. *Let K be a finite algebraic extension of $\mathbb{Q}_p$, $f \in \mathcal{H}_{\underline{d}}$ and $\zeta \in \mathbb{S}(K^{n+1})$ with $f(\zeta) = 0$. Then for all $g \in \mathcal{H}_{\underline{d}}$ with $d_{\mathbb{P}}(Kf, Kg) < \mu(f, \zeta)^{-2}$ any Newton sequence $(\zeta_i)_{i \geq 0}$ with respect to g and starting value $\zeta_0 = \zeta$ exists and converges to a zero $\xi \in \mathbb{S}(K^{n+1})$ of g (at linear rate of convergence).*

Proof. Let π be an uniformizer in K. After scaling we can assume that $\|f\| = \|g\| = 1$. Suppose $d_{\mathbb{P}}(Kf, Kg) < \mu(f, \zeta)^{-2}$. Then by lemma 4.2.9 also $d_{\mathbb{P}}(Kf, Kg) < \mu(f, \zeta)^{-1}$. Proposition 4.3.1 implies that $Dg(\zeta)$ restricted to any orthogonal complement of $K\zeta$ is invertible and

$$1 \leq \mu(f, \zeta) = \mu(g, \zeta) = \left\| Dg(\zeta)|_{T_\zeta}^{-1} \right\|. \tag{4.4.2}$$

Thus there exists a $k \geq 0$ such that $\left\| Dg(\zeta)|_{T_\zeta}^{-1} \right\| = |\pi|^{-k}$. Note that $\left\| Dg(\zeta)|_{T_\zeta}^{-1} \right\|$ is independent of the choice of the orthogonal complement. From the proof of proposition 2.7.4 we know that there exists some $\lambda \in \mathbb{S}(K)$ with $\|f - \lambda g\| = d_{\mathbb{P}}(Kf, Kg)$. Let $m \in \mathbb{Z}$ such that $\|g(\zeta)\| = |\pi|^m$. We apply lemma 4.2.2 and get

$$|\pi|^m = \|g(\zeta)\| = \|\lambda g(\zeta)\| = \|(f - \lambda g)(\zeta)\| \leq \|f - \lambda g\| < \mu(f, \zeta)^{-2} = |\pi|^{2k}.$$

Hence $m > 2k$. Let T_ζ be an orthogonal complement of $K\zeta$ to use for the Newton operator. Set $\hat{\zeta} = N_g(\zeta, T_\zeta)$. We can estimate

$$\left\| \hat{\zeta} - \zeta \right\| = \left\| Dg(\zeta)|_{T_\zeta}^{-1} g(\zeta) \right\| \leq \left\| Dg(\zeta)|_{T_\zeta}^{-1} \right\| \|g(\zeta)\| = |\pi|^{-k+m} < 1. \tag{4.4.3}$$

This shows $\left\| \hat{\zeta} \right\| = 1$. We use the Taylor expansion of g around ζ and obtain

$$\begin{aligned} g(\hat{\zeta}) &= \sum_{i=0}^{\infty} \frac{D^i g(\zeta)}{i!} (\hat{\zeta} - \zeta)^i \\ &= g(\zeta) - Dg(\zeta) Dg(\zeta)|_{T_\zeta}^{-1} g(\zeta) + \sum_{i=2}^{\infty} \frac{D^i g(\zeta)}{i!} (\hat{\zeta} - \zeta)^i \\ &= \sum_{i=2}^{\infty} \frac{D^i g(\zeta)}{i!} (\hat{\zeta} - \zeta)^i. \end{aligned} \tag{4.4.4}$$

This yields $\left\| g(\hat{\zeta}) \right\| \leq \max_{i \geq 2} \left\| \frac{D^i g(\zeta)}{i!} \right\| \left\| (\hat{\zeta} - \zeta) \right\|^i$. By lemma 4.2.7 we know that for any i: $\left\| \frac{D^i g(\zeta)}{i!} \right\| \leq \|g\| = 1$. Taking norms in (4.4.4) and using (4.4.3) gives:

$$\left\| g(\hat{\zeta}) \right\| \leq \left\| \hat{\zeta} - \zeta \right\|^2 = |\pi|^{-2k+2m} = |\pi|^m \cdot |\pi|^{-2k+m} < |\pi|^m,$$

from which it follows that $\left\| g(\hat{\zeta}) \right\| \leq |\pi|^{m+1}$. Furthermore, as $m > 2k$, we have that

$$\left\| \zeta - \hat{\zeta} \right\| \mu(f, \zeta) \leq |\pi|^{-k+m} |\pi|^{-k} < 1. \tag{4.4.5}$$

From equation (4.4.2) we know that $\mu(f, \zeta) = \mu(g, \zeta)$. Using equation 4.4.5 together with proposition 4.3.2 we can deduce that $\mu(f, \zeta) = \mu(f, \hat{\zeta})$. In particular, this yields $d_{\mathbb{P}}(Kf, Kg)\mu(f, \hat{\zeta}) \leq d_{\mathbb{P}}(Kf, Kg)\mu(f, \zeta)^2 < 1$. Using proposition 4.3.1 it follows that $\mu(f, \hat{\zeta}) = \mu(g, \hat{\zeta})$. Altogether we have shown that $\mu(g, \zeta) = \mu(g, \hat{\zeta})$.

This implies that for any orthogonal complement $T_{\hat{\zeta}}$ of $K\hat{\zeta}$ $Dg(\hat{\zeta})_{T_{\hat{\zeta}}}$ is invertible and

$$\left\| Dg(\zeta)^{-1}_{T_\zeta} \right\| = \left\| Dg(\hat{\zeta})^{-1}_{T_{\hat{\zeta}}} \right\| = |\pi|^{-m}.$$

We may therefore apply all steps above to $\hat{\zeta}$ and conclude that

1. Any Newton sequence with respect to g and starting value ζ exists.

2. For any i: $\|\zeta_i - \zeta_{i-1}\| \leq |\pi|^{-k+m+i}$.

3. For the i-th iterate ζ_i of a Newton sequence: $\|g(\zeta_i)\| \leq |\pi|^{m+i}$.

This shows that $(\zeta_i)_{i\geq 0}$ is a Cauchy sequence and converges to a zero ξ of g. Since all iterates have norm $= 1$ also $\|\xi\| = 1$. □

Corollary 4.4.6 (affine version of theorem 4.4.1)**.** *Let $f = (f_1, \ldots, f_n)$ be a system of n polynomials over $\mathbb{Q}_p$ in n variables. Suppose we have zero $\zeta \in \mathcal{O}_K$ of f, where K is a finite algebraic extension of $\mathbb{Q}_p$. Denote by $\tilde{f}$ the homogenized f with additional variable X_{n+1}. Clearly $\zeta' = (\zeta, 1)$ is a zero of $\tilde{f}$. Then for all g with $d_{\mathbb{P}}(K\tilde{f}, K\tilde{g}) < \mu(\tilde{f}, \zeta')^{-2}$ the (affine) newton sequence with respect to g and starting value ζ is defined and converges to a zero of g.*

Proof. Both $\tilde{f}$ and $\tilde{g}$ are in the scope of theorem 4.4.1. Therefore any Newton sequence with respect to g and starting value $(\zeta, 1)$ exists and converges to a zero $\xi \in \mathbb{S}(K^{n+1})$ of g. Let us choose one special Newton sequence: As $\zeta \in \mathcal{O}_K$ we have $\|\zeta\| \leq 1$ and $T := \operatorname{Span}\{e_1, \ldots, e_n\}$ is an orthogonal complement of $K\zeta'$. One checks that

$$D\tilde{g}(\zeta')|_T = Dg(\zeta) \text{ and } D\tilde{g}(\zeta')|_T^{-1} g(\zeta) = (Dg(\zeta)^{-1} g(\zeta), 0).$$

Let $\hat{\zeta} = N_g(\zeta', T)$. Then by the preceding equation:

$$\hat{\zeta} = (\zeta - Dg(\zeta)^{-1} g(\zeta), 1).$$

From the proof of theorem 4.4.1 we also know that $\left\| \hat{\zeta} - \zeta' \right\| < 1$. In particular $\left\| \hat{\zeta} \right\| = 1$. This shows that T is an orthogonal complement for $K\hat{\zeta}$, as well. It follows that in every iteration step, T can be chosen to be the orthogonal complement. The Newton sequence that is constructed this way displays the Newton sequence with respect to g and starting value ζ within the first n variables. Any iterate and the limit ξ then have $X_{n+1} = 1$. If $\xi = (\xi_1, \ldots, \xi_n, \xi_{n+1})$, we have $g(\xi_1, \ldots, \xi_n) = 0$. □

Remark 4.4.7. The advantage of this method when compared to the projective γ-theorem (theorem 4.1.15) is that we don't need to know the zero ξ of the system g to find a starting value for the newton algorithm. But it comes at some cost: The rate of convergence has shrunk from quadratic to linear as it was in the Hensel lifting. One can see this very well in the the examples following. Furthermore, we don't need to solve $g(x) = 0 \mod p$, which one needs to apply Kuhlmann's Hensel lifting. The starting value for our Hensel lifting is a zero of the system f that is close to g. We have replaced finding points close to ξ by finding systems close to g of which we know a solution.

Example 4.4.8. We take the system of example 4.1.17: Let $p = 3$, $K = \mathbb{Q}_p$ and $f : K^2 \to K$ with

$$f(X, Y) = X^2 - Y^2.$$

$\zeta = (1, 1)$ is a zero for f. We again set $T = \operatorname{Span}\{e_1\}$. Then $\mu(f, \zeta) = \mu(f, \zeta, T) = \|f\| \left\| Df(\zeta)|_T^{-1} \|\zeta\|^{2-1} \right\|$. One checks that $\|f\| = 1$. This yields

$$\mu(f, \zeta) = \left\| Df(\zeta)|_T^{-1} \right\| = \left\| \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix} \right\| = 1.$$

Let $g(X, Y) = X^2 + 2Y^2 - 6XY$. Then

$$d_{\mathbb{P}}(Kf, Kg) \leq \frac{\|f - g\|}{\|f\|} = \left\| -3 \cdot \binom{2}{2, 0} \cdot Y^2 - 3 \cdot \binom{2}{1, 1} \cdot XY \right\| = 3^{-1}.$$

In particular $d_{\mathbb{P}}(Kf, Kg)\mu(f, \zeta)^2 < 1$. We are in the scope of theorem 4.4.1. First we compute $\mu(g, \zeta)$: Using $Dg(\zeta) = \begin{pmatrix} -4 & -2 \end{pmatrix}$, $Dg(\zeta)|_T^{-1} = \left(\lambda \mapsto \frac{-\lambda}{4} e_1\right)$ and $\|g\| = 1$ gives

$$\mu(g, \zeta) = \mu(g, \zeta, T) = \left\| Dg(\zeta)|_T^{-1} \right\| = \left\| \begin{pmatrix} \frac{-1}{4} \\ 0 \end{pmatrix} \right\| = 1 = \mu(f, \zeta).$$

We give the first steps of the Newton sequence with respect to g and starting value ζ (displayed as fractions).

$$z_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad T_0 = \operatorname{Span}\{e_1\} \quad |g(z_0)| = |-3| = 3^{-1}$$

$$z_1 = \begin{pmatrix} \frac{1}{4} \\ 1 \end{pmatrix} \qquad T_1 = \operatorname{Span}\{e_1\} \quad |g(z_1)| = \left|\tfrac{9}{16}\right| = 3^{-2}$$

$$z_2 = \begin{pmatrix} \frac{25}{64} \\ 1 \end{pmatrix} \qquad T_2 = \operatorname{Span}\{e_1\} \quad |g(z_2)| = \left|\tfrac{-783}{4096}\right| = 3^{-3}$$

$$z_3 = \begin{pmatrix} \frac{5617}{16384} \\ 1 \end{pmatrix} \quad T_3 = \operatorname{Span}\{e_1\} \quad |g(z_3)| = \left|\tfrac{276070400}{8417}\right| = 3^{-4}.$$

(All calculations were made with Maxima 5.28.0.) This also is a good example for corollary 4.4.6. In every step we have used $\operatorname{Span}\{e_1\}$ as the orthogonal complement. All calculations take place in the first entry while the second stays the same. In fact, the sequence can be regarded as the Newton sequence with respect to $g(X, 1) = X^2 - 6X + 2$.

Example 4.4.9. Let $p = 2$, $K = \mathbb{Q}_p$ and $f : K^2 \to K$ with

$$f(X,Y,Z) = \begin{pmatrix} X^2 + XY - Z^2 \\ X^2Z + XYZ + XZ^2 \end{pmatrix}.$$

$\zeta = (1, 0, -1)$ is a zero for f, $\|\zeta\| = 1$ and $T = \operatorname{Span}\{e_2, e_3\}$ is an orthogonal complement of $K\zeta$. We compute

$$\|f\| = \left\| \begin{pmatrix} \binom{2}{2,0,0} \cdot X^2 + \frac{1}{2} \cdot \binom{2}{1,1,0} \cdot XY - \binom{2}{0,0,2} \cdot Z^2 \\ \frac{1}{3} \cdot \binom{3}{2,0,1} \cdot X^2Z + \frac{1}{6} \cdot \binom{3}{1,1,1} \cdot XYZ + \frac{1}{3} \cdot \binom{3}{1,0,2} \cdot XZ^2 \end{pmatrix} \right\| = 2.$$

Further, $\|Df(\zeta)\| = \begin{pmatrix} 2 & 1 & 2 \\ 1 & -1 & -1 \end{pmatrix}$ and $Df(\zeta)|_T^{-1} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = (-\lambda - 2\mu)e_2 - (\lambda + \mu)e_3$. This yields

$$\mu(f, \zeta) = \mu(f, \zeta, T) = \|f\| \left\|Df(\zeta)|_T^{-1}\right\| = 2 \cdot \max \|-e_2 - e_3\|, \|-2e_2 - e_3\| = 2.$$

Let

$$g(X,Y,Z) = \begin{pmatrix} -3X^2 + XY - 5Z^2 \\ -3X^2Z - 7XYZ + XZ^2 \end{pmatrix}.$$

Then

$$\begin{aligned} d_{\mathbb{P}}(Kf, Kg) &\leq \frac{\|f-g\|}{\|f\|} \\ &= \frac{1}{2} \cdot \left\| \begin{pmatrix} 4 \cdot \binom{2}{2,0,0} \cdot X^2 + 4\binom{2}{0,0,2} \cdot Z^2 \\ \frac{4}{3} \cdot \binom{3}{2,0,1} \cdot X^2 Z + \frac{4}{3} \cdot \binom{3}{1,1,1} \cdot XYZ \end{pmatrix} \right\| = \frac{1}{8}. \end{aligned}$$

Putting these together we see that $d_{\mathbb{P}}(Kf, Kg)\mu(f,\zeta)^2 \leq \frac{1}{8} \cdot 4 < 1$. We may therefore apply theorem 4.4.1. Again, first we compute $\mu(g,\zeta)$: One checks that $\|g\| = 2$ and $Dg(\zeta)|_T^{-1} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \frac{\lambda+2\mu}{25} e_2 + \frac{-7\lambda-\mu}{75} e_3$. This gives

$$\begin{aligned} \mu(g,\zeta) = \mu(g,\zeta,T) &= \|g\| \left\| Dg(\zeta)|_T^{-1} \right\| \\ &= 2 \cdot \max \left\| \frac{1}{25} e_2 - \frac{7}{75} e_3 \right\|, \left\| \frac{2}{25} e_2 - \frac{1}{75} e_3 \right\| \\ &= 2 \\ &= \mu(f,\zeta). \end{aligned}$$

We give the first steps of the Newton sequence with respect to g and starting value ζ (displayed in the decimal system).

$$z_0 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \qquad T_0 = \mathrm{Span}\{e_2, e_3\} \quad \|g(z_0)\| = \left\| \begin{pmatrix} -8 \\ 4 \end{pmatrix} \right\| = 2^{-2}$$

$$z_1 = \begin{pmatrix} 1 \\ 0 \\ \frac{-1}{5} \end{pmatrix} \qquad T_1 = \mathrm{Span}\{e_1, e_2\} \quad \|g(z_1)\| = \left\| \begin{pmatrix} \frac{-16}{5} \\ \frac{16}{25} \end{pmatrix} \right\| = 2^{-4}$$

$$z_2 = \begin{pmatrix} \frac{113}{241} \\ \frac{16}{1205} \\ \frac{-1}{5} \end{pmatrix} \qquad T_2 = \mathrm{Span}\{e_2, e_3\} \quad \|g(z_2)\| = \left\| \begin{pmatrix} \frac{-247808}{290405} \\ \frac{231424}{1452025} \end{pmatrix} \right\| = 2^{-11}$$

$$z_3 = \begin{pmatrix} \frac{113}{241} \\ \frac{346481392}{1291508155} \\ \frac{894289}{5358955} \end{pmatrix} \quad T_2 = \mathrm{Span}\{e_2, e_3\} \quad \|g(z_2)\| = \left\| \begin{pmatrix} \frac{-773094113280}{1148735947681} \\ \frac{-81379958798155776}{333598662886300805} \end{pmatrix} \right\| = 2^{-28}.$$

(All calculations were made with Maxima 5.28.0.)

Bibliography

[1] Baker, A. (2002): *An Introduction to p-adic Numbers and p-adic Analysis.* Department of mathematics, university of Glasgow. Glasgow.

[2] Blum, L., Cucker, F., Shub, M., Smale, S. (1998): *Complexity and real computation.* Springer. NY et al.

[3] Blum, L., Shub, M., Smale, S. (1989): *On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines.* In: Bull. Amer. Math. Soc. 21, 1-46.

[4] Bürgisser, P., Cucker, F. (2013): *Condition.* Grundlehren der mathematischen Wissenschaften, No. 349. Springer.

[5] Cox, D., Little, J., O'Shea, D. (2007): *Ideals, Varieties and Algorithms.* Springer. NY.

[6] Freankel, A.S., Yesha, Y. (1979): *Complexity of problems in games, graphs and algebraic equations.* In: Discrete Applied Mathematics 1, 15-30. North-Holland Publishing Company.

[7] Kuhlmann, F. (2011): *Maps on ultrametric spaces, Hensels Lemma, and differential equations over valued fields.* In: Comm. in Alg. 39, 1730-1776.

[8] Maller, M., Whitehead, J. (1997): *Computational Complexity over the p-adic numbers.* In: Journal of complexity 13, 195-207. Elsevier.

[9] Robert, A.M. (2000): *A course in p-adic analysis.* Springer. NY et al.

[10] Schikhof, W.H. (1984): *Ultrametric calculus: An introduction to p-adic analysis.* Cambridge University Press. Cambridge et al.

[11] Schneider, P. (2011): *p-adic Lie Groups.* Springer. Berlin et al.

Eidesstattliche Erklärung zur Masterarbeit

Ich versichere, die Masterarbeit selbstständig und lediglich unter Benutzung der angegebenen Quellen und Hilfsmittel verfasst zu haben.
Ich erkläre weiterhin, dass die vorliegende Arbeit noch nicht im Rahmen eines anderen Prüfungsverfahrens eingereicht wurde.

Göttingen, den 26.09.2013