Chapter 5

Polynomials over valued fields

5.1 The value of a polynomial

Suppose we have a polynomial ring in several variables and a valuation on its quotient field. Then we would like to know the value of any given polynomial or rational function under this valuation. This is in general not an easy problem, and the answer depends on how the valuation is defined on the rational function field. This will be studied at several points later in this book. Here, we will present a particularly easy and therefore important special case.

Theorem 5.1 *Let $(L|K,P)$ be an extension of valued fields. Take a set*

$$\mathcal{T} \;=\; \{x_i, y_j \mid i\in I\,,\, j\in J\}$$

of elements in L such that the values vx_i, $i\in I$, are rationally independent over vK, $vy_j=0$ for all $j\in J$, and the residues y_jv, $j\in J$, are algebraically independent over Kv.

If we write

$$f \;=\; \sum_k c_k \prod_{i\in I} x_i^{\mu_{k,i}} \prod_{j\in J} y_j^{\nu_{k,j}} \in K[x_i,y_j\mid i\in I, j\in J] \tag{5.1}$$

in such a way that for every $k\neq \ell$ there is some i such that $\mu_{k,i}\neq\mu_{\ell,i}$ or some j such that $\nu_{k,j}\neq\nu_{\ell,j}$, then

$$vf \;=\; \min_k \left(v(c_k \prod_{i\in I} x_i^{\mu_{k,i}} \prod_{j\in J} y_j^{\nu_{k,j}})\right) \;=\; \min_k \left(vc_k \,+\, \sum_{i\in I}\mu_{k,i} vx_i\right)\,.$$

That is, the value of the polynomial f is equal to the least of the values of its monomials. In particular, this implies that the elements x_i,y_j, $i\in I$, $j\in J$, are algebraically independent over K, and that

$$\begin{aligned} vK(\mathcal{T}) &= vK\oplus\bigoplus_{i\in I}\mathbb{Z}vx_i\\ K(\mathcal{T})v &= Kv\,(y_jv\mid j\in J)\,. \end{aligned}$$

Further, v is uniquely determined on $K(x_i,y_j\mid i\in I, j\in J)$ by its restriction to K and the values vx_i, and the residue map is uniquely determined on $K(x_i,y_j\mid i\in I,j\in J)$ by these data together with the residues y_jv.

Proof: Every element of $K(\mathcal{T})$ is the quotient of two polynomials in finitely many elements $x_1,\ldots,x_r,y_1,\ldots,y_s\in\mathcal{T}$. We take a polynomial f in these variables and write it in the form (5.1) where we replace I by $\{1,\ldots,r\}$ and J by $\{1,\ldots,s\}$ in such a way that the subsequent conditions are satisfied. These conditions together with the rational independence of the values vx_i over vK imply that

$$\begin{aligned} v(c_k\prod_{1\le i\le r} x_i^{\mu_{k,i}}\prod_{1\le j\le s} y_j^{\nu_{k,j}}) &= vc_k+\sum_{1\le i\le r}\mu_{k,i}vx_i\\ &= vc_\ell+\sum_{1\le i\le r}\mu_{\ell,i}vx_i \;=\; v(c_\ell\prod_{1\le i\le r} x_i^{\mu_{\ell,i}}\prod_{1\le j\le s} y_j^{\nu_{\ell,j}}) \end{aligned}$$

if and only if $vc_k=vc_\ell$ and $\mu_{k,i}=\mu_{\ell,i}$ for $1\le i\le r$. So there is exactly one tuple $(\mu_{k,1},\ldots,\mu_{k,r})$ for which this value is minimal; without loss of generality we may assume that this tuple and the minimal value is assumed exactly for $k=1,\ldots,t$; we have that $vc_1=\ldots=vc_t$. The corresponding monomials in f have minimal value and the other have higher value, so we will obtain that vf is equal to this minimal value once we prove the first equation in

$$v\left(\sum_{k=1}^t c_k\prod_i x_i^{\mu_{k,i}}\prod_j y_j^{\nu_{k,j}}\right) \;=\; v(c_1\prod_i x_i^{\mu_{1,i}}\prod_j y_j^{\nu_{1,j}})\;;$$

the second equation holds because $vy_j=0$ for all j. Since we can write

$$\sum_{k=1}^t c_k\prod_i x_i^{\mu_{k,i}}\prod_j y_j^{\nu_{k,j}} \;=\; c_1\prod_i x_i^{\mu_{1,i}}\cdot\sum_{k=1}^t\frac{c_k}{c_1}\prod_j y_j^{\nu_{k,j}}\;,$$

we just have to show that

$$v\left(\sum_{k=1}^t\frac{c_k}{c_1}\prod_j y_j^{\nu_{k,j}}\right) \;=\; 0\;. \tag{5.2}$$

As $v\frac{c_k}{c_1}=0$ for $1\le k\le t$, we find that

$$\left(\sum_{k=1}^t\frac{c_k}{c_1}\prod_j y_j^{\nu_{k,j}}\right)v \;=\; \sum_{k=1}^t\frac{c_k}{c_1}v\,\prod_j (y_jv)^{\nu_{k,j}}\;\neq\;0$$

because the residues y_jv are algebraically independent over Kv and the coefficients $\frac{c_k}{c_1}v$ of the linear combination are non-zero. This proves that (5.2) holds.

Further, it follows that

$$vf \;=\; vc_1+\sum_{i=1}^r\mu_{1,i}vx_i\;\in\;vK\,\oplus\,\bigoplus_{i\in I}\mathbb{Z}vx_i\;.$$

Since the value of a quotient f/g is $vf-vg$, we see that the values of all non-zero elements in $K(\mathcal{T})$ lie in $vK\oplus\bigoplus_{i\in I}\mathbb{Z}vx_i$. Conversely, by the choice of quotients of suitable polynomials, one shows that every value in $vK\oplus\bigoplus_{i\in I}\mathbb{Z}vx_i$ appears as a value of some element in $K(\mathcal{T})$. Observe that the prescription of the values vx_i determines which monomials in a given polynomial are the ones of least value, and that their value is computed by the above formula only by use of the restriction of v to K and the values vx_i. Hence, v is uniquely determined on $K(\mathcal{T})$ by these data, together with the assumption on the residues of the y_j.

Now assume that the value of some quotient f/g is zero; we want to determine its residue. Since $vf = vg$, the summands of minimal value in f and g all contain x_i to the same power ν_i, for every $i \in \{1,\ldots,r\}$. Dividing numerator and denominator by $c\prod_{i=1}^r x_i^{\nu_i}$ with a suitable constant $c \in K$, we obtain a quotient where numerator and denominator have value zero and where the summands of minimal value are polynomials in $K[y_1,\ldots,y_s]$. Since the residues of numerator and denominator are only depending on the residues of these summands of minimal value 0, our assertion now follows from the fact that the residue of every polynomial in $\mathcal{O}[y_1,\ldots,y_s]$ lies in $Kv(y_jv \mid j \in J)$. The prescribed values vx_i determine uniquely the summands of least value 0, and the prescribed residues y_jv determine uniquely their residues. Hence, the residue map is uniquely determined on $K(\mathcal{T})$ by these data. $\square$

Let us state the following almost trivial observation:

Lemma 5.2 *Let (K,v) be a valued field and $L = K(t_i \mid i \in I)$ an extension field of K. Assume that w is a map from L into an ordered abelian group Γ which extends v. Then w is a valuation on L if and only if it satisfies axioms* (V0), (VT), (VH) *on the ring $K[t_i \mid i \in I]$, and $w(f/g) = wf - wg$ for every $f,g \in K[t_i \mid i \in I]$.*

Proof: Suppose that w satisfies the above conditions. We show that it then satisfies the triangle inequality (VT) on all of L; the other details are left to the reader. Let $f_1,f_2,g_1,g_2 \in K[t_i \mid i \in I]$. Then $v(f_1/g_1 - f_2/g_2) = v(f_1g_2 - f_2g_1) - vg_1g_2 \geq \min\{vf_1g_2, vf_2g_1\} - vg_1g_2 = \min\{v(f_1/g_1), v(f_2/g_2)\}$. $\square$

Theorem 5.3 *Let (K,v) be a valued field, α_i, $i \in I$, elements in some ordered abelian group extension Γ of vK which are rationally independent over vK. Take a set $\mathcal{T} = \{x_i, y_j \mid i \in I, j \in J\}$ of elements algebraically independent over K. For $f \in K[\mathcal{T}]$ given as in (5.1), define*

$$vf \;:=\; \min_k v\left(c_k \prod_{i\in I} x_i^{\mu_{k,i}} \prod_{j\in J} y_j^{\nu_{k,j}}\right) \;=\; \min_k \left(vc_k + \sum_{i\in I}\mu_{k,i}\alpha_i\right) \tag{5.3}$$

and extend v to the rational function field $K(\mathcal{T})$ by setting $v(f/g) = vf - vg$. Then v is a valuation on $K(\mathcal{T})$, and it is the unique valuation extending v from K to $K(\mathcal{T})$ such that $vx_i = \alpha_i$, $i \in I$, $vy_j = 0$, $j \in J$, and that the residues y_jv, $j \in J$, are algebraically independent over Kv.

Proof: In view of the foregoing lemma, it suffices to show that v defined on the polynomial ring $K[\mathcal{T}]$ by (5.3) satisfies axioms (V0), (VT), (VH). Its definition on $K(\mathcal{T})$ is then canonically given by the rule $v(f/g) = vf - vg$.

The right hand side in (5.3) is ∞ if and only if all $c_k = 0$, that is, if and only if $f = 0$. This proves (V0).

For the proof of the triangle inequality (VT), we observe that each monomial in the sum of two polynomials f and g is the sum of two monomials $c\prod_{i\in I} x_i^{\nu_i}\prod_{j\in J} y_j^{\nu_j}$ and $c'\prod_{i\in I} x_i^{\nu_i}\prod_{j\in J} y_j^{\nu_j}$ in f and g, respectively. We have that

$$\begin{aligned} v(c+c')\prod_{i\in I} x_i^{\nu_i}\prod_{j\in J} y_j^{\nu_j} &= v(c+c')+\sum_{i\in I}\nu_i\alpha_i \;\geq\; \min\{vc,vc'\}+\sum_{i\in I}\nu_i\alpha_i \\ &= \min\{vc+\sum_{i\in I}\nu_i\alpha_i, vc'+\sum_{i\in I}\nu_i'\alpha_i\} \\ &\geq \min\{vf,vg\};. \end{aligned}$$

So we see that the value of every monomial in $f+g$ is $\geq \min\{vf,vg\}$, which by definition yields that $v(f+g)\geq\min\{vf,vg\}$. This shows that v satisfies (VT) on $K[x_i\mid i\in I]$.

We wish to show that v satisfies (VH) on $K[x_i\mid i\in I]$. If we multiply two polynomials f and g, then we obtain the monomials of minimal value in the product fg by multiplying out the monomials of minimal value in f and in g. As in the proof of Theorem 5.1 we can write the monomials of minimal value in f as $c_k\prod_{i=1}^r x_i^{\mu_{1,i}}\prod_{j=1}^s y_j^{\nu_{k,j}}$, $1\leq k\leq t$, and the monomials of minimal value in g as $c_k'\prod_{i=1}^r x_i^{\mu_{1,i}'}\prod_{j=1}^s y_j^{\nu_{k,j}'}$, $1\leq k\leq t'$. Without loss of generality, we may assume that among all the tuples $(\nu_{k,1},\ldots,\nu_{k,s})$ the one with $k=1$ is the lexicographically smallest, and we may assume the same for the tuples $(\nu_{k,1}',\ldots,\nu_{k,s}')$. Then among all the possible products of pairs of the above monomials,

$$c_1c_1'\prod_{i=1}^r x_i^{\mu_{1,i}+\mu_{1,i}'}\prod_{j=1}^s y_j^{\nu_{1,j}+\nu_{1,j}'}$$

is the unique one with lexicographically minimal tuple of exponents for the y_j. Consequently, the coefficient for $\prod_{i=1}^r x_i^{\mu_{1,i}+\mu_{1,i}'}\prod_{j=1}^s y_j^{\nu_{1,j}+\nu_{1,j}'}$ in the product fg is just $c_1c_1'\neq 0$. Hence by our definition of v, we find that

$$v(fg)\;\leq\; vc_1c_1'+\sum_i(\mu_{1,i}+\mu_{1,i}')\alpha_i \;=\; vc_1+\sum_i\mu_{1,i}\alpha_i+vc_1'+\sum_{i\in I}\mu_{1,i}'\alpha_i \;=\; vf+vg\;.$$

On the other hand, if $d\prod_{i=1}^r x_i^{\kappa_i}\prod_{j=1}^s y_j^{\lambda_j}$ is any monomial appearing in fg, then the coefficient d is a linear combination of products cc' where c,c' are coefficients of monomials $c\prod_{i=1}^r x_i^{\mu_i}\prod_{j=1}^s y_j^{\nu_j}$ and $c'\prod_{i=1}^r x_i^{\mu_i'}\prod_{j=1}^s y_j^{\nu_j'}$ in f and g, respectively, for which $\mu_i+\mu_i'=\kappa_i$ and $\nu_i+\nu_i'=\lambda_i$. Then $vc\geq vf-\sum_{i=1}^r\mu_i\alpha_i$ and $vc'\geq vf-\sum_{i=1}^r\mu_i'\alpha_i$. Thus, $vcc'= vf+vg-\sum_{i=1}^r(\mu_i+\mu_i')\alpha_i = vf+vg-\sum_{i=1}^r\kappa_i\alpha_i$, and by the ultrametric triangle law on (K,v), $vd\geq vf+vg-\sum_{i=1}^r\kappa_i\alpha_i$, showing that $vd\prod_{i=1}^r x_i^{\kappa_i}\prod_{j=1}^s y_j^{\lambda_j}\geq vf+vg$. Together with our previous result, this proves that $v(fg)=vf+vg$.

The uniqueness of v on $K(x_i\mid i\in I)$ was already stated in Theorem 5.1. □

An important special case of our definition given in (**??**) appears when $I=\emptyset$. Then we have

$$v\left(\sum_k c_k x_i^{\mu_{k,i}}\prod_{j\in J} y_j^{\nu_{k,j}})\right) \;:=\; \min_k vc_k\;. \tag{5.4}$$

The induced extension of v from K to $K(y_j\mid j\in J)$ is called the **Gauß valuation** or **functional valuation**. It can be described by saying that the value of a polynomial is the minimum of the values of its coefficients. In particular, $vf \geq 0$ if and only if all coefficients are integral. That is,

$$K[X_1, \ldots, X_n] \cap \mathcal{O}_{K(X_1,\ldots,X_n)} = \mathcal{O}_K[X_1, \ldots, X_n] .$$

Note that a monic polynomial always has value ≤ 0 and that a monic polynomial with integral coefficients has value 0. A polynomial $f \in K[X_1, \ldots, X_n]$ is called **primitive** if $vf = 0$, that is, if the coefficients of minimal value are units in $\mathcal{O}_{\mathbf{K}}$. The latter is the classical definition of "primitive". With the help of the Gauß valuation, it follows at once from Theorem 5.3 that in the present case, *the product of two primitive polynomials is again primitive.* We also get the so-called **Gauß' Lemma** for granted:

Lemma 5.4 *Let $f, g, h \in K[X_1, \ldots, X_n]$ such that $f = gh$.*

a) **(Gauß' Lemma, first form)** *If f and g are primitive, then so is h.*

b) **(Gauß' Lemma, second form)** *If f and g are monic and f has integral coefficients, then also h is monic, and g and h have integral coefficients.*

Lemma 5.5 *Let (K, v) be a valued field and ζ_j , $j \in J$, algebraically independent over $\overline{K}$. Choose elements y_j , $j \in J$ which are algebraically independent over K. Then v and its associate residue map extend in a unique way from K to the rational function field $K(y_j \mid j \in J)$ such that $\overline{y}_j = \zeta_j$ for every $j \in J$. For this extension,*

$$vK(y_j \mid j \in J) = vK \quad \text{and} \quad \overline{K(y_j \mid j \in J)} = \overline{K}(\zeta_j \mid j \in J) .$$

On $K[y_j \mid j \in J]$, the valuation v is given by

$$v\left(\sum_{\underline{\nu}} c_{\underline{\nu}}\, y_1^{\nu_1} \cdot \ldots \cdot y_n^{\nu_n}\right) = \min_{\underline{\nu}} vc_{\underline{\nu}} ,$$

and on $\mathcal{O}_{\mathbf{K}}[y_j \mid j \in J]$, the residue map is given by

$$\overline{\sum_{\underline{\nu}} c_{\underline{\nu}}\, y_1^{\nu_1} \cdot \ldots \cdot y_n^{\nu_n}} = \sum_{\underline{\nu}} \overline{c_{\underline{\nu}}}\, \overline{y}_1^{\nu_1} \cdot \ldots \cdot \overline{y}_n^{\nu_n} .$$

Proof: Let P be the place associated with v on K. It is a homorphism from $\mathcal{O}_{\mathbf{K}}$ onto $\overline{K}$. Since the elements y_j , $j \in J$, are algebraically independent over K and the elements ζ_j , $j \in J$, are algebraically independent over $\overline{K}$, we can extend P to a homomorphism from $\mathcal{O}_{\mathbf{K}}[y_j \mid j \in J]$ onto $\overline{K}[\zeta_j \mid j \in J]$ by setting $y_jP = \zeta_j$. Via the rule $(f/g)P = fP/gP$, this homomorphism extends to a place of the quotient field $K(y_j \mid j \in J)$ of $\mathcal{O}_{\mathbf{K}}[y_j \mid j \in J]$ (cf. also Theorem 4.10). We have constructed an extension of P from K to $K(y_j \mid j \in J)$, hence the valuation associated with P on the latter field is also an extension of v.

The formula for the value of polynomials is just a special case of (6.5) since $vy_j = 0$ for all $j \in J$. The uniqueness of v on $K(y_j \mid j \in J)$ was already stated in Lemma 6.35. $\square$

5.2 The value of coefficients and roots

Let $\mathbf{K} = (K, v)$ be an arbitrary valued field and

$$f = c_n X^n + c_{n-1} X^{n-1} + \ldots + c_1 X + c_0 \in K[X] \tag{5.5}$$

an arbitrary polynomial. In this section, we will consider the valuation theoretical relation between the coefficients and the roots of f. Since the roots may not lie in K, we extend v to $\tilde{K}$ and call this extension again v. To begin with, we treat the case of a monic polynomial f, that is, $c_n = 1$. Write

$$f = \prod_{i=1}^{n} (X - a_i) \quad \text{with} \quad a_i \in \tilde{K} \; .$$

Recall that the coefficients c_i are, up to the sign, elementary symmetric polynomials in the roots a_i. In fact, if $0 < j \leq n$ then $(-1)^j c_{n-j}$ is equal to the j-th elementary symmetric polynomial in $a_1, \ldots, a_n$:

$$s_j(a_1, \ldots, a_n) := \sum_{1 \leq i_1 < \ldots < i_j \leq n} a_{i_1} \cdot \ldots \cdot a_{i_j} \; .$$

In particular, we observe that

$$vc_{n-1} = v(a_1 + \ldots + a_n) \geq \min_i va_i \quad \text{and} \quad vc_0 = v(a_1 \cdot \ldots \cdot a_n) = va_1 + \ldots + va_n \; .$$

Further, if all roots a_i lie in $\mathcal{O}_{\tilde{\mathbf{K}}}$, then all coefficients c_i lie in $\mathcal{O}_{\mathbf{K}}$ (observe that the latter does not depend on the chosen extension of v from K to $\tilde{K}$). More precisely, if $va_i \geq \alpha$ for all i, then $vc_{n-j} \geq j\alpha$ for $0 < j \leq n$.

Without loss of generality, we may assume that the roots a_i are numbered such that $va_1 \leq va_2 \leq \ldots \leq va_n$. Suppose that $a_1, \ldots, a_\ell$ are all roots of value < 0 (which means that f has precisely $n - \ell$ integral roots). Then the value of $a_1 \cdot \ldots \cdot a_\ell$ is smaller than the value of any other product of at most ℓ distinct roots a_i. This yields that $a_1 \cdot \ldots \cdot a_\ell$ is the unique summand of minimal value in the ℓ-th elementary symmetric polynomial in $a_1, \ldots, a_n$. Consequently, $vc_{n-\ell} = va_1 \cdot \ldots \cdot a_\ell = va_1 + \ldots + va_\ell$. It also yields that this value is smaller than the value of any summand in the j-th elementary symmetric polynomial, for every $j < \ell$. Consequently, $vc_{n-\ell} < vc_{n-j}$ for $j < \ell$. Moreover, $va_1 \cdot \ldots \cdot a_\ell$ is smaller or equal to the value of any other product of distinct roots a_i. This yields that $vc_{n-\ell} \leq vc_{n-j}$ for $j \geq \ell$. Hence, among all coefficients of smallest value, $vc_{n-\ell}$ is the one with maximal index. Since $c_n = 1$, this also holds if all roots are integral, that is, if $\ell = 0$.

Suppose that precisely $a_1, \ldots, a_m$ are of minimal value among the elements $a_1, \ldots, a_n$. Then $a_1 \cdot \ldots \cdot a_m$ is the unique summand of minimal value in the m-th elementary symmetric polynomial in $a_1, \ldots, a_n$. Consequently, $vc_{n-m} = va_1 \cdot \ldots \cdot a_m = mva_1$, and c_{n-m} is integral if and only if $a_1, \ldots, a_m$ are integral. For every $j \leq m$, every product of j distinct roots a_i has value at least jva_1. On the other hand, for $j > m$ every product of j distinct roots a_i has value $> jva_1$ since it has at least one factor of value $> va_1$. Consequently, $vc_{n-j} \geq jva_1$ for $j \leq m$, and $vc_{n-j} > jva_1$ for $j > m$. Hence, among all coefficients for which $\frac{1}{j} vc_{n-j}$ is minimal, vc_{n-m} is the one with minimal index. Further, we see that $\min_i va_i = va_1 = \min_j \frac{1}{j} vc_{n-j}$.

If $0 < m < n$ is such that $va_m < va_{m+1}$, then as before, $a_1 \cdot \ldots \cdot a_m$ is the unique summand of minimal value in the m-th elementary symmetric polynomial in $a_1, \ldots, a_n$. This yields that $vc_{n-m} = va_1 \cdot \ldots \cdot a_m = va_1 + \ldots + va_m$. If moreover $m < \mu < n$ is such that $va_\mu < va_{\mu+1}$, then $vc_{n-\mu} - vc_{n-m} = va_{m+1} + \ldots + va_\mu$. In particular, if $1 < m < n$ and $va_{m-1} < va_m < va_{m+1}$, then $va_m = vc_{n-m} - vc_{n-m+1} \in vK$. A similar formula holds for the roots a_1 and a_n if they are the unique root of minimal resp. maximal value. Indeed, if $va_1 < va_2$, then $va_1 = vc_{n-1}$, which we can write as $vc_{n-1} - vc_n \in vK$ because of $c_n = 1$. If $va_{n-1} < va_n$, then by what we have shown, $vc_1 = va_1 + \ldots + va_{n-1}$. Since $vc_0 = va_1 + \ldots + va_n$, it follows that $va_n = vc_0 - vc_1 \in vK$.

Now assume that f is not monic. Then we divide by c_n to obtain a monic polynomial which has the same roots as f. We have to replace vc_i by $vc_i - vc_n$, but this does not change the order between the values of the coefficients. Let us summarize what we have proved:

Lemma 5.6 *Let* $\mathbf{K} = (K, v)$ *and* $f \in K[X]$ *a polynomial of the form (5.5) with roots* $a_1, \ldots, a_n$. *Then:*

a) Assume that f *is monic. Then* $f \in \mathcal{O}_{\mathbf{K}}[X]$ *if and only if all roots* a_i *are integral.*

b) Suppose that among all coefficients of smallest value, vc_j *is the one with maximal index. Then* f *has precisely* j *integral roots. If in addition, the roots* $a_1, \ldots, a_{n-j}$ *are the roots which are not integral, then* $vc_j - vc_n = va_1 + \ldots + va_{n-j}$.

c) The values of the roots are bounded from below by the following formula:

$$\min_i va_i = \min_{1 \le j \le n} \frac{vc_{n-j} - vc_n}{j} .$$

d) If $m \in \{1, \ldots, n\}$ *such that* $va_j \neq va_m$ *for all* $j \neq m$, *then* $va_m \in vK$.

Consider a polynomial of the form (5.5). If all coefficients c_i are integral, that is, if $f \in \mathcal{O}_{\mathbf{K}}[X]$, then $\overline{f}$ will denote the polynomial $\overline{c}_n X^n + \ldots + \overline{c}_1 X + \overline{c}_0 \in \overline{K}[X]$ which we obtain from f by replacing its coefficients by their residues. We will also use the notation fv if it is necessary to indicate the valuation. We call $\overline{f}$ the **reduction of** f. Since the residue map is a ring homomorphism, we have $\overline{f(a)} = \overline{c_n a^n + \ldots + c_1 a + c_0} = \overline{c}_n\, \overline{a}^n + \ldots + \overline{c}_1\, \overline{a} + \overline{c}_0 = \overline{f}(\overline{a})$ for $a \in \mathcal{O}_{\mathbf{K}}$. In particular, if a is a root of f then $\overline{a}$ is a root of $\overline{f}$. The converse is certainly not true if the valuation v is non-trivial on K, since then for every $\overline{a} \in \overline{K}$, the set $a + \mathcal{M}_{\mathbf{K}}$ of elements with residue $\overline{a}$ is infinite and thus, not all of these elements can be zeros of f.

Our definition of the reduction of a polynomial is generalized in the obvious way to polynomials in several variables. When one works with the notions of "polynomial ring" and "rational function field", then variables are viewed as transcendental elements (this ambiguity is also used in the model theory of fields, as we will also do in Chapter 20). Here, this observation leads to the question whether the reduction of a polynomial is actually the application of a place of the rational function field, for which the elements of the polynomial ring are integral. This is in fact true, as we will show now.

Consider the polynomial ring $K[X_1, \ldots, X_n]$. Although it is not quite formally correct, we let $X_1, \ldots, X_n$ represent elements which are algebraically independent over K as well as over $\overline{K}$. So we take $y_j = X_j = \zeta_j$ and apply the Lemma 5.3. We obtain a residue map on $K(X_1, \ldots, X_n)$ which satisfies

$$\overline{f(X_1, \ldots, X_n)} = \overline{f}(X_1, \ldots, X_n) \quad \text{for every } f \in \mathcal{O}_{\mathbf{K}}[X_1, \ldots, X_n] .$$

So it coincides with our definition of the reduction of a polynomial. The associated valuation on $K[X_1, \ldots, X_n]$ assigns to f the minimal value of its coefficients. For a polynomial f in one variable, given in the form (5.5), this reads as follows:

$$vf = \min_{1 \le i \le n} vc_i \,.$$

The next lemma tells about the relation of the roots of f and the roots of its reduction $\overline{f}$.

Lemma 5.7 *Let (K, v) be an arbitrary valued field and let v be extended to $\tilde{K}$. Let $f \in \mathcal{O}_{\mathbf{K}}[X]$ such that $\overline{f} \neq 0$. Then the residue map establishes a bijection from the integral roots of f onto the roots of $\overline{f}$ (counted with multiplicities). If all roots of f have value < 0, then $\overline{f} \in \overline{K}$.*

Proof: Let f be given in the form (5.5). If all roots of f have value < 0, then according to part b) of Lemma 5.6, c_0 is the unique coefficient of f of smallest value. Hence $\overline{f} = \overline{c_0} \in \overline{K} \setminus \{0\}$ does not admit any roots.

Now assume that $\overline{f}$ admits a root. Write $f = c_n \prod_{i=1}^{n}(X - a_i)$ with $a_i \in \tilde{K}$. Define $g := \prod_{i \in I}(X - a_i) \in \mathcal{O}_{(\tilde{K},v)}[X]$ where $I = \{i \mid 1 \le i \le n \wedge va_i \ge 0\}$. We have that $f = gh$ with $h \in \mathcal{O}_{(\tilde{K},v)}[X]$ by virtue of Gauß' Lemma. By definition, all roots of h have value < 0, showing that $\overline{h}$ is a constant. It is a nonzero element in $\overline{K}$ since it is the leading coefficient of $\overline{f} \neq 0$. Consequently, with $\overline{c} := \overline{h} \in \overline{K}$,

$$\overline{f} \;=\; \overline{c} \prod_{va_i \ge 0} (X - \overline{a_i})$$

which shows that the residue map establishes the asserted bijection. $\square$

5.3 Newton polygons

5.4 Continuity of roots

We are now going to prove what is called the "continuity of roots". We need some easy observations.

Lemma 5.8 *Let a be a root of the polynomial (5.5). Then $vf - vc_n \le 0$, and*

$$vc_0 - (n-1)(vf - vc_n) \;\ge\; va \;\ge\; \min_{1 \le j \le n} \frac{vc_{n-j} - vc_n}{j} \;\ge\; vf - vc_n \,.$$

Proof: Since $vf - vc_n = c_n^{-1}f$ is the value of a monic polynomial, it is ≤ 0. So if $\frac{1}{j}(vc_{n-j} - vc_n) \ge 0$, then $\frac{1}{j}(vc_{n-j} - vc_n) \ge vf - vc_n$. If $\frac{1}{j}(vc_{n-j} - vc_n) < 0$, then $\frac{1}{j}(vc_{n-j} - vc_n) > vc_{n-j} - vc_n \ge vf - vc_n$ by definition of vf. Now the lower bounds for va follow from part c) of Lemma 5.6. Now let $a_1 = a, a_2, \ldots, a_n$ be all roots of f. By what we have just proved, $va_i \ge vf - vc_n$ for all i. Consequently, $vc_0 = v\prod_{i=1}^{n} a_i = \sum_{i=1}^{n} va_i \ge va + (n-1)(vf - vc_n)$. This gives the upper bound for va. $\square$

Lemma 5.9 *Let a be a root of the monic polynomial $f \in K[X]$. Further, let $g \in K[X]$ be of degree $n = \deg f$. Then*
$$vg(a) \;\geq\; nvf + v(f-g)\;.$$
If $v(f-g) > 0$, then there is a root $b \in \tilde{K}$ of g such that
$$v(a-b) \;\geq\; vf + \frac{v(f-g)}{n}\;. \tag{5.6}$$

Proof: Write $g(X) = \sum_{i=1}^{n} c_i' X^i$. Then $vg(a) = v(g(a) - f(a)) = v\sum_{i=1}^{n}(c_i' - c_i)a^i \geq \min_i(v(c_i' - c_i) + iva) \geq \min_i v(c_i' - c_i) + \min_i iva \geq v(g-f) + \min_{0\leq i\leq n} ivf = v(f-g) + nvf$ because $vf \leq 0$.

Now we write $g(X) = c_n' \prod_{i=1}^{n}(X - b_i)$ with $b_i \in \tilde{K}$. By what we have just proved, $nvf + v(f-g) \leq vg(a) = vc_n' \prod_{i=1}^{n}(a - b_i) = vc_n' + \sum_{i=1}^{n} v(a - b_i)$. If $v(f-g) > 0$, then $v(c_n - c_n') > 0$ and since $c_n = 1$, this implies $vc_n' = 0$. It follows that at least one of the summands $v(a - b_i)$ is $\geq \frac{1}{n}(nvf + v(f-g)) = vf + \frac{1}{n}v(f-g)$. We choose such an i and set $b := b_i$. □

Lemma 5.10 *Assumptions as in the foregoing lemma. Let $v(f-g) > 0$ and choose b as in that lemma. Then*
$$v\left(\frac{f(X)}{X-a} - \frac{g(X)}{X-b}\right) \;\geq\; 2vf + \frac{v(f-g)}{n}\;.$$

Proof: Since the value of the normed polynomial $(X-a)(X-b)$ is ≤ 0, we have that
$$\begin{aligned} v\left(\frac{f(X)}{X-a} - \frac{g(X)}{X-b}\right) &= v\left(f(X)(X-b) - g(X)(X-a)\right) - v(X-a)(X-b) \\ &\geq v\left(f(X)(X-b) - g(X)(X-a)\right) \\ &\geq \min\{v\left((f(X) - g(X))X\right), v\left(f(X)b - g(X)a\right)\}\;. \end{aligned}$$

In view of $vX = 0$, we find that $v\left((f(X) - g(X))X\right) = v(f-g)$. Further, let us write $f(X)b - g(X)a = f(X)b - f(X)a + f(X)a - g(X)a$. Consequently, $v(f(X)b - g(X)a) \geq \min\{v\left(f(X)b - f(X)a\right), v\left(f(X)a - g(X)a\right)\} = \min\{vf + v(b-a), v(f-g) + va)\}$. In view of the foregoing lemmata, we thus obtain that
$$v\left(\frac{f(X)}{X-a} - \frac{g(X)}{X-b}\right) \geq \min\{v(f-g), vf + vf + \frac{v(f-g)}{n}, v(f-g) + vf\} = 2vf + \frac{v(f-g)}{n}$$
since $v(f-g) > 0$ by assumption and $vf \leq 0$ because f is monic. □

Theorem 5.11 (Continuity of roots)
Let (K, v) be an arbitrary valued field, $0 \leq \alpha \in vK$ and $f, g \in K[X]$. Let $\deg f = \deg g = n$ and c_n, c_n' be the leading coefficients of f and g respectively. If
$$v(f-g) > \beta \quad \textit{with } \beta = n^n\alpha - 3n^n(vf - vc_n) + vc_n\;, \tag{5.7}$$
then we can write $f = c_n \prod_{i=1}^{n}(X - a_i)$ and $g = c_n' \prod_{i=1}^{n}(X - b_i)$ with $a_i, b_i \in \tilde{K}$ in such a way that $v(a_i - b_i) > \alpha$ for $1 \leq i \leq n$.

Proof: Condition (5.7) can also be read as $v(c_n^{-1}f - c_n^{-1}g) > n^n\alpha - 3n^n v c_n^{-1} f$, with $c_n^{-1}f$ a monic polynomial. Since the roots do not change when the polynomial is multiplied with a constant, it thus suffices to prove our assertion under the assumption that f be monic. Then in particular, $vc_n = 0$ and $vf \leq 0$. We use induction on some enumeration of the roots of f. We set $f_1 := f$ and $g_1 := g$ (in this proof, the indeces will *not* indicate derivatives). After taking the first root a_1 and finding a root b_1 of g such that $v(a_1 - b_1) > \alpha$, we repeat the procedure with $f_2 := f_1/(X - a_1)$ and $g_2 := g_1/(X - b_1)$. Having found the pair a_i, b_i of roots of the polynomials f_i, g_i, we define $f_{i+1} := f_i/(X - a_i)$ and $g_{i+1} := g_i/(X - b_i)$. We continue until we arrive at the linear polynomials f_n, g_n. Note that $vf_{i+1} = vf_i - v(X - a_i) \geq vf_i$. Since $\deg f_i = n - i + 1$, the foregoing lemma shows that $v(f_{i+1} - g_{i+1}) \geq 2vf_i + \frac{v(f_i - g_i)}{n-i+1} \geq 2vf + \frac{v(f_i - g_i)}{n-i+1}$. By induction on i, we find that $vf_i \geq vf$ and that

$$v(f_{i+1} - g_{i+1}) \geq 2vf + \frac{2vf}{n-i+1} + \ldots + \frac{2vf}{(n-1)\cdot\ldots\cdot n-i+1} + \frac{v(f-g)}{n\cdot\ldots\cdot n-i+1}\ .$$

Since $vf \leq 0$ and $v(f-g) \geq 0$, all of these values are $\geq 2nvf + \frac{v(f-g)}{n!}$.

Note that every f_i is monic. So we may employ Lemma 5.9 to determine a condition which guarantees that we find a root b_i of g_i which satisfies $v(a_i - b_i) > \alpha$. This condition is that $vf_i + \frac{1}{n-i+1}v(f_i - g_i)$ be $> \alpha$ and that $v(f_i - g_i)$ be > 0. If the latter holds, then the former is satisfied if $vf_i + \frac{1}{n}v(f_i - g_i) > \alpha$. Now we can use the lower bounds that we have computed for vf_i and $v(f_i - g_i)$. We find that our condition is satisfied if $vf + \frac{1}{n}(2nvf + \frac{v(f-g)}{n!}) > \alpha$. This is equivalent to $v(f-g) > n!(n\alpha - 3nvf)$. If this holds, then we also have that $v(f_i - g_i) \geq 2nvf + \frac{v(f-g)}{n!} > n\alpha - nvf \geq n\alpha \geq 0$, because $vf \leq 0$ and α was assumed to be ≥ 0. Since $n^n \geq n! \cdot n$ for every n and $\alpha - 3vf \geq \alpha \geq 0$, we have that $n^n\alpha - 3n^n vf \geq n!(n\alpha - 3nvf)$. This proves that if $v(f-g) > n^n\alpha - 3n^n vf$, then at every induction step we can find a root b_i of g such that $v(a_i - b_i) > \alpha$, as required. □

Our condition (5.7) on β may be quite coarse. However, its meaning is that it gives a value which lies in the convex subgroup of vK generated by the values α, vf and vc_n. This will not essentially change under better estimates. For large enough α, the condition (5.7) can be replaced by "$v(f-g) > n^n\alpha$".

The following result has a long history. See N. Shell [SHE], §15, Theorem 1, for a list of contributors.

Theorem 5.12 *Let (K, v) be an algebraically closed valued field. Then its completion $(K, v)^c$ is also algebraically closed.*

Proof: It suffices to show that $(K, v)^c$ is dense in its algebraic closure since then it follows that both are equal. So let a be an element of this algebraic closure and f its minimal polynomial over K^c. Let $\alpha \in v\widetilde{K^c}$ be given. From Corollary 6.15 we know that $v\widetilde{K^c}|vK^c$ is algebraic, that is, $v\widetilde{K^c}$ lies in the divisible hull of $vK^c = vK$. Hence after enlarging α if necessary, we may assume that $\alpha \in vK$ (and that $\alpha \geq 0$). Since f has coefficients in K^c, we can find a polynomial g with coefficients in K, monic and of the same degree as f, such that $v(f-g)$ is bigger than $n^n\alpha - 3n^n(vf - vc_n) + vc_n$. Then it follows from Theorem 5.11 that there is a root b of g such that $v(a-b) > \alpha$. But $b \in K$ since K is algebraically closed by assumption. Hence, K and thus also K^c is dense in $\widetilde{K^c}$. □

For the conclusion of this section, let us discuss the case of two polynomials f, g not having the same degree. For this case, our above proof does not work. But we can trace this case back to the equal degree case by inverting the variables. Assume that f is given in the form (5.5) and that $g = c'_m X^m + \ldots + c'_0$ with $m \geq n$ and $c'_m \neq 0$. We can also assume that both c_0 and c'_0 are nonzero; otherwise, we just replace them by any constants of value $> n^n\alpha - 3n^n(vf - vc_n) + vc_n$. We can do this also to change the value of c_0 if vc_0 is higher than this value (we will use this fact later). Now we set

$$\begin{aligned}\tilde{f} &:= c_0Y^m + c_1Y^{m-1} + \ldots + c_nY^{m-n}\\ \tilde{g} &:= c'_0Y^m + c'_1Y^{m-1} + \ldots + c'_nY^{m-n} + \ldots + c'_m\end{aligned}$$

with $Y := 1/X$. Then $v\tilde{f} = vf$ and $v(\tilde{f} - \tilde{g}) = v(f - g)$. Further, a is a root of f if and only if $a \neq 0$ and $1/a$ is a root of $\tilde{f}$, and b is a root of g if and only if $b \neq 0$ and $1/b$ is a root of $\tilde{g}$. Note that $\tilde{g}$ does not admit 0 as a root, while 0 appears precisely $m - n$ times as a zero of $\tilde{f}$. Assume that $v\left(\frac{1}{a} - \frac{1}{b}\right) > v\frac{1}{a}$. Then it follows that $v\frac{1}{a} = v\frac{1}{b}$, that is, $va = vb$. Using the lower bound $vf - vc_n$ for the value $va = vb$ of the root a of f, we find that

$$v(a-b) = v\left(\frac{1}{a} - \frac{1}{b}\right) + va + vb \geq v\left(\frac{1}{a} - \frac{1}{b}\right) + 2(vf - vc_n)\;.$$

Now let

$$\alpha' := \alpha - 2(vf - vc_n)$$

and assume that

$$v(f-g) = v(\tilde{f} - \tilde{g}) > n^n\alpha' - 3n^n(vf - vc_0) + vc_0\;. \tag{5.8}$$

Then by Theorem 5.11, $\tilde{f} = c_0Y^{m-n}\prod_{i=1}^n(Y - \frac{1}{a_i})$ and $\tilde{g} = c'_0\prod_{i=1}^m(Y - \frac{1}{b_i})$ such that $v\left(\frac{1}{a_i} - \frac{1}{b_i}\right) > \alpha'$ for $1 \leq i \leq n$, and $v\left(0 - \frac{1}{b_i}\right) = v\frac{1}{b_i} > \alpha'$ for $n+1 \leq i \leq m$. It follows that $vb_i < 2(vf - vc_n) - \alpha$ for $n+1 \leq i \leq m$. If $v\left(\frac{1}{a_i} - \frac{1}{b_i}\right) > v\frac{1}{a_i}$ is satisfied, then it also follows that $v(a_i - b_i) > \alpha$, for $1 \leq i \leq n$. On the other hand, this condition is satisfied whenever $\alpha - 2(vf - vc_n) > -va_i$ for $1 \leq i \leq n$. Using again the lower bound $vf - vc_n$ for the value of the roots of f, we find that the condition is satisfied whenever $\alpha - 2(vf - vc_n) > -(vf - vc_n)$, that is, if $\alpha > vf - vc_n$. But for this, it suffices to assume that $\alpha > 0$.

Hence, we have shown that (5.8) yields that we obtain pairs of roots with $v(a_i - b_i) > \alpha$. We have also mentioned that we can replace c_0 by another constant of value $> n^n\alpha - 3n^n(vf - vc_n) + vc_n =: \alpha_f$ if vc_0 is higher than α_f. Hence we can substitute something larger than α_f, say $\alpha + \alpha_f$ for vc_0 on the right hand side of (5.8). After this substitution, the right hand side of (5.8) will only depend on α, vf and vc_0 and will lie in the convex subgroup generated by them. We have thus proved:

Corollary 5.13 *Let (K, v) be an arbitrary valued field and $f, g \in K[X]$. Let $n := \deg f \leq m := \deg g$ and c_n, c'_m be the leading coefficients of f and g respectively. Then for every $\alpha \in vK$, $\alpha > 0$, there is a value β depending on α, vf and vc_0 and lying in the convex subgroup generated by these values, such that the following holds: If $v(f - g) > \beta$, then we can write $f = c_n\prod_{i=1}^n(X - a_i)$ and $g = c'_m\prod_{i=1}^m(X - b_i)$ with $a_i, b_i \in \tilde{K}$ in such a way that $v(a_i - b_i) > \alpha$ for $1 \leq i \leq n$, and $vb_i < 2(vf - vc_n) - \alpha$ for $n+1 \leq i \leq m$.*

Exercise 5.1 *Let* $\mathbf{K} = (K, v)$ *be an arbitrary valued field and let* a *be algebraic with minimal polynomial* f *over* K. *Prove that*

a) $v\mathrm{Tr}_{K(a)|K}(a) \geq \min_i va_i$ *and* $v\mathrm{N}_{K(a)|K}(a) = va_1 + \ldots + va_n$,

c) $va \in vK$ *if* a *is separable over* K *and the value of* a *is different from the values of the other conjugates of* a.

Exercise 5.2 *Let the assumptions be as in Lemma 5.9, with the exception that we allow* f *to be not monic and* $m := \deg g$ *to be bigger than* $\deg f$. *Let* c_n , c'_m *be the leading coefficients of* f *and* g *respectively. Show that* $vg(a) \geq m(vf - vc_n) + v(f - g)$ *and that there is a root* b *of* g *such that* $v(a - b) \geq vf - vc_n + \frac{1}{m}v(f - g) - \frac{1}{m}vc'_m$. *Explain why the latter formula in case of* $m > n$ *is not appropriate to describe the growth of* $v(a - b)$ *in dependence on* $v(f - g)$.

Exercise 5.3 *Let* f *be a polynomial of the form (5.5). In order to compute the value of all roots from the value of the coefficients, repeat the procedure that we have applied to the roots* $a_1, \ldots, a_m$ *of minimal value, preceding to Lemma 5.6. Cf. the method of* **Newton polygons** *as described e.g. in [KOB].*

Exercise 5.4 *Let* (K, v) *be a valued field and extend* v *to* $\tilde{K}$. *Let* $(K(a)|K, v)$ *be a finite vs-defectless extension and* f *the minimal polynomial of* a *over* K. *Show that if* $g \in K[X]$ *of degree* $\deg g = \deg f$ *and if* $b \in \tilde{K}$ *is a root of* g, *then also* $(K(b)|K, v)$ *is a finite vs-defectless extension with* $e(K(b)|K, v) = e(K(a)|K, v)$ *and* $f(K(b)|K, v) = f(K(a)|K, v)$ *if* $v(f - g)$ *is large enough (where* v *is the Gauß valuation on* $K(X)$ *). Discuss also the case of* $\deg g > \deg f$. *(Hint: choose a standard valuation basis of* $(K(a)|K, v)$ *and use the idea of Lemma 6.20.)*

Exercise 5.5 *Let* (K, v) *be a valued field and* $a_1, \ldots, a_n \in K$. *Show that for every* $\gamma \in vK$ *there is some* $\delta \in vK$ *such that the following holds: If* $b_1, \ldots, b_n \in K$ *satisfy* $v(a_i - b_i) > \delta$ *and if* $i_1, \ldots, i_m \in \{1, \ldots, n\}$, $m \leq n$, *then*

$$v\left(\prod_{j=1}^{m}(X - a_{i_j}) - \prod_{j=1}^{m}(X - b_{i_j})\right) > \gamma \, .$$

Can the condition $m \leq n$ *be dropped?*

5.5 Polynomial maps

Take any $n \in \mathbb{N}$. For any system $f = (f_1, \ldots, f_n)$ of n polynomials in n variables with coefficients in K, we denote by $J_f(b)$ its Jacobian matrix at $b \in K^n$. We will denote by $J_f^*(b)$ the adjoint matrix of $J_f(b)$.

Proposition 5.14 *a) Take a polynomial* $f \in \mathcal{O}[X]$ *and* $b \in \mathcal{O}$ *such that*

$$s := f'(b) \neq 0 \, .$$

Then f *induces a pseudo-linear map with pseudo-slope* s *from* $b + s\mathcal{M}$ *into* $f(b) + s^2\mathcal{M}$.

b) Take n *polynomials in* n *variables* $f_1, \ldots, f_n \in \mathcal{O}[X_1, \ldots, X_n]$ *and* $b \in \mathcal{O}^n$ *such that*

$$s := \det J_f(b) \neq 0$$

for $f = (f_1, \ldots, f_n)$. *If* $vs = 0$, *then* $J_f(b)$ *is a pseudo-companion of* f *on* $b + \mathcal{M}$ *and* f *induces an embedding from* $b + \mathcal{M}$ *into* $f(b) + \mathcal{M}$ *with value map* $\varphi = \mathrm{id}$.

In the general case, $J_f^*(b)\,f$ *induces a pseudo-linear map with pseudo-slope* s *from* $b + s\mathcal{M}^n$ *into* $J_f^*(b)f(b) + s^2\mathcal{M}^n$

Proof: Note that whenever we prove pseudo-linearity, the assertions about the range of the functions will follow from Proposition **??**.

a): For a polynomial f in one variable over a field of arbitrary characteristic, we denote by $f^{[i]}$ its i-th formal derivative (cf. [KA], [KU4]). These polynomials are defined such that the following Taylor expansion holds in arbitrary characteristic:

$$f(b+\varepsilon) \;=\; f(b)+\sum_{i=1}^{\deg f} \varepsilon^i f^{[i]}(b)\;. \tag{5.9}$$

Note that $f' = f^{[1]}$. Since $f \in \mathcal{O}[X]$, we have that $f^{[i]} \in \mathcal{O}[X]$. Since $b \in \mathcal{O}$, we also have that $f^{[i]}(b) \in \mathcal{O}$. Now take $y, z \in b + s\mathcal{M}$. Write $y = b + \varepsilon_y$ and $z = b + \varepsilon_z$ with $\varepsilon_y, \varepsilon_z \in s\mathcal{M}$. Then by (5.9),

$$\begin{aligned} f(y)-f(z) \;&=\; (\varepsilon_y-\varepsilon_z)f'(b)+\sum_{i=2}^{\deg f}(\varepsilon_y^i-\varepsilon_z^i)f^{[i]}(b) \\ &=\; s(y-z)+S(b,\varepsilon_y,\varepsilon_z)\;. \end{aligned} \tag{5.10}$$

Since

$$\begin{aligned} &\varepsilon_y^i-\varepsilon_z^i \;= \\ &=\; (\varepsilon_y-\varepsilon_z)(\varepsilon_y^{i-1}+(i-1)\varepsilon_y^{i-2}\varepsilon_z+\ldots+(i-1)\varepsilon_y^{i-2}\varepsilon_z^{i-2}+\varepsilon_y^{i-1}) \\ &\in\; (\varepsilon_y-\varepsilon_z)s\mathcal{M} \end{aligned}$$

for every $i \geq 2$, and since $f^{[i]}(b) \in \mathcal{O}$, we find that

$$S(b,\varepsilon_y,\varepsilon_z) \in (\varepsilon_y-\varepsilon_z)s\mathcal{M} \;=\; s(y-z)\mathcal{M}\;.$$

This proves that

$$v(f(y)-f(z)-s(y-z)) \;=\; vS(b,\varepsilon_y,\varepsilon_z) \;>\; vs(y-z) \tag{5.11}$$

which implies that (**??**) holds. This proves a).

b): We write $J = J_f(b)$ and $J^* = J_f^*(b)$. Then $JJ^* = (\det J)E = sE$ where E is the $n\times n$ identity matrix. Note that $J, J^* \in \mathcal{O}^{n\times n}$ by our assumptions on f and b. If $y \in K^n$ then we can write $y = cz$ with $c \in K$, $vc = vy$, $z \in \mathcal{O}^n$ and $vz = 0$. Then $Jy = cJz \in c\mathcal{O}^n$, hence $vJy = vc + vJz \geq vc = vy$. Similarly, $vJ^*y \geq vy$ for all $y \in K^n$.

Take $\varepsilon_1, \varepsilon_2 \in s\mathcal{M}^n$. The multidimensional Taylor expansion gives the following analogue of (5.10):

$$f(b+\varepsilon_1)-f(b+\varepsilon_2) \;=\; J(\varepsilon_1-\varepsilon_2)+S(b,\varepsilon_1,\varepsilon_2) \tag{5.12}$$

with

$$vS(b,\varepsilon_1,\varepsilon_2) \;>\; vs(\varepsilon_1-\varepsilon_2)\;. \tag{5.13}$$

Assume first that $vs = 0$. Then also $J^{-1} = \frac{1}{s}J^* \in \mathcal{O}^{n\times n}$, so for all $y \in K^n$, $vJ^{-1}y \geq vy$. But then, $vy = vEy = vJ^{-1}Jy \geq vJy \geq vy$, so equality must hold. We find that for all $y \in K^n$, $vJy = vy$ and similarly, $vJ^*y = vy$. In particular, this yields that J induces a value-preserving automorphism of the valued abelian group $(\mathcal{M}^n, +)$, and an isomorphism of ultrametric spaces from $\mathcal{M}^n$ onto $\mathcal{M}^n$ with value map $\varphi = \mathrm{id}$, with inverse maps induced by J^{-1}. From (5.12) and (5.13) we obtain that for $y = b + \varepsilon_1$ and $z = b + \varepsilon_2$ in $b + \mathcal{M}$,

$$v(f(y) - f(z) \; - \; J(y - z)) \; > \; vs(y - z) \; = \; v(y - z) \; = \; vJ(y - z) \; .$$

This proves that J is a pseudo-companion of f on $b + \mathcal{M}$. From Proposition 2.41 we infer that f induces an embedding from $b + \mathcal{M}$ into $f(b) + J\mathcal{M} = f(b) + \mathcal{M}$ with value map $\varphi = \mathrm{id}$.

Now we turn to the general case. We compute:

$$\begin{aligned} J^*f(y) - J^*f(z) \;&=\; J^*(f(b + y - b) - f(b + z - b)) \\ &=\; J^*J(y - z) \; + \; J^*S(b, y - b, z - b) \\ &=\; s(y - z) \; + \; J^*S(b, y - b, z - b) \; . \end{aligned}$$

By (5.13),

$$vJ^*S(b, y - b, z - b) \; \geq \; vS(b, y - b, z - b) \; > \; vs(y - z) \; .$$

Hence,

$$v\left(J^*f(y) - J^*f(z) - s(y - z)\right) \; = \; vJ^*S(b, y - b, z - b) \; > \; vs(y - z) \; .$$

This proves our assertion for the map $J_f^*(b)\, f$. □

Note that in the one-dimensional case ($n = 1$), writing $\det J_f(b) = f'(b)$ and $J_f^*(b) = 1$ makes the definition of $f_{\langle b \rangle}$ in the one-dimensional case a special case of the definition for the multi-dimensional case.

If $vs > 0$ in the multi-dimensional case, then in general $J_f(b)$ will not be a pseudo-companion of f. It is necessary to transform f in order to obtain suitable pseudo-companions. We have shown above that this can be done so that one even obtains pseudo-linear functions.

From Proposition 5.14 together with Propositions **??** and 2.41, we obtain:

Theorem 5.15 *Assume that (K, v) is spherically complete.*

a) Take a polynomial $f \in \mathcal{O}[X]$ and $b \in \mathcal{O}$ such that $s := f'(b) \neq 0$. Then f induces a pseudo-linear isomorphism of ultrametric spaces from $b + s\mathcal{M}$ onto $f(b) + s^2\mathcal{M}$, with pseudo-slope s.

b) Take n polynomials in n variables $f_1, \ldots, f_n \in \mathcal{O}[X_1, \ldots, X_n]$ and $b \in \mathcal{O}^n$ such that $s := \det J_f(b) \neq 0$ for $f = (f_1, \ldots, f_n)$. If $vs = 0$, then f induces an embedding of ultrametric spaces from $b + \mathcal{M}$ onto $f(b) + \mathcal{M}$.

In the general case, $J_f^*(b)\, f$ induces a pseudo-linear isomorphism of ultrametric spaces from $b + s\mathcal{M}^n$ onto $J_f^*(b)\, f(b) + s^2\mathcal{M}^n$, with pseudo-slope s.

5.6 The Newton Algorithm

Let us introduce and discuss the ultrametric version of the **Newton Algorithm** which is known from analysis. We take a valued field (K, v) and a polynomial $f(X) \in \mathcal{O}[X]$. From the Taylor expansion as given in Lemma 24.59 we infer the existence of some $\tilde{h}(X,Z) \in K[X,Z]$ such that

$$f(Z) \;=\; f(X) + f'(X)(Z-X) + (Z-X)^2 H_f(X,Z)\;.$$

Since f has coefficients in $\mathcal{O}$, it follows that $H_f(X,Z) \in \mathcal{O}[X,Z]$ and $f'(X) \in \mathcal{O}[X]$. Hence if $b, c \in \mathcal{O}$, then $vH_f(b,c) \geq 0$. If we set

$$c \;:=\; b - \frac{f(b)}{f'(b)}\;, \tag{5.14}$$

then we obtain

$$f(c) \;=\; f(b) + f'(b)(c-b) + (c-b)^2 H_f(b,c) \;=\; (c-b)^2 H_f(b,c)\;,$$

whence

$$vf(c) \;\geq\; 2v(c-b) \;=\; 2(vf(b) - vf'(b))\;.$$

From now on we assume that

$$vf(b) \;>\; 2vf'(b)\;; \tag{5.15}$$

then we obtain

$$vf(c) \;\geq\; vf(b) + (vf(b) - 2vf'(b)) \;>\; vf(b)\;.$$

Now we wish to iterate the procedure (5.14), that is, we take d and c in place of c and b, respectively, in (5.14). The question arises whether (5.15) still holds with c in place of b. By our assumption (5.15), we have that $v(c-b) > vf'(b)$, so that, by the Taylor expansion for f' as given in Lemma 24.59) and by the ultrametric triangle law,

$$vf'(c) \;=\; v\left(f'(b) + (c-b)G_{f'}(b,c)\right) \;=\; \min\{vf'(b), v(c-b)G_{f'}(b,c)\} \;=\; vf'(b)\;,$$

since $G_{f'}(X,Y) \in \mathcal{O}[X,Y]$. So we have, indeed,

$$vf(c) \;>\; vf(b) \;>\; 2vf'(b) \;=\; 2vf'(c)\;,$$

and we obtain

$$vf(d) \;\geq\; vf(c) + (vf(c) - 2vf'(c)) \;>\; vf(c)\;.$$

We set $c_0 := b$ and

$$c_{i+1} \;:=\; c_i - \frac{f(c_i)}{f'(c_i)} \tag{5.16}$$

for all integers $i \geq 0$. Then by induction, $vf'(c_i) = vf(b)$,

$$vf(c_{i+1}) \;\geq\; vf(c_i) + (vf(c_i) - 2vf'(b))\;,$$

$vf(c_i) - 2vf'(c_i) \geq vf(b) - 2vf'(b)$, and therefore,

$$vf(c_i) \;\geq\; vf(b) + i \cdot (vf(b) - 2vf'(b))\;. \tag{5.17}$$

Hence if vK is archimedean, then the sequence of values $vf(c_i)$ is cofinal in vK, or in other words, $f(c_i)$ becomes arbitrarily close to zero with increasing i. The algorithm can be seen as a successive refinement of the tentative root b. However, the question whether this algorithm converges will only make sense if we can show that the refinements c_i form a Cauchy sequence — or at least a pseudo Cauchy sequence in case vK is not archimedean. But this is clear, as

$$v(c_{i+1} - c_i) \;=\; vf(c_i) - vf'(c_i) \;=\; vf(c_i) - vf'(b)$$

and these values are strictly increasing with i.

Suppose that a is a limit of the (pseudo) Cauchy sequence $(c_i)_{i\in\mathbb{N}}$. That is, $v(a-c_i) = v(c_{i+1}-c_i)$ for all i. We compute, using the Taylor expansion as given in Lemma 24.59,

$$\begin{aligned} vf(a) &= v\left(f(c_i) + (a-c_i)G_f(a,c_i)\right) \;\geq\; \min\{vf(c_i), v(a-c_i)G_f(a,c_i)\}\\ &= \min\{vf(c_i), v(c_{i+1}-c_i)G_f(a,c_i)\} \;=\; \min\{vf(c_i), vf(c_i)-vf'(b)+vG_f(a,c_i)\}\\ &\geq vf(c_i) - vf'(b)\,. \end{aligned}$$

If vK is archimedean, then this means that $vf(a) = \infty$, that is, $f(a) = 0$. Hence if (K,v) is complete of rank 1, then the limit a exists and is a root of f. Further, $v(a-b) = v(a-c_0) = v(c_1-c_0) = vf(b) - vf'(b) > vf'(b)$. We summarize:

Theorem 5.16 *Take a complete field (K,v) of rank 1, $f \in \mathcal{O}[X]$ and $b \in \mathcal{O}$ such that $vf(b) > 2vf'(b)$. Then the Newton Algorithm converges to a root a of f which satisfies $v(a-b) > vf'(b)$.*

If (K,v) is not of rank 1, then a may not be a root of f. One could start the algorithm again, with a in place of b. If (K,v) is spherically complete, then by a transfinite induction one obtains a pseudo Cauchy sequence indexed by some limit ordinal, having a limit which is a root of f. In order to avoid transfinite induction, one can use much more elegant approaches, such as using fixed point theorems or our own approach which considers polynomials as maps on the valued field (as detailed in the previous section). The use of these approaches for finding roots of polynomials will be discussed in Section 9.2.

For the conclusion of this section, we show that the algorithm can be simplified by replacing $f'(c_i)$ by the constant $f'(b)$. Then our recursion reads as follows:

$$c_{i+1} \;:=\; c_i - \frac{f(c_i)}{f'(b)}\,. \tag{5.18}$$

With b, c as above, we have that

$$f'(c) \;=\; f'(b) + (c-b)G_{f'}(b,c)$$

If we set

$$d' \;:=\; c - \frac{f(c)}{f'(b)}\,,$$

then

$$\begin{aligned} f(d') &= f(c) + f'(c)(d'-c) + (d'-c)^2 H_f(c,d')\\ &= f(c) + f'(b)(d'-c) + (c-b)G_{f'}(b,c)(d'-c) + (d'-c)^2 H_f(c,d')\\ &\ (d'-c)[(c-b)G_{f'}(b,c) + (d'-c)H_f(c,d')]\,, \end{aligned}$$

whence

$$\begin{aligned} vf(d') &= v(d'-c) + v\left((c-b)G_{f'}(b,c) + (d'-c)H_f(c,d')\right) \\ &\geq vf(c) - vf'(b) + \min\{vf(b) - vf'(b) + vG_{f'}(b,c)\,,\, vf(c) - vf'(b) + vH_f(c,d')\} \\ &\geq vf(c) - vf'(b) + \min\{vf(b) - vf'(b)\,,\, vf(c) - vf'(b)\} \;=\; vf(c) + vf(b) - 2vf'(b) \\ &> vf(c)\,. \end{aligned}$$

It follows that we again obtain the estimate (5.17). So Theorem **??** remains true when the recursion formula (5.14) is replaced by (5.18). However, the convergence will in general not be as good as in the original Newton Algorithm. Indeed, we leave it as an exercise to the reader to show that with the recursion (5.14), the i in the estimate (5.17) can be replaced by $2i+1$.

Remark 5.17 For the usual Newton method of calculus it is also known that one can work with a fixed denominator, although as in the ultrametric case convergence may not be as fast. Note that one may even replace $f'(b)$ by any element c' such that $v(f'(b)-c') > vf'(b)$. This is why the refinement (9.2) in Section 9.2 works with denominator 1.

Exercise 5.6 *Take any field k and any prime p. Consider the polynomial $X^p - X - t$ over the field $k((t))$. Apply the Newton Algorithm with both recursion formulas. Assuming that $p = \operatorname{char} k$, express the root obtained by the algorithm as a power series in $k((t))$. What happens if the polynomial $X^p - X - t$ is replaced by the polynomial $X^p - X - 1/t$?*