CONTINUITY OF ROOTS FOR POLYNOMIALS OVER VALUED FIELDS

H. ĆMIEL, F.-V. KUHLMANN AND P. SZEWCZYK

Abstract. We study connections between polynomials which are close to each other, i.e., whose respective coefficients are close in the topology induced by a valuation. This paper consists of both an overview on known root continuity theorems and new results on the subject. We present theorems which have been published over the years, correcting and improving some of their formulations and proofs. We also give a sketch of several approaches to root continuity, such as employing the Newton Polygon and induction on the degree of the polynomial. We study the behavior of the irreducible factors of a given polynomial, the extensions generated by its roots and invariants connected to that polynomial under transition to a second polynomial which is sufficiently close. Further, we present applications of root continuity to the study of valued field extensions and ramification theory, including the theory of the defect.

1. Introduction

In this paper we analyze the relations between polynomials whose respective coefficients are close in the topology induced by the valuation. The basic result on the topic of root continuity states that if two polynomials are close, then so are their roots under a suitable pairing. The coefficients of the polynomials are taken to lie in a valued field (K, v). We choose an algebraic closure $\widetilde{K}$ of K and extend v from K to $\widetilde{K}$; we will denote this extended valuation by v as well. Then we extend v further to the polynomial ring $\widetilde{K}[x]$ by the Gauß valuation, which we will introduce in Section 2.1. Two polynomials are close under the Gauß valuation if for each $i \geq 0$ their coefficients of x^i are close under the valuation v of K.

A number of results from the literature will be formulated here in a corrected or improved form (e.g. Theorem 14). Theorem 16 is an improved version of theorems which can be found e.g. in [5] and [13]. Given any ε in the value group, if two polynomials f and g are sufficiently close to each other, then their roots can be paired in such a way that the value of their

Date: September 26, 2022.

Key words and phrases. Valued fields; Continuity of roots; Newton Polygon; Gauss valuation.

The first and second authors were supported by Opus grant 2017/25/B/ST1/01815 from the National Science Centre of Poland.

difference is larger than ε. Moreover, (possibly under additional assumptions) we show that several invariants for f are the same as for g, including the degree, the minimal value of the roots and the value

$$\mathrm{kras}(f) = \max\{v(\alpha - \alpha') \mid \alpha \neq \alpha' \text{ are roots of } f\}.$$

When f has only one root, we take $\mathrm{kras}(f)$ to be the value of that root.

Many of the known results will be presented with simplified proofs. In some cases we are able to both improve the formulation and simplify the proofs (e.g. in Theorem 25, which is a version of a result stated in [1]).

Further, we present several approaches to root continuity. As an example, we will consider convergent nets of polynomials in Theorem 31 (whose original formulation can be found in [12]) and in Corollary 32, which presents a converse to that theorem. Other approaches include induction on the degree of the polynomials (Theorem 50) and employing the Newton Polygon. This method in particular is developed in Sections 3 and 5. We study the situation when the Newton Polygons of two polynomials f and g coincide along an interval. Theorem 10 says that we can give a precise statement about the numbers and the values of roots of f and g, which gives more detailed information than the results in [6]. Moreover, we do not require f and g to be of the same degree, nor do we require both to be monic, so our result can be readily adapted to a generalization of [6, Theorem 2]. We also present an alteration of a result found in [2] (Theorem 23).

Apart from the connection between the roots of polynomials which are close, we also find relations between their other attributes. We employ an approach of deriving a separable polynomial from a given polynomial (see Lemma 41), which allows us to study polynomials which are not necessarily separable. We are then able to find connections between the irreducible factors of the polynomials and the extensions generated by their roots. An example of this can be found in Theorem 36 (which originates from [13]) and Theorem 42. Both results state that under a suitable pairing, the extensions generated by a root of the irreducible factors of the polynomials in question are isomorphic (either over the ground field, or over its henselization), and that the splitting fields of each pair of factors are the same. Moreover, we prove that polynomials which are close to each other define extensions with the same ramification theoretical invariants (Theorem 48).

2. Preliminaries

2.1. **Notation and basics of valuation theory.** Let $(K, +, \cdot, 0, 1)$ be a field and v a (Krull) valuation on K. That is, $v : K \to \Gamma \cup \{\infty\}$ for some ordered Abelian group Γ, where ∞ is an element greater than every element of Γ, and for all $a, b \in K$:

(a) $v(a) = \infty \iff a = 0$,
(b) $v(a + b) \geq \min\{v(a), v(b)\}$ (*ultrametric triangle law*),
(c) $v(a \cdot b) = v(a) + v(b)$.

If no confusion arises, for an element a we will write va in place of $v(a)$.

We will employ basic facts from valuation theory without proving them. For background on valuations we refer the reader to sources such as [4], [5], [8], and [10, Chapter 2].

Let x be an independent variable. We extend v from K to $\tilde{K}$, and then further to the polynomial ring $\widetilde{K}[x]$ by the *Gauß valuation*, which we will once again denote by v:

$$v\left(\sum_{i=0}^{n} a_i x^i\right) := \min_{0\leq i\leq n} va_i.$$

Observe that the Gauß valuation satisfies conditions (a)–(c). It can be extended in a canonical way to a valuation on the rational function field $\widetilde{K}(x)$ by means of setting $v(\frac{f}{g}) := vf - vg$. When we consider a valuation v on $\widetilde{K}[x]$ or $\widetilde{K}(x)$, it will always be the Gauß valuation.

Throughout this paper we will use the following notation for $f, g \in K[x]$:

$$\begin{cases} f(x) = \sum_{i=0}^{n} a_i x^i = a_n \prod_{i=1}^{n}(x-\alpha_i), & \alpha_i \in \widetilde{K},\ a_i \in K, \\ \\ g(x) = \sum_{i=0}^{m} b_i x^i = b_m \prod_{i=1}^{m}(x-\beta_i), & \beta_i \in \widetilde{K},\ b_i \in K, \end{cases} \tag{1}$$

with $m, n \geq 1$.

For an arbitrary valued field (K, v), we denote by vK its value group, by $\mathcal{O}_K$ its valuation ring and by Kv its residue field. Observe that $vK[x] = vK$. Note that $v\widetilde{K}$ is the divisible hull of vK and $\widetilde{K}v$ is the algebraic closure of Kv. Even if vK is not divisible, we will use quotients $\frac{\delta}{n}$ for $\delta \in vK$ and $n \in \mathbb{N}$, working in $v\widetilde{K}$.

The residue of an element $a \in \mathcal{O}_K$ will be denoted by av. For a polynomial $f \in \mathcal{O}_K[x]$ given by (1), we will write

$$(fv)(x) := \sum_{i=0}^{n} (a_i v) x^i \in (Kv)[x].$$

A polynomial $f \in K[x]$ will be called *separable* if it has only *simple roots*, that is, roots of multiplicity 1. An element $\alpha \in \widetilde{K}$ will be called *separable over* K if it is a root of a separable polynomial over K. Similarly, an algebraic extension $L|K$ will be called *separable* if each element in L is separable over K. The set consisting of all elements in $\widetilde{K}$ separable over K is a field, called the *separable-algebraic closure*, which we will denote by K^{sep}. If an algebraic extension $L|K$ (or polynomial f or element α) is not separable, then we will call it *inseparable*. If f only admits one root, then it will be called *purely inseparable*. Similarly, $L|K$ is *purely inseparable* if each element $a \in L$ is a root of a purely inseparable polynomial over K. Observe that in our notation, linear polynomials are both separable and purely inseparable.

A valued field (K, v) is called *Henselian* if the extension of v to $\widetilde{K}$ is unique, or equivalently, if it satisfies the assertion of Hensel's Lemma (see e.g. [4, Corollary 16.6], [5, Theorem 4.1.3]): *Take $f \in \mathcal{O}_K[x]$. If fv has a simple root $\zeta \in Kv$, then f admits a root $\alpha \in \mathcal{O}_K$ such that $\alpha v = \zeta$.*

The *henselization* K^h of (K, v) is the minimal algebraic extension of K which is Henselian (with respect to the fixed extension of v to $\widetilde{K}$).

Let L be an arbitrary algebraic extension of K. We will denote by $\mathrm{Gal}(L|K)$ the set of automorphisms of L leaving K elementwise fixed. In particular, we will write $\mathrm{Gal}\, K := \mathrm{Gal}(K^{\mathrm{sep}}|K)$. If $\sigma \in \mathrm{Gal}\, L|K$ and $\alpha \in L$, then we will write $\sigma\alpha$ in place of $\sigma(\alpha)$.

Let $L|K$ and $F|K$ be algebraic extensions of K. We say that elements $\alpha_1, \ldots, \alpha_n \in L$ are *K-linearly independent* if for every $c_1, \ldots, c_n \in K$, $\sum_{i=1}^n c_i\alpha_i = 0$ implies that $c_i = 0$ for $1 \le i \le n$. We say that $L|K$ is *linearly disjoint from $F|K$* if for every $n \in \mathbb{N}$ and every choice of K-linearly independent elements $\alpha_1, \ldots, \alpha_n \in L$, these elements will also be F-linearly independent. This relation is symmetric (see e.g. [3, Proposition 11.6.1] for more details), thus we can say that L and F are *linearly disjoint over K*. In this case we have $[L : K] = [L.F : F]$, where the *compositum* $L.F$ is the smallest subfield of $\widetilde{K}$ that contains both L and F.

2.2. **The Taylor expansion.** In this section we introduce the "characteristic blind" Taylor expansion for polynomials. This means that it does not contain any denominators of natural numbers which in positive characteristic could be equal to 0. Throughout this section, we assume (K, v) to be an arbitrary valued field and we take $f \in K[x]$ as in (1). The Taylor expansion of f employs the following *Hasse–Schmidt derivatives*:

$$\partial_i f(x) := \sum_{j=i}^{n} a_j \binom{j}{i} x^{j-i} = \sum_{j=0}^{n-i} a_{j+i} \binom{j+i}{i} x^j, \quad 0 \le i \le n. \tag{2}$$

The polynomials $\partial_i f$ yield the following polynomial identity, which is called the *characteristic blind Taylor expansion for the polynomial f*:

$$f(x+y) = \sum_{0 \le i \le n} \partial_i f(y) x^i \ . \tag{3}$$

Observe that every natural number n, taken as the element of K defined as the n-fold sum of 1, has a nonnegative value under each valuation on K. Indeed, we have $v1 = 0$, hence $vn = v(1 + \ldots + 1) \ge v1 = 0$. By the definition of the Gauß valuation, for every $j \in \{0, \ldots, n\}$ we have:

$$v\partial_i f = \min_{i \le j \le n} va_j \binom{j}{i} \ge \min_{i \le j \le n} va_j \ge \min_{0 \le j \le n} va_j = vf.$$

Lemma 1. *Take $c \in K$ and a polynomial $f \in K[x]$ of degree n, then*

$$v\partial_i f(c) \ge vf + \min\{0, (n-i)vc\} \quad \textit{for} \quad 0 \le i \le n.$$

Proof. We will employ Equation (2). If $vc \geq 0$, then we have:

$$\begin{aligned} v\partial_i f(c) &\geq \min_{i\leq j\leq n}\left\{va_j + v\binom{j}{i} + (j-i)vc\right\} \\ &\geq \min_{i\leq j\leq n} va_j = vf = vf + \min\{0,(n-i)vc\}. \end{aligned}$$

If $vc < 0$, then we have:

$$\begin{aligned} v\partial_i f(c) &\geq \min_{i\leq j\leq n}\left\{va_j + v\binom{j}{i} + (j-i)vc\right\} \geq \min_{i\leq j\leq n} va_j + (n-i)vc \\ &= vf + (n-i)vc = vf + \min\{0,(n-i)vc\}. \end{aligned}$$

□

Definition 2. For $f \in K[x]$ and $c \in K$, we set $f_c(x) := f(x+c)$.

Lemma 3. *Take $c \in K$. Given polynomials $f, g \in K[x]$, we have:*

$$v(f_c - g_c) \geq v(f-g) + \min\{0, \deg(f-g)vc\}.$$

In particular, if f and g are monic polynomials of degree n, then

$$v(f_c - g_c) \geq v(f-g) + \min\{0,(n-1)vc\}.$$

Proof. Set $h(x) := f(x) - g(x)$ and $r := \deg h$. Then from (3) we obtain:

$$v(f_c - g_c) = vh(x+c) = v\left(\sum_{0\leq i<r} \partial_i h(c)x^i\right) = \min_{0\leq i<r} v\partial_i h(c).$$

Now we use Lemma 1 to conclude:

$$\min_{0\leq i<r} v\partial_i h(c) \geq \min_{0\leq i<r}\,(vh + \min\{0,(r-i)vc\}) = vh + \min\{0, rvc\}.$$

If $\deg g = \deg f = n$ and both f and g are monic, then $r \leq n-1$ and so the above value is greater than or equal to $vh + \min\{0,(n-1)vc\}$. □

3. The Newton Polygon

We introduce a number of definitions and results on the *Newton Polygon*, as well as results on root continuity connected with this notion.

Consider a monic polynomial $f \in K[x]$ given by (1), that is, $a_n = 1$. Observe that the coefficients are symmetric functions in the roots:

$$a_k = s_{n-k}(\alpha_1, \ldots, \alpha_n).$$

We enumerate the roots so that their values form a non-decreasing sequence. Further, we enumerate the distinct values appearing in this sequence as

$$\gamma_1 < \gamma_2 < \ldots < \gamma_s.$$

In other words, for a suitable sequence

$$0 = j_0 < j_1 < \ldots < j_{s-1} < j_s = n$$

of natural numbers we have

$$v\alpha_{j_{\ell-1}+1} = v\alpha_j = v\alpha_{j_\ell} = \gamma_\ell$$

for $1 \le \ell \le s$ and $j_{\ell-1} < j \le j_\ell$. We set $k_\ell := n - j_\ell$, so we have:

$$0 = k_s < k_{s-1} < \ldots < k_1 < k_0 = n.$$

Then the multiplicity of the value γ_ℓ is

$$m_\ell := j_\ell - j_{\ell-1} = k_{\ell-1} - k_\ell.$$

We observe that $s_{j_\ell}(\alpha_1, \ldots, \alpha_n)$ is a sum of products of j_ℓ roots of f. Since $v\alpha_{j_\ell} < v\alpha_{j_\ell+1}$ for $\ell < s$, the unique product of minimal value must be $\prod_{i=1}^{j_\ell} \alpha_i$. This shows that

$$va_{k_\ell} = vs_{j_\ell}(\alpha_1, \ldots, \alpha_n) = v \prod_{1\le i\le j_\ell} \alpha_i = \sum_{1\le i\le j_\ell} v\alpha_i. \tag{4}$$

Therefore,

$$va_{k_\ell} - va_{k_{\ell-1}} = \sum_{j=j_{\ell-1}+1}^{j_\ell} v\alpha_j = m_\ell\gamma_\ell,$$

Consequently,

$$\gamma_\ell = \frac{va_{k_\ell} - va_{k_{\ell-1}}}{j_\ell - j_{\ell-1}} = -\frac{va_{k_{\ell-1}} - va_{k_\ell}}{k_{\ell-1} - k_\ell}. \tag{5}$$

Thus, if we enter the pairs $(0, va_0), \ldots, (k, va_k), \ldots, (n, va_n)$ in Cartesian coordinates representing $(\mathbb{N}\cup\{0\})\times vK\cup\{\infty\}$ and draw a line going through the point (k_ℓ, va_{k_ℓ}) and the point $(k_{\ell-1}, va_{k_{\ell-1}})$, then by (5), $-\gamma_\ell$ is the slope of this line. (If 0 is a root of f, then the first line is vertical.) Then we can compute the values of all roots of f and their multiplicities once we are able to recognize the numbers k_ℓ from the values of the coefficients of f.

First we observe that for $0 < \ell \le s$, the slope $-\gamma_\ell$ of the line going through (k_ℓ, va_{k_ℓ}) and $(k_{\ell-1}, va_{k_{\ell-1}})$ is smaller than the slope $-\gamma_{\ell-1}$ of the next line. This shows that the segments of the respective lines form the graph of an upward convex piecewise linear function from $(0, va_0)$ to (n, va_n).

Now we determine the location of the remaining points (k, va_k). Assume that $k_\ell < k < k_{\ell-1}$, that is, $j_{\ell-1} = n - k_{\ell-1} < n - k < n - k_\ell = j_\ell$. Then the products of minimal value in $s_{n-k}(\alpha_1, \ldots, \alpha_n)$ are of the form $\prod_{i=1}^{j_{\ell-1}} \alpha_i$ times a product of $n - k - j_{\ell-1}$ roots of value γ_ℓ. Hence, using (5)

$$\begin{aligned} va_k &\ge va_{k_{\ell-1}} + (n - k - j_{\ell-1})\gamma_\ell \\ &= va_{k_\ell} + (k_{\ell-1} - k_\ell)(-\gamma_\ell) + (k - k_{\ell-1})(-\gamma_\ell) \\ &= va_{k_\ell} + (k - k_\ell)(-\gamma_\ell), \end{aligned}$$

which shows that the point (k, va_k) lies on or above the line going through (k_ℓ, va_{k_ℓ}) and $(k_{\ell-1}, va_{k_{\ell-1}})$.

We note that multiplying f with a nonzero leading coefficient a_n will only shift the graph up or down by va_n but will not change neither slopes nor roots. Therefore, we obtain the values of an arbitrary polynomial of degree n in exactly the same way as above. We note:

The function described by the graph we have constructed is the largest upward convex piecewise linear function from $(0, va_0)$ *to* (n, va_n) *for which all points* (j, va_j), $0 \le j \le n$, *lie on or above its graph.*

The function described above will be denoted by NP_f. It can be seen as a function from $\mathbb{Q}^{\ge 0}$ to $v\widetilde{K} \cup \{\infty\}$. Here, we set $\mathrm{NP}_f(i) = \infty$ for $i > n$ and if $\gamma_s = \infty$, then we also set $\mathrm{NP}_f(i) = \infty$ for $i < m_s$. We will refer to both the graph and the corresponding function as the *Newton Polygon of* f. With this notion we can associate a finite set of points $(i, \mathrm{NP}_f(i))$ for $i \in \{0, \dots, n\}$. The points $(k_\ell, va_{k_\ell}) = (k_\ell, \mathrm{NP}_f(k_\ell))$, $0 \le \ell \le s$, are called the *vertices* of the corresponding Newton Polygon, the segments of the lines connecting one vertex (k_ℓ, va_{k_ℓ}) with the next vertex $(k_{\ell-1}, va_{k_{\ell-1}})$, $1 \le \ell \le s$, its *faces*, and the respective positive integers $k_{\ell-1} - k_\ell$ the *length of the face*. In terms of these notions, we have shown:

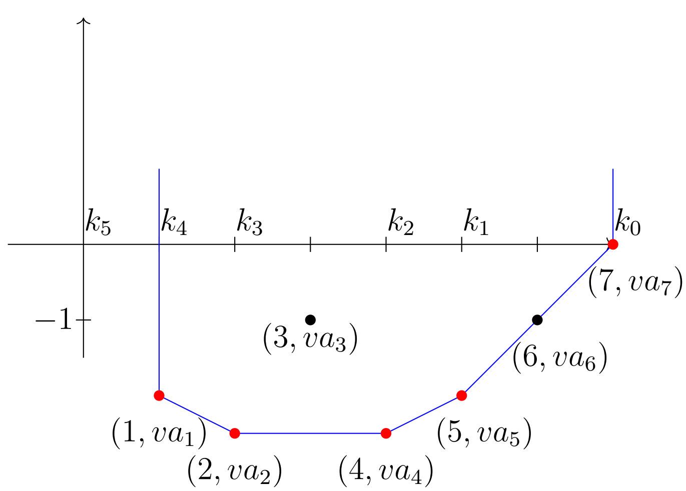
Theorem 4. *Take a polynomial* $f \in K[x]$ *as in (1). If the Newton Polygon of* f *has a face of length* k *with slope* $-\gamma$, *then* f *has exactly* k *roots of value* γ *(counted with multiplicity). In other words,*

$$v(\alpha_i) = \mathrm{NP}_f(n - i) - \mathrm{NP}_f(n - i + 1). \tag{6}$$

Example 5. Consider the field $\mathbb{Q}(\sqrt{2})$ with the 2-adic valuation v, that is, $v(2) = 1$. Take the polynomial

$$f(x) = x^7 + \frac{x^6}{2} + \frac{x^5}{4} + \frac{x^4}{4\sqrt{2}} + \frac{x^3}{2} + \frac{x^2}{4\sqrt{2}} + \frac{x}{4}.$$

The Newton Polygon of f is represented by the blue graph in the following picture:



The vertices of NP_f are represented by the red dots on the graph and are of the form (k_i, va_{k_i}). Note that the corresponding function $\mathrm{NP}_f(k)$ has value ∞ for $0 \leq k < 1$ and for $k > 7$, and is a piecewise linear function for $1 \leq k \leq 7$. From Equation (6) we know that the values of the roots of f are as follows:

$$\begin{aligned}
v\alpha_7 &= \mathrm{NP}_f(0) - \mathrm{NP}_f(1) = \infty - (-2) = \infty = \gamma_5,\\
v\alpha_6 &= \mathrm{NP}_f(1) - \mathrm{NP}_f(2) = -2 - (-2\tfrac{1}{2}) = \tfrac{1}{2} = \gamma_4,\\
v\alpha_5 &= \mathrm{NP}_f(3) - \mathrm{NP}_f(2) = 0 = \gamma_3,\\
v\alpha_4 &= \mathrm{NP}_f(4) - \mathrm{NP}_f(3) = 0 = \gamma_3,\\
v\alpha_3 &= \mathrm{NP}_f(5) - \mathrm{NP}_f(4) = -\tfrac{1}{2} = \gamma_2,\\
v\alpha_2 &= \mathrm{NP}_f(6) - \mathrm{NP}_f(5) = -1 = \gamma_1,\\
v\alpha_1 &= \mathrm{NP}_f(7) - \mathrm{NP}_f(6) = -1 = \gamma_1.
\end{aligned}$$

The following theorem allows us to study the connections between the Newton Polygons of polynomials which are sufficiently close to each other.

Theorem 6. *Consider the polynomials f and g as in (1) with f monic and $m \geq n$. For the polynomial f let the integers k_ℓ and the slopes γ_ℓ be defined as above. Fix some $\varepsilon \geq 0$, assume that $v(f-g) > n\varepsilon$ and that the set*

$$\{\ell \in \{1\dots,s\} \mid \gamma_\ell \leq \varepsilon\}$$

is nonempty. If ℓ_ε is the maximum of this set, then $\mathrm{NP}_f(k) = \mathrm{NP}_g(k)$ for $k \in \{k_{\ell_\varepsilon}, \dots, n\}$.

Proof. Fix any index $\ell \in \{0,\dots,\ell_\varepsilon\}$. By (4) we have:

$$va_{k_\ell} = \sum_{1\leq i\leq j_\ell} v\alpha_i \leq j_{\ell_\varepsilon}\gamma_{\ell_\varepsilon} \leq n\cdot\max\{0,\gamma_{\ell_\varepsilon}\} \leq n\varepsilon.$$

Since $v(a_{k_\ell} - b_{k_\ell}) \geq v(f-g) > n\varepsilon$, it follows that $va_{k_\ell} = vb_{k_\ell}$. As NP_f is upward convex, for $k_\ell \leq k \leq n$ we have $\mathrm{NP}_f(k) \leq \max\{va_{k_\ell}, va_n\} \leq n\varepsilon$. Since the point (k, va_k) lies on or above the polygon, we have $va_k \geq \mathrm{NP}_f(k)$. Hence, $vb_k \geq \min\{va_k, v(a_k - b_k)\} \geq \min\{\mathrm{NP}_f(k), n\varepsilon\} = \mathrm{NP}_f(k)$, so that also the point (k, vb_k) lies on or above NP_f. This shows that the points $(k_\ell, vb_{k\ell})$ are vertices of NP_g. Therefore, from $k = k_{\ell_\varepsilon}$ to $k = n$ we have $\mathrm{NP}_f(k) = \mathrm{NP}_g(k)$. □

Remark 7. With the notation and assumptions of the above theorem, we have $k_{\ell_\varepsilon} < n$. Indeed, the value $\gamma_{\ell_\varepsilon}$ corresponds to the face represented by the linear function on the real interval $[k_{\ell_\varepsilon}, k_{\ell_\varepsilon - 1}]$. Since the value $\gamma_{\ell_\varepsilon}$ exists by assumption, also $k_{\ell_\varepsilon - 1} \leq n$ exists and thus $k_{\ell_\varepsilon} < k_{\ell_\varepsilon - 1} \leq n$.

The above theorem tells us that along a certain interval, the Newton Polygons of the polynomials f and g have the same vertices and slopes. This yields a connection between the values of the roots of f and g. To give more details about this connection, we will require the following two lemmas.

Lemma 8. *Take f and g satisfying the assumptions of Theorem 6. Then all the slopes of NP_g located on the left of the coordinate k_{ℓ_ε} are strictly smaller than $-\varepsilon$. In particular, $(k_{\ell_\varepsilon}, vb_{k_{\ell_\varepsilon}})$ is a vertex of NP_g.*

Proof. By Theorem 6, $\mathrm{NP}_f(k) = \mathrm{NP}_g(k)$ for $k \in \{k_{\ell_\varepsilon}, \dots, n\}$. We claim that proving the first assertion will prove the second assertion. Indeed, the slope located on the right of k_{ℓ_ε} is equal to $-\gamma_{\ell_\varepsilon}$, whereas by the first assertion, the slope located on the left of this coordinate is strictly smaller than $-\varepsilon \leq -\gamma_{\ell_\varepsilon}$. This means that the point $(k_{\ell_\varepsilon}, \mathrm{NP}_g(k_{\ell_\varepsilon}))$ is a vertex of NP_g, which then by definition must be equal to the point $(k_{\ell_\varepsilon}, vb_{k_{\ell_\varepsilon}})$.

Suppose that the face of NP_g which is located on the left of the coordinate k_{ℓ_ε} has slope greater than or equal to $-\varepsilon$. Let (k, vb_k) be the left vertex of this face for some $0 \leq k < k_{\ell_\varepsilon}$. On the one hand, we have:

$$vb_k \leq \sum_{i=1}^{\ell_\varepsilon}(k_{i-1}-k_i)\gamma_i+(k_{\ell_\varepsilon}-k)\varepsilon \leq \sum_{i=1}^{\ell_\varepsilon}(k_{i-1}-k_i)\varepsilon+(k_{\ell_\varepsilon}-k)\varepsilon = (n-k)\varepsilon \leq n\varepsilon.$$

On the other hand, the supposed slope of NP_g located to the left of the coordinate k_{ℓ_ε} is greater than or equal to $-\varepsilon$, whereas the corresponding slope of NP_f is strictly smaller than $-\varepsilon$. Since $\mathrm{NP}_g(k_{\ell_\varepsilon}) = \mathrm{NP}_f(k_{\ell_\varepsilon})$, this means that

$$vb_k = \mathrm{NP}_g(k) < \mathrm{NP}_f(k) \leq va_k.$$

As a result,

$$vb_k = v(a_k - b_k) \geq v(f-g) > n\varepsilon,$$

which gives us a contradiction. □

Lemma 9. *Take polynomials f and g as in (1), with f monic and $m \geq n$. Take any $\varepsilon \geq 0$ and assume that $v(f-g) > n\varepsilon$. If $\varepsilon \geq (1-\frac{m}{n})\gamma_1$, then all slopes of NP_g along the interval $[n, m]$ are strictly greater than $-\gamma_1$. In particular, if under these assumptions NP_f and NP_g coincide along some interval $[k, n]$ for $k < n$, then (n, vb_n) is a vertex of NP_g.*

Proof. Since $v(f-g) > 0$, we have $vb_n = va_n = 0$ and $vb_i > 0$ for $i > n$. Hence, $\mathrm{NP}_g(n) \leq vb_n = 0$ and $\mathrm{NP}_g(k) > 0$ if the point (k, vb_k) is the leftmost vertex located to the right of the coordinate n. This in particular means that along the interval $[n, m]$, NP_g has positive slope. Hence, the assertions of our lemma are satisfied if $\gamma_1 \geq 0$.

Assume now that $\gamma_1 < 0$ and that $\varepsilon \geq (1-\frac{m}{n})\gamma_1$. The assertions of the lemma are satisfied if and only if for all $i \in \{n+1, \dots, m\}$ the point (i, vb_i) lies above the line going through the point $(n, \mathrm{NP}_g(n))$ with slope $-\gamma_1$ (note that by the convexity of NP_g those points cannot lie below that line). This in turn is equivalent to saying that for all $i > n$ the line going through $(n, \mathrm{NP}_g(n))$ and (i, vb_i) has slope strictly greater than $-\gamma_1$, that is,

$$vb_i > \mathrm{NP}_g(n) - (i-n)\gamma_1. \tag{7}$$

By assumption and since $\mathrm{NP}_g(n) \leq vb_n = 0$, we have:

$$vb_i \geq v(f-g) > n\varepsilon \geq (n-m)\gamma_1 \geq (n-i)\gamma_1 \geq \mathrm{NP}_g(n) - (i-n)\gamma_1.$$

Therefore (7) holds, and so our lemma is proved. □

Take any $a \in K$, $\gamma \in vK$. The set $B^\circ_\gamma(a) := \{b \in K \mid v(a-b) > \gamma\}$ will be called the *open ultrametric ball of radius* γ *centered at* a. Similarly, we define the set $S_\gamma(a) := \{b \in K \mid v(a-b) = \gamma\}$ and call it the *ultrametric sphere of radius* γ *centered at* a.

We define $n_b(f,\gamma,a)$ and $n_s(f,\gamma,a)$ to be the number of roots of f (counted with multiplicities) in $B^\circ_\gamma(a)$ and $S_\gamma(a)$, respectively. We will often consider those numbers in the case where $a = 0$. For brevity, we will write $n_b(f,\gamma) := n_b(f,\gamma,0)$ and $n_s(f,\gamma) := n_s(f,\gamma,0)$. Then $n_s(f,\gamma)$ is the number of roots of f with value γ, and $n_b(f,\gamma)$ is the number of roots of f with value strictly greater than γ.

Theorem 10. *Let the polynomials f and g be as in Theorem 6. Then:*

(a) *$n_s(g,\gamma) = n_s(f,\gamma)$ if $\gamma_1 < \gamma < \varepsilon$ or if $\gamma = \varepsilon$.*
(b) *$n_s(g,\gamma_1) \geq n_s(f,\gamma_1)$ and equality holds if and only if the point (n, vb_n) is a vertex of NP_g.*
(c) *$n_b(g,\gamma) = n_b(f,\gamma)$ for $\gamma_1 \leq \gamma \leq \varepsilon$ and $n_b(g,\gamma) \geq n_b(f,\gamma)$ for $\gamma < \gamma_1$.*
(d) *$k_{\ell_\varepsilon} = n_b(f,\gamma_{\ell_\varepsilon}) = n_b(g,\gamma_{\ell_\varepsilon}) = n_b(g,\varepsilon) = n_b(f,\varepsilon)$. In particular, if $\gamma'_{\ell'_\varepsilon}$ is chosen for g in the same manner as $\gamma_{\ell_\varepsilon}$ was chosen for f, then $\gamma'_{\ell'_\varepsilon} = \gamma_{\ell_\varepsilon}$.*
(e) *g has $m - k_1$ roots of value $\leq \gamma_1$.*
(f) *g has $m - n$ roots of value $< \gamma_1$ if and only if the point (n, vb_n) is a vertex of NP_g.*

Proof. By Theorem 6 we have $\mathrm{NP}_f(k) = \mathrm{NP}_g(k)$ for $k \in \{k_{\ell_\varepsilon}, \ldots, n\}$. This fact combined with Lemma 8 yields the following observations:

(i) Along the segment $[k_{\ell_\varepsilon}, k_1]$, all the vertices and slopes of NP_g are precisely the same as the respective vertices and slopes of NP_f. (It is possible that $k_{\ell_\varepsilon} = k_1$, and in this case we only know that the Newton Polygons of f and g share a common vertex at k_1.)
(ii) All the slopes of both NP_g and NP_f along the interval $[0, k_{\ell_\varepsilon}]$ are strictly smaller than $-\varepsilon \leq -\gamma_{\ell_\varepsilon}$. The slopes of NP_g and NP_f located on the right of the coordinate k_{ℓ_ε} are greater than or equal to $-\gamma_{\ell_\varepsilon}$.
(iii) NP_g has a face of slope $-\gamma_1$ along the interval $[k_1, k']$ for some $n \leq k' \leq m$. This face contains the corresponding face of NP_f of slope $-\gamma_1$ which runs along the interval $[k_1, n]$. The lengths of those two faces are equal if and only if $k' = n$. This happens if and only if the point (n, vb_n) is a vertex of NP_g.
(iv) With k' as in (iii), all the slopes of NP_g along the interval $[k', m]$ are strictly greater than $-\gamma_1$.

We will now combine those facts with Theorem 4.

Assertion (a) for $\gamma_1 < \gamma \leq \varepsilon$ follows from observations (i) and (ii). The remaining case is $\gamma_1 = \gamma = \varepsilon \geq 0$ because by assumption, $\gamma_1 \leq \varepsilon$. Then in particular $\varepsilon \geq (1 - \frac{m}{n})\gamma_1$. We can thus use Lemma 9 to find that $k' = n$ for k' as in observation (iii). This finishes the proof of assertion (a).

We will now prove assertion (c). Note that $n_b(f,\varepsilon) = n_b(g,\varepsilon)$ by observation (ii). For $\gamma_1 \le \gamma < \varepsilon$ we have $n_b(f,\gamma) = n_b(f,\varepsilon) + \sum_{i\in I} n_s(f,\gamma_i)$, where $I = \{2 \le i \le k_{\ell_\varepsilon} \mid \gamma < \gamma_i\}$ is a (possibly empty) set of indices. By assertion (a), $n_s(f,\gamma_i) = n_s(g,\gamma_i)$ for all $i \in I$ and $n_s(g,\delta) = 0$ for each value δ such that $\gamma_1 < \delta \le \varepsilon$ and δ is not of the form γ_i for some $i \in I$. Thus,

$$n_b(g,\gamma) = n_b(g,\varepsilon) + \sum_{i\in I} n_s(g,\gamma_i) = n_b(f,\varepsilon) + \sum_{i\in I} n_s(f,\gamma_i) = n_b(f,\gamma).$$

For $\gamma < \gamma_1$, we have $n_b(f,\gamma) = n \le n_b(g,\gamma)$.

Assertion (d) follows from (ii), assertion (b) follows from (iii), and assertions (e) and (f) follow from (iii) and (iv). □

The above computations only need that for all i either $va_i = vb_i$ or $va_i, vb_i \ge n\varepsilon$. Thus the results can be generalized, as in [6], to the case where f and g are polynomials over two different valued fields with their respective value groups contained in a common ordered Abelian group.

Corollary 11. *Let the polynomials f and g satisfy the assumptions of Theorem 6. Then $n_s(g,\gamma) = n_s(f,\gamma)$ for all γ such that $\gamma_1 \le \gamma \le \varepsilon$ if and only if (n, vb_n) is a vertex of NP_g. This holds in particular if $m = n$ or, more generally, if $\varepsilon \ge (1 - \frac{m}{n})\gamma_1$. Moreover, if $m = n$, then $n_s(g,\gamma) = n_s(f,\gamma)$ for all $\gamma \le \varepsilon$.*

Proof. The first assertion follows directly from parts (a) and (b) of Theorem 10. The particular case for $\varepsilon \ge (1 - \frac{m}{n})\gamma_1$ holds by Lemma 9. Note that for $m = n$ this condition reads $\varepsilon \ge 0$, which was our original assumption on ε. If $m = n$, then by part (f) of Theorem 10 we have $n_s(g,\gamma) = n_s(f,\gamma) = 0$ for $\gamma < \gamma_1$. □

Observe that for $c \in K$ the map $x \mapsto x - c$ induces a bijection between the roots of f and those of f_c (as in Definition 2), and a bijection between $B^\circ_\gamma(c)$ and $B^\circ_\gamma(0)$. Thus $n_b(f,\gamma,c) = n_b(f_c,\gamma,0)$.

Corollary 12. *Take $\varepsilon > 0$ and $c \in K$. Let $f, g \in K[x]$ be two monic polynomials of degree n. If*

$$v(f - g) > n\varepsilon - \min\{0, (n-1)vc\}, \tag{8}$$

then $n_b(f,\varepsilon,c) = n_b(g,\varepsilon,c)$.

Proof. From the observation before the corollary we have:

$$n_b(f,\varepsilon,c) = n_b(f_c,\varepsilon,0) \quad \text{and} \quad n_b(g,\varepsilon,c) = n_b(g_c,\varepsilon,0).$$

By Lemma 3 and by (8) we have:

$$v(f_c - g_c) \ge v(f - g) + \min\{0, (n-1)vc\} > n\varepsilon.$$

By part (c) of Theorem 10, $n_b(f_c,\varepsilon,0) = n_b(g_c,\varepsilon,0)$, thus $n_b(f,\varepsilon,c) = n_b(g,\varepsilon,c)$. □

Remark 13. Observe that the assumption $\varepsilon \geq 0$ already implies that $\varepsilon \geq (1-\frac{m}{n})\gamma_1$ if $\gamma_1 \geq 0$. Hence, this assumption is only relevant if $\gamma_1 < 0$.

If we wish to have a bound that does not require computing the value of a root of f, then we may use the fact that $\gamma_1 \geq vf$ (cf. Lemma 18), thus $(1-\frac{m}{n})vf \geq (1-\frac{m}{n})\gamma_1$; so the condition in the above lemma can be replaced by $\varepsilon \geq (1-\frac{m}{n})vf$.

Note that in the case where $\gamma_1 < 0$, the bound for ε depends also on the degree of the polynomial g. We will now show that it is impossible to specify a bound which is independent of m. Consider the field $\mathbb{Q}$ with the 2-adic valuation v and the polynomial $f(x) = 1 + \frac{1}{2}x + \frac{1}{2}x^2 + x^3$. For $j \geq 1$ we define $g_j(x) = f(x) + 2^j x^{j+3}$. Then the rightmost face of the Newton Polygons of f and g_j has slope 1. The length of that face is 1 for f and $j+1$ for g_j. This means that f has one root of value $\gamma_1 = -1$, whereas each g_j has $j+1$ roots of value -1. However, $v(f-g_j) = j$, which means that for every $\varepsilon \in v\mathbb{Q} = \mathbb{Z}$ there exists j such that $v(f-g_j) > n\varepsilon$, but f and g_j do not have the same number of roots of value γ_1.

4. Basic results

In this section we present a number of results on the basic principle of root continuity. We give possible bounds for the value $v(f-g)$ which guarantee that the roots of f and g are sufficiently close to each other under a suitable pairing. The following is a result which can be found in [13, Theorem 30.26] and [5, Theorem 2.4.7]. This theorem will be a consequence of Theorem 16 below (as observed in Remark 17).

Theorem 14. *Let f be a separable monic polynomial. Then for every $\varepsilon \in vK$ there exists $\delta \in vK$ such that the following holds:*

If g is a monic polynomial such that $v(f-g) > \delta$, then $\deg g = \deg f$, and for each root α of f there is a root β of g such that $v(\alpha-\beta) \geq \varepsilon$. Moreover, if $\varepsilon > \mathrm{kras}(f)$, then the choice of β is unique and g is separable.

The original version of the above theorem given in [13] has a slightly different formulation. It states that for an arbitrary ε, the choice of β such that $v(\alpha-\beta) \geq \varepsilon$ is unique. However, this is not true for any ε, as can be seen in the following simple example.

Example 15. Consider $K = \mathbb{Q}$ with the 2-adic valuation v on $\mathbb{Q}$, extended to $\mathbb{Q}[x]$ through the Gauß valuation. Take the polynomials

$$f(x) = g(x) = (x-1)(x+1).$$

Choose $\varepsilon = \mathrm{kras}(f) = 1$. We have $v(f-g) > \delta$ for each $\delta \in vK$, but for the root $\alpha := 1$ of f, both roots $\beta_1 := 1$ and $\beta_2 := -1$ of g satisfy:

$$v(\alpha-\beta_1) \geq \varepsilon \quad \text{and} \quad v(\alpha-\beta_2) \geq \varepsilon.$$

Thus, the pairing between the roots of f and g is not unique.

We will now prove a theorem which provides more detailed information than Theorem 14. Let (K, v) be an arbitrary valued field, and take $c \in K$, $\varepsilon \in vK$. We define the set

$$B_\varepsilon(c) := \{b \in K \mid v(c - b) \geq \varepsilon\}$$

and call it the *(closed) ultrametric ball of radius ε centered at c.*

For $f \in K[x]$ as in (1), we define:

$$\gamma(f) := \min_{1 \leq i \leq n} v\alpha_i, \quad \gamma^*(f) := \min\{\gamma(f), 0\}. \tag{9}$$

Theorem 16. *Take monic polynomials $f, g \in K[x]$ and set*

$$\varepsilon := \frac{v(f - g)}{n} + \gamma^*(f).$$

If $\varepsilon > 0$, then the following assertions hold:

(a) $\deg g = \deg f$,
(b) *for each root β of g there is a root α of f such that $v(\alpha - \beta) \geq \varepsilon$,*
(c) *for each root α of f there is a root β of g such that $v(\alpha - \beta) \geq \varepsilon$,*
(d) *$\gamma^*(f) = \gamma^*(g)$, and if $\varepsilon > \gamma(f)$, then $\gamma(f) = \gamma(g)$.*

If in addition, f is separable and $\varepsilon > \mathrm{kras}(f)$, then:

(e) *the root α in assertion (b) and the root β in assertion (c) are uniquely determined,*
(f) *g is separable,*
(g) *for every root α of f the ultrametric ball $B_\varepsilon(\alpha)$ contains precisely one root of f and precisely one root of g,*
(h) $\mathrm{kras}(f) = \mathrm{kras}(g)$.

Proof. Let f and g be given by (1). Since $\varepsilon > 0$, we have:

$$v(f - g) = n\varepsilon - n\gamma^*(f) > -n\gamma^*(f) \geq 0. \tag{10}$$

Suppose that $\deg g \neq \deg f$. Then $g - f$ is a monic polynomial, thus $v(f - g) \leq 0$, which contradicts (10). Therefore, we must have $\deg g = \deg f$ and we have proved assertion (a).

To prove assertion (b), suppose that there exists a root β of g such that $v(\alpha_i - \beta) < \varepsilon$ for every i. Then

$$vf(\beta) = \sum_{i=1}^{n} v(\beta - \alpha_i) < n\varepsilon. \tag{11}$$

Assume first that $v\beta \geq 0$. By Lemma 1 with $i = 0$ applied to $f - g$ we have

$$n\varepsilon \leq n\varepsilon - n\gamma^*(f) = v(f - g) \leq v(f(\beta) - g(\beta)) = vf(\beta),$$

which contradicts (11).

Assume now that $v\beta < 0$. Again by Lemma 1, we obtain that $vf(\beta) \geq v(f - g) + nv\beta$, so

$$n\varepsilon - n\gamma^*(f) = v(f - g) \leq vf(\beta) - nv\beta < n\varepsilon - nv\beta. \tag{12}$$

This implies that $v\beta < \gamma^*(f)$. But this means that for all i we have $v\beta < v\alpha_i$ and therefore $v(\beta - \alpha_i) = v\beta$. Hence,

$$vf(\beta) = \sum_{i=1}^{n} v(\beta - \alpha_i) = nv\beta.$$

Combining this with (12), we obtain that

$$v(f - g) \leq vf(\beta) - nv\beta = 0,$$

which contradicts (10). This shows that assertion (b) holds.

Now we will show with the same methods that for every root α of f there exists a root β of g such that $v(\alpha - \beta) > \varepsilon$. Suppose there exists a root α of f such that for every root β of g we have $v(\alpha - \beta) < \varepsilon$. Then

$$vg(\alpha) = \sum_{j=1}^{n} v(\alpha - \beta) < n\varepsilon. \tag{13}$$

If $v\alpha \geq 0$, then as before we apply Lemma 1 for $i = 0$ to $f - g$ to obtain that $n\varepsilon \leq v(f - g) \leq v(f(\alpha) - g(\alpha)) = vg(\alpha)$, which contradicts (13).

If $v\alpha < 0$, then by Lemma 1, $vg(\alpha) - nv\alpha \geq v(f - g)$. Thus,

$$n\varepsilon - n\gamma^*(f) = v(f - g) \leq vg(\alpha) - nv\alpha < n\varepsilon - nv\alpha,$$

whence $v\alpha < \gamma^*(f) \leq \gamma(f) = \min_i v\alpha_i \leq v\alpha$, a contradiction. This shows that assertion (c) holds.

To prove assertion (d), assume first that $\varepsilon > \gamma(f)$. Take k, ℓ such that $v\alpha_\ell = \gamma(f)$ and $v\beta_k = \gamma(g)$. By part (c), there exists a root β of g such that

$$v(\alpha_\ell - \beta) \geq \varepsilon > \gamma(f) = v\alpha_\ell.$$

Thus $\gamma(f) = v\alpha_\ell = v\beta \geq \gamma(g)$. By part (b), there exists a root α of f such that $v(\alpha - \beta_k) \geq \varepsilon$. Since $\varepsilon > \gamma(f)$ and $v\alpha \geq \gamma(f)$, we have:

$$\gamma(g) = v\beta_k \geq \min\{v(\alpha - \beta_k), v\alpha\} \geq \gamma(f).$$

This shows that $\gamma(f) = \gamma(g)$.

It remains to prove that $\gamma^*(f) = \gamma^*(g)$ always holds. If $\varepsilon > \gamma(f)$, then this is a consequence of the equality $\gamma(f) = \gamma(g)$. Now assume that $\varepsilon \leq \gamma(f)$; this implies that $\gamma(f) > 0$. Take β_k as above and use part (b) to find a root α of f such that $v(\alpha - \beta_k) \geq \varepsilon > 0$. Since $v\alpha \geq \gamma(f) > 0$, we obtain that $\gamma(g) = v\beta_k > 0$. Consequently, $\gamma^*(f) = 0 = \gamma^*(g)$.

Assume now that f is separable and $\varepsilon > \mathrm{kras}(f)$. We know by assertion (b) that for every root β of g there is a root α of f such that $v(\beta - \alpha) \geq \varepsilon$. Suppose that for some $i \neq j$ we have $v(\beta - \alpha_i) \geq \varepsilon$ and $v(\beta - \alpha_j) \geq \varepsilon$. It follows that $v(\alpha_i - \alpha_j) \geq \varepsilon > \mathrm{kras}(f) = \max_{i \neq j} v(\alpha_i - \alpha_j)$, which is a contradiction. This shows that α is uniquely determined, and we also see that the ultrametric balls $B_\varepsilon(\alpha_i)$, $1 \leq i \leq \deg f$, are pairwise disjoint. By assertion (c), each of the $\deg f$ balls $B_\varepsilon(\alpha_i)$ contains at least one root of g.

As $\deg f = \deg g$, this root is uniquely determined and g is separable. This proves parts (e), (f) and (g).

Take any two distinct roots β_k and β_ℓ of g. Let α_i and α_j be the distinct roots of f such that $\beta_k \in B_\varepsilon(\alpha_i)$ and $\beta_\ell \in B_\varepsilon(\alpha_j)$. Then $v(\alpha_i - \beta_k) \geq \varepsilon > v(\alpha_i - \alpha_j)$ and $v(\alpha_j - \beta_\ell) \geq \varepsilon > v(\alpha_i - \alpha_j)$, whence

$$v(\beta_k - \beta_\ell) = \min\{v(\alpha_i - \beta_k), v(\alpha_j - \beta_\ell), v(\alpha_i - \alpha_j)\} = v(\alpha_i - \alpha_j).$$

Therefore, every value $v(\beta_k - \beta_\ell)$ appears among the values $v(\alpha_i - \alpha_j)$. Since for each distinct α_i and α_j we can also find β_k and β_ℓ as above, we see that also every value $v(\alpha_i - \alpha_j)$ appears among the values $v(\beta_k - \beta_\ell)$. This implies that $\mathrm{kras}(f) = \mathrm{kras}(g)$ and concludes the proof of our theorem. □

Remark 17. Statement (a) of Theorem 16 holds already under the assumption that f, g are monic and $v(f - g) > 0$, which is weaker than the assumption of the theorem.

Further, observe that Theorem 16 implies Theorem 14. Indeed, if we choose any $\varepsilon > 0$, then by Theorem 16 for every polynomial g such that

$$v(f - g) > \delta := n\varepsilon - n\gamma^*(f),$$

the claims of Theorem 14 are satisfied.

If we choose $\varepsilon \leq 0$, then it suffices to take any positive value ε' and assume that the above inequality holds with ε' in place of ε. Then the claims of Theorem 14 are satisfied with the original value ε.

The three following lemmas can be found in varying forms in sources such as [1] (3.4.1, Proposition 3.4.1/1), [8] (Lemma 5.8, Lemma 5.9) and [12] (Lemma 1-3, Lemma 1-4). They present useful observations which will allow us to prove more refined results on the continuity of roots throughout the next sections.

Lemma 18. *If α is a root of a monic polynomial $f \in K[x]$, then $v\alpha \geq vf$. In particular, $\gamma^*(f) \geq vf$.*

Proof. Since f is monic, we have $vf \leq 0$, so the first claim is satisfied if $v\alpha \geq 0$. Thus we may assume that $v\alpha < 0$. Write f as in (1) with $a_n = 1$. Since

$$nv\alpha = v(\alpha^n) = v\left(\sum_{0 \leq i < n} a_i \alpha^i\right) \geq \min_{0 \leq i < n} \{va_i + iv\alpha\},$$

we have that

$$v\alpha \geq \min_{0 \leq i < n} \{va_i + (i + 1 - n)v\alpha\} \geq \min_{0 \leq i < n} va_i \geq vf.$$

□

In view of the above lemma, we can replace the term $\gamma^*(f)$ by vf in the definition of ε in Theorem 16. This proves useful in case we have no immediate knowledge of the roots of f. Indeed, vf is straightforward to obtain as opposed to the value $\gamma^*(f)$ which requires computing the slopes of the Newton Polygon NP_f.

Lemma 19. *Let $f, g \in K[x]$ be polynomials of degree $n \geq 1$. Assume that f is monic and take a root α of f. Then $vg(\alpha) \geq v(f-g) + nvf$.*

Proof. Write $f(x),\ g(x)$ as in (1) with $m = n$, and choose a root α of f. We apply Lemma 1 for $i = 0$ together with Lemma 18 and the fact that $vf \leq 0$ to obtain that $vg(\alpha) \geq v(f-g) + \min\{0, nv\alpha\} \geq v(f-g) + nvf$. □

The following lemma is sometimes (e.g. in [1] and [12]) cited as a separate result on the continuity of roots. It is a generalization of one of the results given in Theorem 14, since we don't require the polynomial g to be monic. This lemma will be employed, directly or indirectly, to prove a number of results (see e.g. Theorem 21, Theorem 31 and Theorem 50).

Lemma 20. *Let $f, g \in K[x]$ be polynomials of degree $n \geq 1$, assume that f is monic and let α be a root of f. If g is monic or $v(f-g) > 0$, then there exists a root β of g such that*

$$v(\beta - \alpha) \geq vf + \frac{v(f-g)}{n}.$$

Proof. Write f and g as in (1) with $m = n$. We first claim that $vb_n = 0$. This is true if g is monic. If $v(f-g) > 0$, then

$$0 < v(f-g) = \min_i \{v(a_i - b_i)\} \leq v(1 - b_n),$$

which also implies that $vb_n = 0$. Suppose that for every root β of g we have:

$$v(\beta - \alpha) < vf + \frac{v(f-g)}{n}.$$

Since $vb_n = 0$, we thus obtain that

$$vg(\alpha) = \sum_{i=1}^{n} v(\beta_i - \alpha) \leq n \cdot \max_{1 \leq i \leq n} v(\beta_i - \alpha) < nvf + v(f-g),$$

which contradicts Lemma 19. □

The following theorem is a direct application of Lemma 20 and Theorem 16. We employ the results and methods which were already introduced, in order to formulate a root continuity theorem which does not require the polynomials in question to be monic. We are, however, assuming that they are of equal degree. Another price to pay for the generalization is that the bound in the following theorem can be worse than the one in Theorem 16.

Theorem 21. *Let $f \in K[x]$ be as in (1) and take $\varepsilon > 0$. If $g \in K[x]$ is a polynomial of degree n such that*

$$v(f-g) \geq n\varepsilon - nvf + (n+1)va_n, \tag{14}$$

then assertions (b)–(d) of Theorem 16 hold. Moreover, if f is separable and $\varepsilon > \mathrm{kras}(f)$, then also assertions (e)–(h) of Theorem 16 hold.

Proof. Observe that Equation (14) is equivalent to:

$$v(a_n^{-1}f - a_n^{-1}g) \geq n\varepsilon - nv(a_n^{-1}f).$$

We will work with $\hat{g} := a_n^{-1}g$ and with the monic polynomial $\hat{f} := a_n^{-1}f$. Both polynomials have the same roots as g and f respectively. Our assumption can now be written as:

$$v(\hat{f} - \hat{g}) \geq n\varepsilon - nv\hat{f} \geq n\varepsilon > 0.$$

Fix any root α of $\hat{f}$. From Lemma 20 we infer that there exists a root β of $\hat{g}$ such that

$$v(\alpha - \beta) \geq v\hat{f} + \frac{v(\hat{f} - \hat{g})}{n} \geq v\hat{f} + \varepsilon - v\hat{f} = \varepsilon.$$

This proves assertion (c) of Theorem 16.

Observe that $v(1 - a_n^{-1}b_n) \geq v(\hat{f} - \hat{g}) > 0$. This implies that $v(a_n^{-1}b_n) = 0$, and so $va_n = vb_n$. Moreover, since $\varepsilon > 0$ and $vf \leq va_n$, (14) implies that $v(f - g) > va_n$, which in turn implies that $vf = vg$. Working with $b_n^{-1}f$ and the monic polynomial $b_n^{-1}g$, our assumption now states that:

$$v(b_n^{-1}f - b_n^{-1}g) \geq n\varepsilon - nv(b_n^{-1}g).$$

Thus we can repeat the above method to prove part (b) of Theorem 16.

To prove the further assertions, we observe that the arguments for assertions (d)–(h) in the proof of Theorem 16 do not use the assumption that the polynomials in question are monic. Thus we can employ the now proved assertions (b) and (c) and repeat the rest of the proof of Theorem 16. □

5. Using the Newton Polygon method and improving the bounds

Krasner's constant of an element $\alpha \in K^{\mathrm{sep}} \setminus K$ is defined as follows:

$$\mathrm{kras}_K(\alpha) = \max\{v(\alpha - \sigma\alpha) \mid \sigma \in \mathrm{Gal}\, K \wedge \sigma\alpha \neq \alpha\}.$$

Note that if α is a root of a separable polynomial $f \in K[x]$, then $\mathrm{kras}(f) \geq \mathrm{kras}_K(\alpha)$.

The first theorem stated in this section is an application of the Newton Polygon. Its formulation and proof are alterations of [2, Theorem 1]. At the cost of modifying the bound given in [2], we are able to drop the assumptions on the polynomials in question to be of the same degree, both monic and separable, and to have integral coefficients.

To prove the second part of the theorem, we will employ the following version of Krasner's Lemma (see e.g. [4, 16.8], [5, Theorem 4.1.7]).

Lemma 22. *Let (K, v) be a Henselian valued field. Then for every element $\alpha \in K^{\mathrm{sep}}$ the following holds: if $\beta \in K^{\mathrm{sep}} \setminus K$ satisfies $v(\alpha - \beta) > \mathrm{kras}_K(\alpha)$, then $\alpha \in K(\beta)$.*

Theorem 23. *Let (K, v) be a valued field and take $f, g \in K[x]$, written as in (1) with f monic and $m \geq n$. Fix an $\varepsilon \geq \max\{0, \mathrm{kras}(f)\}$ and assume that*

$$v(f-g) > n\varepsilon - \deg(f-g)\gamma^*(f). \tag{15}$$

Then, after suitably rearranging indices, for every $k \in \{1, \ldots, n\}$ we have $v(\alpha_k - \beta_k) > \varepsilon$.

If in addition (K, v) is Henselian and f and g are separable, then for each k we have $K(\alpha_k) \subseteq K(\beta_k)$.

Proof. Choose a root α of f and consider $f_\alpha(x) := f(x+\alpha)$, $g_\alpha(x) := g(x+\alpha)$. Denote by a'_i and b'_i the respective coefficients of f_α and g_α. We will now prove a number of results on the Newton Polygons of f_α and g_α.

If α is a root of f of multiplicity $t \geq 1$, then 0 is a root of f_α of multiplicity t. Hence for $0 \leq i \leq t-1$ we have $\mathrm{NP}_{f_\alpha}(i) = va'_i = \infty$. Moreover,

$$\left\{\begin{array}{rcl} \mathrm{NP}_{f_\alpha}(t) & = & va'_t = v(s_{n-t}(\alpha_1 - \alpha, \ldots, \alpha_n - \alpha)) \\ & = & v\left(\sum_{i_1 < i_2 < \ldots < i_{n-t}} (\alpha_{i_1} - \alpha) \cdot \ldots \cdot (\alpha_{i_{n-t}} - \alpha)\right) \\ & = & v\left(\prod_{j \in J} (\alpha_j - \alpha)\right) = \sum_{j \in J} v(\alpha - \alpha_j) \leq (n-t)\varepsilon, \end{array}\right. \tag{16}$$

where $J \subset \{1, \ldots, n\}$ is the set of all $n-t$ indices j such that $\alpha_j \neq \alpha$. By Lemma 3 we have:

$$\begin{array}{rcl} v(f_\alpha - g_\alpha) & \geq & v(f-g) + \min\{0, \deg(f-g)v\alpha\} \\ & \geq & v(f-g) + \deg(f-g)\gamma^*(f) > n\varepsilon. \end{array}$$

Therefore, for $0 \leq i \leq t-1$ we have:

$$vb'_i = v(b'_i - a'_i) \geq v(f_\alpha - g_\alpha) > n\varepsilon. \tag{17}$$

Assume now that f is not purely inseparable. We apply Theorem 6 to the polynomials f_α and g_α and the value ε. Using the notation of that theorem, we see that $\gamma_s = \infty > \varepsilon$. Since f has at least two distinct roots, the property $\varepsilon \geq \mathrm{kras}(f)$ implies that $\varepsilon \geq \gamma_{s-1}$. Hence $\ell_\varepsilon = s-1$, and since 0 is a root of f_α of multiplicity t, we have $k_{\ell_\varepsilon} = t$. As a result,

$$\mathrm{NP}_{f_\alpha}(i) = \mathrm{NP}_{g_\alpha}(i) \quad \text{for} \quad t \leq i \leq n. \tag{18}$$

For the time being, we write the indices of roots of g in such a way that

$$v(\beta_1 - \alpha) \geq v(\beta_2 - \alpha) \geq \ldots \geq v(\beta_m - \alpha). \tag{19}$$

Then $\beta_1 - \alpha, \ldots, \beta_t - \alpha$ are the t roots of g_α whose value exceeds that of the remaining roots of g_α. This means that their values correspond to the leftmost slopes of NP_{g_α}. In particular, we have:

$$v(\beta_t - \alpha) = \mathrm{NP}_{g_\alpha}(t-1) - \mathrm{NP}_{g_\alpha}(t). \tag{20}$$

By Equation (18) with $i = t$ and Equation (16) we have:

$$\mathrm{NP}_{g_\alpha}(t) = \mathrm{NP}_{f_\alpha}(t) \leq (n-t)\varepsilon. \tag{21}$$

From equation (17) we have that the second coordinate of each point (i, vb_i) on the interval $[0, t-1]$ is strictly above $n\varepsilon$. Since $\mathrm{NP}_{g_\alpha}(t) \leq (n-t)\varepsilon$, this means that each slope located on the left of the point $(t, \mathrm{NP}_{g_\alpha}(t))$ must be strictly less than the slope of the line going through $(0, n\varepsilon)$ and $(t, (n-t)\varepsilon)$. The slope of this line is $-\varepsilon$, and thus each of the roots $\beta_1 - \alpha_1, \ldots, \beta_t - \alpha_t$ has value strictly greater than ε.

Consider the set $J = \{j \in \{1, \ldots, n\} \mid \alpha_j = \alpha\}$ containing t elements. We renumber the roots $\beta_1, \ldots, \beta_t$ by the indices in J so that they are paired with the roots α_j, $j \in J$. Then we have $v(\beta_j - \alpha_j) > \varepsilon$ for all $j \in J$.

To the roots α_j, $j \in J$ we have now assigned the corresponding roots β_j. We claim that if we repeat this construction for another root $\alpha := \alpha_l$, $l \notin J$, then none of the so assigned roots β_l can be equal to β_j for any $j \in J$. Suppose that $\beta_l = \beta_j$ for some $j \in J$. Since $\varepsilon \geq \mathrm{kras}(f)$, we have:

$$v(\alpha_j - \alpha) \geq \min\{v(\alpha_j - \beta_j), v(\alpha - \beta_j)\} > \varepsilon \geq v(\alpha_j - \alpha).$$

We have now shown that for every root α of f of multiplicity t there exist at least t roots of g which satisfy our claim. Moreover, the argument above yields that those roots of g cannot be assigned to a root distinct from α. We can thus renumber the roots of g such that $v(\alpha_k - \beta_k) > \varepsilon$, assigning indices from $\{n+1, \ldots, m\}$ to the roots of g which were not chosen to be paired with any root of f.

If f is purely inseparable, then we have $\mathrm{NP}_{f_\alpha}(i) = \infty$ for $i < n$, $\mathrm{NP}_{f_\alpha}(n) = 0$. Recall from (17) that $vb_i > n\varepsilon$ for $0 \leq i < n$. Moreover, we have $v(b'_n - a'_n) \geq v(f_\alpha - g_\alpha) > 0$, hence $\mathrm{NP}_{g_\alpha}(n) = vb'_n = 0$. We write the indices $1, \ldots, n$ of roots of g as in (19). Then, by the same argument as before,

$$v(\beta_i - \alpha_k) > \varepsilon, \quad 1 \leq i \leq n.$$

Hence, also in the case when f is purely inseparable, we can renumber the roots of g in such a way that $v(\alpha_k - \beta_k) > \varepsilon \geq \mathrm{kras}(f)$.

To prove the last assertion, observe that the separability of f together with the above property implies that $v(\alpha_k - \beta_k) > \varepsilon \geq \mathrm{kras}_K(\alpha_k)$. Hence by Lemma 22, $K(\alpha_k) \subseteq K(\beta_k)$ for each k. This finishes the proof. □

Note that by Theorem 23, each ball $B^\circ_\varepsilon(\alpha_k)$ contains at least t_k roots of g. The following theorem gives us a more precise result for some roots of f.

Theorem 24. *Take $f, g \in K[x]$, written as in (1) with f monic and $m \geq n$. Fix an $\varepsilon \geq \max\{0, \mathrm{kras}(f)\}$ and assume that*

$$v(f-g) > n\varepsilon - \deg(f-g)\gamma^*(f).$$

If $v\alpha_k > \gamma(f)$, then there are precisely t_k roots of g (counted with multiplicity) in the ball $B^\circ_\varepsilon(\alpha_k)$.

If in addition $\varepsilon \geq (1 - \frac{m}{n})\gamma(f)$, then the same holds for each α_k such that $v\alpha_k = \gamma(f)$. In this case, for $n < k \leq m$ we have $v\beta_k < \gamma(f)$.

Proof. We employ Theorem 23 to find an enumeration of the roots of f and g such that $v(\alpha_k - \beta_k) > \varepsilon$ for $1 \leq k \leq n$. In particular, every ball $B^\circ_\varepsilon(\alpha_k)$ has at least t_k roots of g. As in the previous proof, we find that $B^\circ_\varepsilon(\alpha_k) \cap B^\circ_\varepsilon(\alpha_j) = \emptyset$ if $\alpha_k \neq \alpha_j$.

As in the construction of NP_f, denote by γ_i the values of roots of f in increasing order, with γ_s being the largest. Since $\varepsilon \geq \mathrm{kras}(f)$, we must have $\varepsilon \geq \gamma_{s-1} \geq \gamma_1 = \gamma(f)$ if f has at least two distinct values of roots, and $\varepsilon \geq \gamma_s = \gamma_1 = \gamma(f)$ if all roots of f have one value.

We will first assume that $\varepsilon \geq (1 - \frac{m}{n})\gamma(f)$ and prove the claim for any root α_k of f. Since $v\alpha_k \geq \gamma_1$, the fact that $v(\alpha_k - \beta_k) > \varepsilon \geq \gamma_1$ implies that also $v\beta_k \geq \gamma_1$. By Lemma 9 combined with Theorem 4, all the roots β_i, $i \in \{n+1, \ldots, m\}$, have value strictly less than γ_1. This means that there are precisely n roots of g which are eligible to be paired up with roots of f. We combine this with our previous observation to obtain that each ball $B_\varepsilon(\alpha_k)$ contains precisely t_k roots of g.

Now take any root α_k of f such that $v\alpha_k > \gamma_1$. Since the condition $\varepsilon \geq (1 - \frac{m}{n})\gamma_1$ holds if $\gamma_1 \geq 0$, we may assume that $\gamma_1 < 0$. Then $v(\alpha_k - \beta_k) \geq \varepsilon \geq 0$ implies that also $v\beta_k > \gamma_1$. Since $\varepsilon \geq \gamma_1$, the assumptions of Theorem 6 are satisfied. In particular, NP_g and NP_f coincide along the interval on which NP_f assumes slope $-\gamma_1$, that is, the interval $[k_1, n]$. If we show that on the left of the coordinate k_1, NP_g has slope strictly less than $-\gamma_1$, by Theorem 4 we will obtain that f and g have the same number of roots of value strictly greater than γ_1. We can then use the same argument as above to conclude that each ball $B_\varepsilon(\alpha_k)$ contains precisely t_k roots of g.

Since there exists a root of f of value greater than $\gamma(f)$, f has at least two distinct values of roots. We therefore must have $k_1 \neq 0$. Suppose that NP_g continues on the left of the point (k_1, va_{k_1}) with slope $-\gamma_1$. Let (i_0, vb_{i_0}), $i_0 < k_1$, be the vertex of NP_g which represents the left end of the face of NP_g that has slope $-\gamma_1$. Then $vb_{i_0} < \mathrm{NP}_f(i_0) \leq va_{i_0}$. Since $\gamma_1 < 0$ and $va_n = vb_n = 0$, this also means that $vb_{i_0} < 0$. But this implies that

$$0 > vb_{i_0} = v(b_{i_0} - a_{i_0}) \geq v(f - g) > 0,$$

which gives us a contradiction. □

The following result was presented in [1] for complete normed fields (see [1], Sect. 3.4, Proposition 1 and further results). In the present paper, its formulation has been adapted to work with valuations of arbitrary rank that are not necessarily complete. The completeness of the field is only used in [1] to obtain a unique extension of the norm from the field to its algebraic closure. However, the statement remains true when considering any valued field (K, v) and choosing any extension of v to an algebraic closure of K. Instead of restating the original proof, we note that this theorem is a special case of Theorem 24, where $m = n$. We are also able to specify a bound in our assumptions, replacing the original epsilon-delta formulation.

Theorem 25. *Take any monic polynomial $f \in K[x]$ of degree $n \geq 1$ and let α be a root of f of multiplicity t. Choose an element $\varepsilon \in vK$ such that $\varepsilon \geq \max\{0, \mathrm{kras}(f)\}$. Assume that for a monic polynomial $g \in K[x]$ of degree n we have:*

$$v(f-g) > n\varepsilon - (n-1)\gamma^*(f).$$

Then g has exactly t roots (counted with multiplicities) in $B_\varepsilon^\circ(\alpha)$.

Note that as was the case of Lemma 9, the bound for $v(f-g)$ in Theorems 23 and 24 depends on both f and g. Similarly to the observation in Remark 13, this bound cannot be made independent of the polynomial g. This is illustrated in the following example.

Example 26. We claim that there exists a monic polynomial f and polynomials g_j, $j \geq 1$, such that $\{v(f-g_j) \mid j \geq 1\}$ is cofinal in vK, but for any value $\varepsilon \in vK$ and for any root α of f, the ball $B_\varepsilon^\circ(\alpha)$ contains either no roots of g_j or all roots of g_j. In particular, we will show that $B_\varepsilon^\circ(\alpha)$ contains no roots of g_j if $\varepsilon \geq 0$.

We consider the field $\mathbb{Q}$ with the 2-adic valuation v. We set $f(x) = x - \frac{1}{2}$, $g_j(x) = f(x) + 2^j x^{j+1}$. Then $v(f-g_j) = j$. Note that all roots of g_j have value -1, same as the only root $\alpha = \frac{1}{2}$ of f. Indeed, the Newton Polygons of g_j contain only one face of length $j+1$.

Fix any positive integer j and take $g := g_j$. To look at the values $v(\frac{1}{2}-\beta)$ for any given root β of g, we will consider the polynomial $g(x+\frac{1}{2})$. The roots of $g(x+\frac{1}{2})$ are of the form $\beta - \frac{1}{2}$, where β is any root of g. We compute:

$$g\left(x+\frac{1}{2}\right) = x + 2^j\left(x+\frac{1}{2}\right)^{j+1} = x + \sum_{i=0}^{j+1}\binom{j+1}{i}2^{i-1}x^i =: \sum_{i=0}^{j+1} b_i x^i.$$

Observe that $vb_0 = -1$ and $vb_{j+1} = j$. We claim that for $1 \leq i \leq j$ we have $vb_i \geq i-1$. Note that $b_1 = 1 + \binom{j+1}{1}$, hence $vb_1 \geq 0$. For $1 < i \leq j$ we have:

$$vb_i = v\left(\binom{j+1}{i}2^{i-1}\right) \geq v\left(2^{i-1}\right) = i-1.$$

This means that $\mathrm{NP}_{g(x+\frac{1}{2})}$ contains precisely one face with slope 1, whose left and right endpoints are $(0,-1)$ and $(j+1, j)$, respectively.

Hence for each root β of g we have $v(\frac{1}{2}-\beta) = -1$. In particular, the ball $B_\gamma^\circ(\frac{1}{2})$ contains precisely $j+1$ roots of g if $\gamma < -1$ and it contains no roots of g for $\gamma \geq -1$.

Example 27. We will now show that the bound for $v(f-g)$ in Theorem 23 is sharp if we take $\varepsilon := \max\{0, \mathrm{kras}(f)\}$. To this end, we will again consider $\mathbb{Q}$ with the 2-adic valuation v and the polynomial $f(x) = x - \frac{1}{2}$. It has a single root $\alpha = \frac{1}{2}$ of value -1, so $\gamma^*(f) = -1$, $\mathrm{kras}(f) = -1$ and $\varepsilon = 0$. This time, we take $g_j(x) := f(x) + 2^j x^j$ for $j \in \mathbb{N}$. Then g_1 has the single root $\beta = \frac{1}{6}$, thus $v(\alpha - \beta) = -v(3) = 0$. On the other hand,

$v(f-g) = v(2x) = 1$ and therefore, $n\varepsilon - \deg(f-g)\gamma^*(f) = 1 = v(f-g)$. This proves that the bound given in (15) is sharp.

By using the polynomials g_j defined above we can construct examples with polynomials that are arbitrarily close to f. We observe that $v(f-g_j) = j$. We fix an arbitrary $j \in \mathbb{N}$ and set $g := g_j$.

As in Example 26, we see that

$$g\left(x+\frac{1}{2}\right) = x + \sum_{i=0}^{j} \binom{j}{i} 2^i x^i =: \sum_{i=0}^{j+1} b_i x^i.$$

Then $vb_0 = vb_1 = 0$, $vb_j = j$, and $vb_i \geq i$ for $1 < i < j$. This means that the Newton Polygon of $g(x+\frac{1}{2})$ has two faces: one with slope 0 and length 1, and the other with slope $\frac{j}{j-1}$ and length $j-1$. Hence, $g(x+\frac{1}{2})$ has one root of value 0 and $j-1$ roots of value $-\frac{j}{j-1}$. If β is any root of $g(x)$, then $\beta - \frac{1}{2}$ is a root of $g(x+\frac{1}{2})$ and therefore, $v(\alpha-\beta) = v(\beta-\frac{1}{2}) \leq 0$. On the other hand,

$$v(f-g) = j = n \cdot 0 - j \cdot (-1) = n\varepsilon - \deg(f-g)\gamma^*(f).$$

This again shows that the strict inequality in (15) is necessary even when the polynomials f and g are close to each other.[1]

We now focus on a different approach to proving root continuity theorems, which can be found in [6] and [7]. Similarly to Theorems 23 and 24, the methods presented here allow us to improve the results given in Section 4. To prove the following result, we will use the theory introduced in Section 3 in the particular case where $\deg f = \deg g$.

Theorem 28. *Take $\varepsilon > 0$, and two monic polynomials $f, g \in K[x]$ as in (1) with $m = n$. Assume that*

$$v(f-g) > n\varepsilon - (n-1)\gamma^*(f). \tag{22}$$

Then, after suitably rearranging indices, $v(\alpha_i - \beta_i) > \varepsilon$ for every i.

Proof. Choose roots $\alpha_{i_1}, \dots \alpha_{i_\ell}$ of f such that the balls $B^\circ_\varepsilon(\alpha_{i_1}), \dots, B^\circ_\varepsilon(\alpha_{i_\ell})$ are disjoint, and such that every root of f is contained in one of these balls. For each $j \in \{1, \dots, \ell\}$ and $\gamma^*(f)$ given by (9) we have:

$$n\varepsilon - (n-1)\gamma^*(f) \geq n\varepsilon - \min\{0, (n-1)v\alpha_{i_j}\}.$$

Combined with (22), this shows that condition (8) is satisfied. Thus by Corollary 12 for $c = \alpha_{i_j}$ we have $n_b(f, \varepsilon, \alpha_{i_j}) = n_b(g, \varepsilon, \alpha_{i_j})$. We can thus enumerate the roots of g by connecting them to the roots of f that are in the same ball. □

[1] This example appears to suggest that the first part of Theorem 23 remains true if "$>$" in condition (15) and the subsequent assertion is replaced by "$\geq$". However, in this case the respective condition on ε in Theorem 23 should have "$\geq$" replaced by "$>$". This and related aspects will be studied in more detail in the PhD thesis of the first author.

Theorem 29. *Take $\varepsilon > 0$, and two polynomials $f, g \in K[x]$ as in (1) with $m = n$ such that f is monic and $v(f-g) > n\varepsilon - nvf$. Then there is an enumeration of the roots of g such that $v(\alpha_i - \beta_i) > \varepsilon$ for every i.*

Proof. Recall from Lemma 18 that $vf \leq \gamma^*(f)$. Assume that $g = b_n g_0$, with g_0 a monic polynomial. We wish to show that

$$v(f - g_0) > n\varepsilon - (n-1)\gamma^*(f).$$

Since f is monic, we have $vf \leq 0$, hence $v(f-g) > \varepsilon > 0$ implies $vf = vg$. Moreover, as in the proof of Lemma 20, we see that $vb_n = 0$. We compute, using the hypothesis of the theorem:

$$\begin{aligned} v(g-g_0) &= v((b_n - 1)g_0) = v(b_n - 1) + vg_0 \geq v(f-g) + vg_0 \\ &> n\varepsilon - nvf + vg - vb_n = n\varepsilon - (n-1)vf. \end{aligned}$$

As a result, we obtain that

$$v(f-g_0) \geq \min\{v(f-g), v(g-g_0)\} > n\varepsilon - (n-1)vf \geq n\varepsilon - (n-1)\gamma^*(f).$$

Applying Theorem 28 to f and g_0 in place of g yields the required result. $\square$

Remark 30. Note that if f and g are monic with $\deg f = \deg g = n$, then $\deg(f-g) \leq n-1$. Hence under the additional assumption that $\varepsilon \geq \mathrm{kras}(f)$, Theorem 23 generalizes Theorem 28. Similarly, since $\gamma^*(f) \leq vf \leq 0$ and $\deg(f-g) \leq n$ for f and g of degree n, with the same additional assumption Theorem 23 generalizes Theorem 29. However, Theorems 28 and 29 are useful if $v(f-g)$ is a small positive value, that is, if $0 < v(f-g) < \mathrm{kras}(f)$. Consider $f(x) = x^2 - 16$ and $g(x) = x^2 - 4$ in the field $\mathbb{Q}$ with the 2-adic valuation. In this case, Theorem 23 does not work since $v(f-g) < \mathrm{kras}(f)$, but taking $\varepsilon = \frac{1}{2}$ allows us to use Theorem 28.

6. Convergent nets of polynomials

A different approach to root continuity can be found in [12]. Instead of looking at a polynomial which is 'close' to a given polynomial, we consider convergent nets of polynomials and study the behavior of their roots.

A *directed* set $(I, \leq)$ is a partially ordered set such that for all $i, j \in I$ there is $k \in I$ such that $k \geq i$ and $k \geq j$. We call $J \subseteq I$ *cofinal* in I if for every $i \in I$ there is $j \in J$ such that $j \geq i$. Note that a cofinal subset of a directed set is itself directed. Moreover, if I_1 is cofinal in I_2 and I_2 is cofinal in I_3, then I_1 is cofinal in I_3.

A *net* in a set X is a function $\varphi : I \to X$, where I is a directed set; we will denote it by $(x_i)_{i \in I}$. For $Y \subseteq X$, we say that $(x_i)_{i \in I}$ is *ultimately in* Y if there is some $i_0 \in I$ such that $x_i \in Y$ for each $i \in I$ with $i \geq i_0$.

Now assume that X is a topological space. An element $x \in X$ is a *limit of the net* $(x_i)_{i \in I}$ if for every open neighborhood U_x of x, $(x_i)_{i \in I}$ is ultimately in U_x. This fact shall be written as follows: $(x_i)_{i \in I} \to x$. In this case we will say that the net $(x_i)_{i \in I}$ is *convergent* and that it *converges* to x.

Finally, let I and J be directed sets. We say that $(x_j)_{j\in J}$ is a *subnet of* $(x_i)_{i\in I}$ if J is a cofinal subset of I.

We leave it to the reader to observe that if $I = I_1 \cup \ldots \cup I_n$, then there exists $k \in \{1,\ldots,n\}$ such that I_k is cofinal in I. In particular, if under these assumptions $(x_i)_{i\in I}$ is a net, then $(x_i)_{i\in I_k}$ is a subnet.

A particular case of convergence that will be considered in this paper is given by the topology induced by a valuation v on a valued field or ring X. In this setting, we have $(x_i)_{i\in I} \to x$ if for all $r \in vX$ there is some $i_0 \in I$ such that $v(x_i - x) > r$ for each $i \in I$ with $i \geq i_0$.

The following result can be found in [12, Lemma 1-6].

Theorem 31. *Let (K, v) be a valued field and let $(I, \leq)$ be a directed set. Consider a net $(f_i)_{i\in I}$ of monic polynomials in $K[x]$ of degree n. Moreover, let $f \in K[x]$ be the limit of $(f_i)_{i\in I}$ with respect to the valuation v, and for each $i \in I$ choose a root β_i of f_i. Then there exists a root α of f and a cofinal subset $J \subseteq I$ such that $(\beta_j)_{j\in J} \to \alpha$.*

Proof. Choose $(f_i)_{i\in I}$, $\beta_i \in \widetilde{K}$ and $f \in K[x]$ as in the theorem. Note that since $(f_i)_{i\in I} \to f$, the set $I_0 := \{i \in I \mid v(f_i - f) > 0\}$ is cofinal in I. If we find J cofinal in I_0 which satisfies our claim, then J will also be cofinal in I. We can therefore assume without loss of generality that $v(f_i - f) > 0$ for all $i \in I$. Since f is a limit of monic polynomials of degree n, it is itself a monic polynomial of degree n. This fact combined with $v(f_i - f) > 0$ shows that $vf = vf_i$ for all $i \in I$.

Let $\alpha_1, \ldots, \alpha_n \in \widetilde{K}$ be all the (not necessarily distinct) roots of f. For each $k \in \{1, \ldots, n\}$ define:

$$J(\alpha_k) := \left\{ i \in I \mid v(\alpha_k - \beta_i) \geq \frac{v(f - f_i)}{n} + vf \right\}.$$

By Lemma 20 we have that for each $i \in I$ there exists $k \in \{1, \ldots, n\}$ such that $i \in J(\alpha_k)$, that is:

$$I = J(\alpha_1) \cup \ldots \cup J(\alpha_n).$$

There exists a root α of f such that $J(\alpha)$ is cofinal in I. Set $J := J(\alpha)$. Then $(f_j)_{j\in J}$ is a net convergent to f, that is, for all $r \in vK$ there is some $j_0 \in J$ such that $v(f - f_j) > r$ for each $j \in J$ with $j \geq j_0$. Fix any element $\varepsilon \in vK$. We wish to show that, ultimately, $v(\alpha - \beta_j) > \varepsilon$. We know that there is some $j_1 \geq j_0$ such that $v(f - f_j) \geq n\varepsilon - nvf$ for each $j \in J$ with $j \geq j_1$. Thus for all $j \in J$ such that $j \geq j_1$ the following holds:

$$v(\alpha - \beta_j) \geq \frac{v(f - f_j)}{n} + vf \geq \frac{n\varepsilon - nvf}{n} + vf = \varepsilon.$$

□

From Theorem 31 we know that if $(f_i)_{i\in I} \to f$, then each net of roots β_i of f_i contains a subnet convergent to some root α of f. Theorem 25 yields the following converse result which shows that each root α of f is a limit of a suitable choice of $(\beta_i)_{i\in I}$.

Corollary 32. *Let $(I,\leq)$ be a directed set and consider a net $(f_i)_{i\in I}$ of monic polynomials in $K[x]$ of degree n with limit $f\in K[x]$. Choose a root α of f of multiplicity t. Then for every $\varepsilon\geq \max\{0,\mathrm{kras}(f)\}$ there exists $i_0\in I$ such that for every $i\geq i_0$, each of the polynomials f_i has precisely t roots in $B^\circ_\varepsilon(\alpha)$. In particular, there exists a net $(\beta_i)_{i\in I}$ of elements of $\widetilde{K}$ such that β_i is a root of f_i for each $i\in I$, and $(\beta_i)_{i\in I}\to\alpha$.*

In fact, this result can also be obtained by directly applying part (c) of Theorem 16.

7. Applications

In this section we will prove a number of results using the root continuity theorems presented earlier. Moreover, we are able to say more about the roots and the irreducible factors of polynomials which are sufficiently close to each other. As before, we consider a valued field (K,v) and the extension $(\widetilde{K},v)$. All algebraic extensions of K will be equipped with the corresponding restriction of v. We let K^c be the completion of (K,v), equipped with the canonical extension of v. The first result in this section is an application of Theorem 16. It can also be found in [13, Theorem 32.19].

Theorem 33. *The completion of a Henselian field is again Henselian.*

Proof. Take a monic polynomial $f\in\mathcal{O}_{K^c}[x]$ and assume that fv has a simple root $\zeta\in(K^c)v=Kv$. We wish to show that f admits a root $\alpha\in\mathcal{O}_{K^c}$ such that $\alpha v=\zeta$. Extend the valuation v to the algebraic closure of K^c. Since ζ is a simple root of fv, there is a unique root α of f which under this extension satisfies $\alpha v=\zeta$. If we show that $\alpha\in K^c$, then the proof will be finished. To this end, fix any $\varepsilon>0$. We wish to show the existence of an element $\beta\in K$ such that $v(\alpha-\beta)\geq\varepsilon$.

By the definition of K^c, for each $\delta>0$ we can find $g\in K[x]$ such that $v(f-g)>\delta$. Since f has integral coefficients, it follows that $g\in\mathcal{O}_K[x]$. We employ part (c) of Theorem 16 to find that if $v(f-g)$ is large enough, then g has a root β such that $v(\alpha-\beta)\geq\varepsilon$. We have to show that $\beta\in K$. Note that $v(f-g)>0$, thus we have $gv=fv$, so that also gv admits ζ as a simple root. The field (K,v) is assumed to be Henselian, so there is a root $\beta_0\in\mathcal{O}_K$ of g such that $\beta_0 v=\zeta$. Since $v(\alpha-\beta)\geq\varepsilon>0$, we have $\beta v=\alpha v=\zeta=\beta_0 v$. But ζ is a simple root of gv, so $\beta=\beta_0\in K$. □

For the next result, we need the following technical lemma.

Lemma 34. *Take an arbitrary valued field (L,v) and $c_1,\dots,c_n,d_1,\dots,d_n \in L$ such that $vc_i \le 0$ for all i. Take $\varepsilon \ge 0$ and assume that for $1 \le j \le n$,*

$$v(c_j - d_j) > \varepsilon - v \prod_{1\le i\le n} c_i .$$

Then for every subset $I \subseteq \{1,\dots,n\}$,

$$v\left(\prod_{i\in I} c_i - \prod_{i\in I} d_i\right) > \varepsilon. \tag{23}$$

Proof. Observe that since $vc_i \le 0$ for all i, the value of each product of the c_i also does not exceed 0. Since $\varepsilon \ge 0$, it follows that $v(c_j - d_j) > 0$ for all j, which in turn implies that $vc_j = vd_j$.

By induction we show that for $1 \le k \le n$,

$$v\left(\prod_{1\le i\le k} c_i - \prod_{1\le i\le k} d_i\right) > \varepsilon - v \prod_{k+1\le i\le n} c_i > \varepsilon, \tag{24}$$

where for $k = n$ we have $v\prod_{i=n+1}^{n} c_i = v1 = 0$.

Given $I \subseteq \{1,\dots,n\}$, we can without loss of generality renumber the elements c_i so that $I = \{1,\dots,k\}$ for some k. Then (24) will prove (23).

Observe that (24) holds for $k = 1$ because

$$v(c_1 - d_1) > \varepsilon - v \prod_{1\le i\le n} c_i \ge \varepsilon - v \prod_{2\le i\le n} c_i .$$

Now assume that $1 \le k < n$ and that (24) holds for k. We compute:

$$\begin{aligned} v\left(\prod_{1\le i\le k+1} c_i - \prod_{1\le i\le k+1} d_i\right) &= \\ &= v\left(c_{k+1}\prod_{1\le i\le k} c_i - d_{k+1}\prod_{1\le i\le k} d_i\right) \\ &= v\left(c_{k+1}\prod_{1\le i\le k} c_i - d_{k+1}\prod_{1\le i\le k} c_i + d_{k+1}\prod_{1\le i\le k} c_i - d_{k+1}\prod_{1\le i\le k} d_i\right) \\ &\ge \min\left\{ v(c_{k+1}-d_{k+1}) + v\prod_{1\le i\le k} c_i ,\, vd_{k+1} + v\left(\prod_{1\le i\le k} c_i - \prod_{1\le i\le k} d_i\right)\right\}. \end{aligned}$$

By the assumption of our lemma,

$$v(c_{k+1}-d_{k+1}) + v\prod_{1\le i\le k} c_i > \varepsilon - v\prod_{1\le i\le n} c_i + v\prod_{1\le i\le k} c_i \ge \varepsilon - v\prod_{k+2\le i\le n} c_i .$$

By our induction assumption,

$$\begin{aligned} vd_{k+1} + v\left(\prod_{1\le i\le k} c_i - \prod_{1\le i\le k} d_i\right) &> vd_{k+1} + \varepsilon - v \prod_{k+1\le i\le n} c_i \\ &= vc_{k+1} + \varepsilon - v \prod_{k+1\le i\le n} c_i \\ &= \varepsilon - v \prod_{k+2\le i\le n} c_i . \end{aligned}$$

This shows that (24) holds for $k+1$ in place of k and completes the proof of our lemma. □

In the special case where (K, v) is a valued field and the rational function field $K(x)$ is endowed with the Gauß valuation, we can take $c_i = x - \alpha_i$ and $d_i = x - \beta_i$. Then $vc_i \le 0$, $v(c_j - d_j) = v(\alpha_j - \beta_j)$ and $v\prod_{i=1}^n c_i = v\prod_{i=1}^n (x - \alpha_i)$. Thus with $L = K(x)$, the above lemma yields:

Corollary 35. *Take a valued field (K, v) and $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in K$. Choose a nonnegative value $\varepsilon \in vK$ and assume that for $1 \le j \le n$,*

$$v(\alpha_j - \beta_j) > \varepsilon - v \prod_{1\le i\le n} (x - \alpha_i).$$

Then for every subset $I \subseteq \{1, \dots, n\}$,

$$v\left(\prod_{i\in I}(x - \alpha_i) - \prod_{i\in I}(x - \beta_i)\right) > \varepsilon. \tag{25}$$

The following is Theorem 32.20 from [13]. It allows us to obtain information on the irreducible factors of polynomials which are sufficiently close to each other. We give a proof by use of Theorem 16.

Theorem 36. *Let (K, v) be a Henselian field and $f = f_1 \cdot \ldots \cdot f_r$ where $f_1, \dots, f_r$ are distinct monic separable irreducible polynomials over K. Then for every $\varepsilon > \max\{0, \mathrm{kras}(f)\}$ there is some $\delta \in vK$ such that for every monic polynomial $g \in K[x]$ satisfying $v(f - g) > \delta$ we have $g = g_1 \cdot \ldots \cdot g_r$, where $g_1, \dots, g_r$ are distinct monic separable irreducible polynomials over K and for each $k \in \{1, \dots, r\}$ the following assertions hold:*

(a) $\deg f_k = \deg g_k$ *and* $v(f_k - g_k) > \varepsilon$,
(b) *for every root α of f_k there exists a root β of g_k such that $K(\alpha) = K(\beta)$,*
(c) *f_k and g_k have the same splitting field,*
(d) *for all roots α of f_k and β of g_k, $K(\alpha)$ and $K(\beta)$ are isomorphic over K.*

Proof. Let $n = \deg f$ and choose any $\varepsilon > \max\{0, \mathrm{kras}(f)\}$. By assumption, f has n distinct roots $\alpha_1, \ldots, \alpha_n \in \widetilde{K}$. We take any δ satisfying

$$\delta \geq n\left(\varepsilon - vf - \gamma^*(f)\right).$$

Then the assumption $v(f-g) > \delta$ implies that

$$\varepsilon_0 := \frac{v(f-g)}{n} + \gamma^*(f) > \varepsilon - vf \geq \varepsilon > \max\{0, \mathrm{kras}(f)\}.$$

By part (g) of Theorem 16, for every α_i there exists a unique root β_i of g satisfying $v(\alpha_i - \beta_i) \geq \varepsilon_0$. Consequently, g is separable.

For every $k \in \{1, \ldots, r\}$, we define $g_k = \prod (x - \beta_i)$, where the product is taken over all i such that α_i is a root of f_k. Then the factors g_k are separable and pairwise distinct, and $\deg f_k = \deg g_k$. Thus it suffices to show that each g_k is an irreducible polynomial over K. Let α_i be a root of f_k; then by construction, β_i is a root of g_k. Take $\sigma \in \mathrm{Gal}\, K$ and let α_j be the root of f_k such that $\sigma\alpha_i = \alpha_j$. Since (K, v) is Henselian, we have $v(\alpha_j - \sigma\beta_i) = v\sigma(\alpha_i - \beta_i) = v(\alpha_i - \beta_i) \geq \varepsilon_0$. As β_j is the unique root of g such that $v(\alpha_j - \beta_j) \geq \varepsilon_0$, it follows that $\sigma\beta_i = \beta_j$. Therefore, every $\sigma \in \mathrm{Gal}\, K$ maps the roots of g_k onto roots of g_k, and thus g_k is a polynomial over K. Conversely, let β_i and β_j be two roots of g_k. Since f_k is irreducible over K, we can find $\sigma \in \mathrm{Gal}\, K$ such that $\sigma\alpha_i = \alpha_j$. By the same argument as before we find that $\sigma\beta_i = \beta_j$, which means that g_k must be irreducible.

Since for each $1 \leq i \leq n$ we have $v(\alpha_i - \beta_i) \geq \varepsilon_0 > \varepsilon - vf$, we can employ Corollary 35 for the elements α_i and β_i to obtain:

$$\forall_{1 \leq k \leq r}\ v(f_k - g_k) > \varepsilon.$$

This proves assertion (a).

Fix any root β_i of g. Since g is separable over K, β_i lies in K^{sep}. Assume first that $\beta_i \in K$. Then the corresponding irreducible polynomial g_k is of the form $x - \beta_i$ and consequently, $f_k = x - \alpha_i$. Thus $K(\beta_i) = K = K(\alpha_i)$, in which case assertion (b) of the theorem holds. Now assume that $\beta_i \in K^{\mathrm{sep}} \setminus K$. By our choice of ε_0 and by Krasner's Lemma (Lemma 22) we obtain that $K(\alpha_i) \subseteq K(\beta_i)$. But if k is such that α_i is a root of f_k, then $[K(\alpha_i) : K] = \deg f_k = \deg g_k = [K(\beta_i) : K]$, showing that $K(\alpha_i) = K(\beta_i)$. This proves assertion (b), which readily implies assertions (c) and (d). □

From the above theorem we derive the following result.

Corollary 37. *Let (K, v) be a an arbitrary valued field and take a monic separable polynomial $f \in K[x]$. Assume that f has a factorization into distinct irreducible polynomials over K^h of the form $f = f_1 \cdot \ldots \cdot f_r$. Then for every $\varepsilon > \max\{0, \mathrm{kras}(f)\}$ there is some $\delta \in vK$ such that for every monic polynomial $g \in K[x]$ satisfying $v(f-g) > \delta$ we have $g = g_1 \cdot \ldots \cdot g_r$, where $g_1, \ldots, g_r$ are distinct monic separable irreducible polynomials over K^h. Moreover, for each $k \in \{1, \ldots, r\}$, assertions (a)–(d) of Theorem 36 hold for K^h in place of K.*

For further results it will be useful to employ the notions from [8, Section 7.9]. Let H_1, H_2 be subgroups of a group G. Then for $g \in G$ the set

$$H_1 g H_2 = \{h_1 g h_2 \mid h_1 \in H_1 \wedge h_2 \in H_2\}$$

is called a *double coset of* G. The set of all such double cosets induces an equivalence relation on G, with respect to which two elements g_1, g_2 are equivalent if $H_1 g_1 H_2 = H_1 g_2 H_2$. For each element in the corresponding equivalence class we may then choose a representative. This notion will be employed for subgroups of the group $\operatorname{Gal} K$.

Notation 38. Throughout, we will consider a finite extension $L|K$ and fix the representatives $\iota_1, \ldots, \iota_s \in \operatorname{Gal} K$ of the double cosets

$$\{(\operatorname{Gal} K^h)\iota(\operatorname{Gal} L) \mid \iota \in \operatorname{Gal} K\}.$$

Note that $s \leq (\operatorname{Gal} K : \operatorname{Gal} L) \leq [L : K] < \infty$. For $\iota \in \operatorname{Gal} K$ we denote by $\operatorname{res}_L(\iota) = \iota|_L$ the restriction of ι to L. Further, by $[L : K]_{\mathrm{sep}}$ we denote the degree of the maximal separable subextension of $L|K$, and we set

$$[L : K]_{\mathrm{ins}} := \frac{[L : K]}{[L : K]_{\mathrm{sep}}}.$$

We define the *characteristic exponent of* K to be $\operatorname{charexp} K := \operatorname{char} K$ if $\operatorname{char} K > 0$, and $\operatorname{charexp} K := 1$ otherwise. Then $[L : K]_{\mathrm{ins}}$ is a power of $\operatorname{charexp} K$ for every finite extension $L|K$.

The following two lemmas can be found in [8, Lemma 7.46] in a more general form, with an arbitrary algebraic extension K' in place of K^h. For our purposes the result in the simplified form is sufficient.

Lemma 39. *An automorphism $\iota \in \operatorname{Gal} K$ lies in $\operatorname{Gal} K^h \iota_i \operatorname{Gal} L$ if and only if the isomorphism $\operatorname{res}_{\iota_i L}(\iota\iota_i^{-1}) : \iota_i L \to \iota L$ can be extended to an isomorphism of $(\iota_i L).K^h$ onto $\iota L.K^h$ over K^h.*

Proof. Take $\iota \in \operatorname{Gal} K$. Then an automorphism in $\operatorname{Gal} K$ extends $\operatorname{res}_{\iota_i L}(\iota\iota_i^{-1})$ if and only if it lies in the coset $\iota\iota_i^{-1} \operatorname{Gal} \iota_i L$. This coset is equal to

$$\iota\iota_i^{-1}\iota_i(\operatorname{Gal} L)\iota_i^{-1} = \iota(\operatorname{Gal} L)\iota_i^{-1}.$$

Hence, there is an extension of $\operatorname{res}_{\iota_i L}(\iota\iota_i^{-1})$ to an isomorphism over K^h if and only if $\iota(\operatorname{Gal} L)\iota_i^{-1} \cap \operatorname{Gal} K^h \neq \emptyset$. But this is equivalent to $\iota \in (\operatorname{Gal} K^h)\iota_i \operatorname{Gal} L$. □

Lemma 40. *Consider $L|K$ as in Notation 38 and let K_s be the maximal separable subextension of $L|K$. Assume that $K_s = K(\alpha)$ and take f to be the minimal polynomial of α over K. Let $f = f_1 \cdot \ldots \cdot f_r$ be the factorization of f into irreducible polynomials over K^h. Then $r = s$, and after suitably rearranging indices we have that $\iota_i\alpha$ is a root of f_i, so that $[(\iota_i K_s).K^h : K^h] = \deg f_i$.*

Moreover, the following equalities hold:

$$[L : K]_{\mathrm{ins}} = [(\iota_i L).K^h : K^h]_{\mathrm{ins}}, \quad 1 \leq i \leq s, \tag{26}$$

$$[L:K]=\sum_{1\le i\le s}[(\iota_i L).K^h:K^h]. \tag{27}$$

Proof. Observe that since $K_s|K$ is finite and separable, we can always find α such that $K_s=K(\alpha)$.

We will first prove Equation (26). Note that $L|K_s$ is purely inseparable and thus, $\mathrm{Gal}\,L=\mathrm{Gal}\,K_s$. As $K^h|K$ is separable, so is $(\iota_i K_s).K^h|\iota_i K_s$. Since $\iota_i L|\iota_i K_s$ is purely inseparable, it is linearly disjoint from $(\iota_i K_s).K^h|\iota_i K_s$, and $(\iota_i L).K^h|(\iota_i K_s).K^h$ is purely inseparable. This yields that $[\iota_i L:\iota_i K_s]=[(\iota_i L).K^h:(\iota_i K_s).K^h]$ and that $(\iota_i K_s).K^h|K^h$ is the maximal separable subextension of $(\iota_i L).K^h|K^h$. Hence,

$$[L:K]_{\mathrm{ins}}=[\iota_i L:\iota_i K_s]=[(\iota_i L).K^h:(\iota_i K_s).K^h]=[(\iota_i L).K^h:K^h]_{\mathrm{ins}}.$$

Consider α and $f=f_1\cdot\ldots\cdot f_r$ as in the assumption of the lemma. Then for $\iota\in\mathrm{Gal}\,K$, $\mathrm{res}_{\iota_i K_s}(\iota\iota_i^{-1}):\iota_i K_s\to\iota K_s$ can be extended to an isomorphism of $(\iota_i K_s).K^h$ onto $(\iota K_s).K^h$ over K^h if and only if $\iota_i\alpha$ and $\iota\alpha$ are roots of the same irreducible factor. By Lemma 39 we have that $\iota_i\alpha$ and $\iota\alpha$ are roots of the same irreducible factor if and only if $\iota\in\mathrm{Gal}\,K^h\iota_i\,\mathrm{Gal}\,L$. Since there are s representatives of $\mathrm{Gal}\,K^h\iota_i\,\mathrm{Gal}\,L$, there must be s irreducible factors of f, and we may enumerate them so that $\iota_i\alpha$ is a root of f_i. Then $[(\iota_i K_s).K^h:K^h]=\deg f_i$. Hence,

$$[L:K]_{\mathrm{sep}}=[K_s:K]=\deg f=\sum_{1\le i\le s}\deg f_i=\sum_{1\le i\le s}[(\iota_i K_s).K^h:K^h]. \tag{28}$$

Since the extension $(\iota_i L).K^h|(\iota_i K_s).K^h$ is purely inseparable, we have:

$$[(\iota_i K_s).K^h:K^h]=[(\iota_i L).K^h:K^h]_{\mathrm{sep}}=[(\iota_i L).K^h:K^h]\cdot[(\iota_i L).K^h:K^h]_{\mathrm{ins}}^{-1}.$$

In view of this equality and by Equation (26), multiplying Equation (28) with $[L:K]_{\mathrm{ins}}$ yields Equation (27). □

Assume that $\mathrm{charexp}\,K=p$. For $f\in K[x]$ denote by $\mathrm{ins}\,f$ the *degree of inseparability of f*, that is, the maximal number p^ν which divides every exponent in $f(x)$. In this case $f(x)$ can be written as $\tilde f(x^{p^\nu})$ for some $\tilde f\in K[x]$, and $\mathrm{ins}\,\tilde f=1$.

Lemma 41. *Fix any irreducible polynomial $f\in K[x]$ and let α be a root of f. Assume that $\mathrm{ins}\,f=p^\nu$ and take $\tilde f\in K[x]$ such that $\tilde f(x^{p^\nu})=f(x)$. Let $\tilde f=\tilde f_1\cdot\ldots\cdot\tilde f_r$ be the factorization of $\tilde f$ into irreducible polynomials over K^h and set $f_i(x):=\tilde f_i(x^{p^\nu})$. Then $f=f_1\cdot\ldots\cdot f_r$ is the factorization of f into irreducible polynomials over K^h. Moreover, for $\iota_1,\ldots,\iota_s$ chosen as in Notation 38 for $L=K(\alpha)$, and after suitably rearranging indices, $\iota_i\alpha$ is a root of f_i and $\deg f_i=[(\iota_i K(\alpha)).K^h:K^h]$. In particular, $r=s$.*

Proof. We observe that $\tilde f$ is irreducible over K since every factorization $\tilde f=\tilde g\tilde h$ leads to a factorization $f=\tilde g(x^{p^\nu})\tilde h(x^{p^\nu})$. Moreover, $\tilde f$ is separable. Indeed, if it were inseparable, then by its irreducibility we would have $\tilde f'\equiv 0$.

But this would mean that every exponent in $\tilde{f}$ is divisible by p, which contradicts the construction of $\tilde{f}$.

We have that α is a root of f if and only if α^{p^ν} is a root of $\tilde{f}$. Thus the extension $K(\alpha^{p^\nu})|K$ is separable, and $K(\alpha)|K(\alpha^{p^\nu})$ is purely inseparable. Therefore, if K_s is the maximal separable subextension of K in $K(\alpha)$, then $K_s = K(\alpha^{p^\nu})$. We apply Lemma 40 for $\tilde{f}$ in place of f and α^{p^ν} in place of α. We obtain that $\tilde{f}$ splits into irreducible factors $\tilde{f}_1, \ldots, \tilde{f}_s$ over K^h such that $(\iota_i\alpha)^{p^\nu}$ is a root of $\tilde{f}_i$, and

$$\deg \tilde{f}_i = [(\iota_i K_s).K^h : K^h].$$

In particular, $\tilde{f} = \tilde{f}_1 \cdot \ldots \cdot \tilde{f}_s$ is precisely the factorization of $\tilde{f}$ into irreducible polynomials over K^h, hence $r = s$. We define $f_i(x) := \tilde{f}_i(x^{p^\nu})$, then $f = f_1 \cdot \ldots \cdot f_s$ and $\iota_i\alpha$ is a root of f_i. We observe that $\iota_i K(\alpha) = K(\iota_i\alpha)$ and $\iota_i K_s = K\left((\iota_i\alpha)^{p^\nu}\right)$. Since $K^h|K$ is separable, also $\left(K\left((\iota_i\alpha)^{p^\nu}\right).K^h\right)|K\left((\iota_i\alpha)^{p^\nu}\right)$ is separable and thus linearly disjoint from $K(\iota_i\alpha)|K\left((\iota_i\alpha)^{p^\nu}\right)$. Therefore,

$$\begin{aligned}
[K(\iota_i\alpha).K^h : K^h] &= [(\iota_i K(\alpha)).K^h : K^h],\\
[K((\iota_i\alpha)^{p^\nu}).K^h : K^h] &= [(\iota_i K_s).K^h : K^h],\\
[K(\iota_i\alpha).K^h : K((\iota_i\alpha)^{p^\nu}).K^h] &= [K(\iota_i\alpha) : K((\iota_i\alpha)^{p^\nu})].
\end{aligned}$$

Consequently, the equality

$$[K(\iota_i\alpha).K^h : K^h] = [K(\iota_i\alpha).K^h : K((\iota_i\alpha)^{p^\nu}).K^h] \cdot [K((\iota_i\alpha)^{p^\nu}).K^h : K^h]$$

implies that

$$[(\iota_i K(\alpha)).K^h : K^h] = p^\nu [(\iota_i K_s).K^h : K^h] = p^\nu \deg \tilde{f}_i = \deg f_i.$$

This shows that the f_i are irreducible over K^h. □

The assumption on the separability of f in Corollary 37 can be dropped at the cost of adding an assumption on $\operatorname{ins} f$ and $\operatorname{ins} g$.

Theorem 42. *Let (K, v) be an arbitrary field and take $f \in K[x]$ monic and irreducible over K. Assume that f has a factorization into distinct irreducible polynomials over K^h of the form $f = f_1 \cdot \ldots \cdot f_r$. For $\operatorname{ins} f = p^\nu$ take $\tilde{f} \in K[x]$ such that $\tilde{f}(x^{p^\nu}) = f(x)$. Then for every $\varepsilon > \max\{0, \operatorname{kras}(\tilde{f})\}$ there is some $\delta \in vK$ such that the following holds: If g is any irreducible monic polynomial over K satisfying $\operatorname{ins} g \geq \operatorname{ins} f$ and $v(f - g) > \delta$, then:*

- *$\deg f = \deg g$ and $\operatorname{ins} g = \operatorname{ins} f$,*
- *$g = g_1 \cdot \ldots \cdot g_r$ where $g_1, \ldots, g_r$ are irreducible polynomials over K^h,*
- *for each $k \in \{1, \ldots, r\}$, assertions (a)–(d) of Theorem 36 hold with K^h in place of K.*

Proof. For f as given in (1), we take $\delta \geq \max\{va_i \mid 1 \leq i \leq n \wedge a_i \neq 0\}$. Choose any irreducible monic polynomial g such that $v(f - g) > \delta$. Then $a_i \neq 0$ implies that $b_i \neq 0$, so then $\deg f = \deg g$. Moreover, since $\operatorname{ins} f$ divides every i such that $a_i \neq 0$, we must also have $\operatorname{ins} g \leq \operatorname{ins} f$. Together with the hypothesis that $\operatorname{ins} g \geq \operatorname{ins} f$ we then obtain $\operatorname{ins} g = \operatorname{ins} f$; let us assume that it is equal to p^ν.

Let $\tilde{f},\ \tilde{g} \in K[x]$ be such that $\tilde{f}(x^{p^\nu}) = f(x)$, $\tilde{g}(x^{p^\nu}) = g(x)$. Then $\tilde{f}$ is separable and irreducible over K. Let $\tilde{f} = \tilde{f}_1 \cdot \ldots \cdot \tilde{f}_r$ be the factorization of $\tilde{f}$ into irreducible polynomials over K^h. By Lemma 41, $\tilde{f}_i(x^{p^\nu}) = f_i(x)$; in particular, the polynomials $\tilde{f}_i$ are distinct. Since $v(\tilde{f} - \tilde{g}) = v(\tilde{f}(x^{p^\nu}) - \tilde{g}(x^{p^\nu})) = v(f - g)$, we may apply Corollary 37 to $\tilde{f}$ and $\tilde{g}$ in the place of f and g, enlarging the originally chosen δ if necessary (note that this value only depends on f, not on g). We obtain that the respective factorization of $\tilde{g}$ into distinct irreducible polynomials over K^h is of the form $\tilde{g} = \tilde{g}_1 \cdot \ldots \cdot \tilde{g}_r$.

Set $g_k(x) := \tilde{g}_k(x^{p^\nu}) \in K^h[x]$. Since g is irreducible and $\operatorname{ins} g = \operatorname{ins} f = p^\nu$, by Lemma 41 we see that $g = g_1 \cdot \ldots \cdot g_r$ is precisely the factorization of g into irreducible polynomials over K^h.

Since a root α^{p^ν} of $\tilde{f}_k$ corresponds to a root α of f_k, the last assertion follows from Corollary 37 applied again to $\tilde{f}$ and $\tilde{g}$. □

The irreducibility of g in Theorem 42 is essential for assuring that the corresponding factorization of g over K^h yields irreducible polynomials:

Example 43. We claim that there exists $f \in K[x]$ monic and irreducible over K such that for every $\delta \in vK$ there exists a monic polynomial $g \in K[x]$ satisfying $\operatorname{ins} g = \operatorname{ins} f$ and $v(f - g) > \delta$, but f and g do not split into the same number of irreducible factors over K^h, and assertions (a)–(d) of Theorem 36 do not hold for any choice of $\varepsilon > 0$ and any of the polynomials f_k, g_k, and their respective roots.

Take (k, v) to be the rational function field $\mathbb{F}_p(t)$ with the t-adic valuation, extended canonically to $\mathbb{F}_p((t))$. Since the transcendence degree of $\mathbb{F}_p((t))$ over $\mathbb{F}_p(t)$ is infinite, we can choose an element in $\mathbb{F}_p((t^p))$ transcendental over k. For example, take $z := \sum_{i=1}^{\infty} t^{p \cdot i!} \in \mathbb{F}_p((t))$ and define $K := k(z)$. Consider the purely inseparable extension $k(z^{\frac{1}{p}})|K$ of degree p. Observe that $z^{\frac{1}{p}} = \sum_{i=1}^{\infty} t^{i!} \in \mathbb{F}_p((t))$, thus $z^{\frac{1}{p}} \in K^c$. Take f to be the minimal polynomial of $z^{\frac{1}{p}}$ over K, that is, $f(x) = x^p - z$.

To prove our claim, fix any element $\delta \in vK$ and assume without loss of generality that $\delta > 0$. Since $z^{\frac{1}{p}} \in K^c$, we can find an element $\beta \in K$ such that $v(z^{\frac{1}{p}} - \beta) > \delta$. (In fact, we can take $\beta = \sum_{i=1}^{n} t^{i!}$ for n large enough.) Consider the polynomial $g(x) = x^p - \beta^p \in K[x]$, then $\operatorname{ins} f = \operatorname{ins} g$ and $v(f - g) = v(z - \beta^p) > p\delta > \delta$. Since $K^h|K$ is separable and $f(x)$ is purely inseparable, we cannot have $z^{\frac{1}{p}} \in K^h$, so $f(x)$ must be irreducible over K^h. On the other hand, g splits into p linear factors already over K, so in particular over K^h. Clearly, $K^h(z^{\frac{1}{p}})$ cannot be equal nor isomorphic to $K^h(\beta) = K^h$ over K^h.

For the next results we will require a number of properties of the henselization K^h. The extension $(K^h|K,v)$ is *immediate*, which means that the natural embeddings of vK in vK^h and Kv in K^hv are onto ([5, Corollary 5.3.8]). Note that each algebraic extension of K^h is again a Henselian field. Thus if $(K,v)\subset(E,v)\subset(\widetilde{K},v)$, then we have:

$$(E^h,v)=(E.K^h,v). \tag{29}$$

Take any $\sigma\in\mathrm{Gal}(\widetilde{K}|K)$. Then the map

$$v\sigma=v\circ\sigma : E\ni a\mapsto v(\sigma a)\in v\widetilde{K}$$

is a valuation on E which extends K. In fact, all extensions of v from K to E are *conjugate*. That is, all the extensions are of the form $v\sigma$, where σ is an embedding of E in $\widetilde{K}$ over K ([5, Theorem 3.2.15]).

Consider the group

$$G^d:=G^d(\widetilde{K}|K,v):=\{\sigma\in\mathrm{Gal}\,K\mid v(\sigma x)=vx\text{ for all }x\in\widetilde{K}\}$$

called the *decomposition group of* $(\widetilde{K}|K,v)$, and its *fixed field*

$$K^d:=\mathrm{Fix}(G^d):=\{a\in\widetilde{K}\mid\sigma(a)=a\text{ for all }\sigma\in G^d\}.$$

We will also write $(\widetilde{K}|K,v)^d$ in place of K^d to specify which valuation we are considering. This field is the henselization of K with respect to the valuation v. For more details on ramification theory, see e.g. [4, Chapter 3], [5, Sect. 5.2], [8, Chapter 7] or [10, Sect. 2.9].

For the convenience of the reader we include a number of results on G^d and K^d. The following statements can be found in [8, Sect. 7].

Lemma 44. *Take $\iota,\sigma,\tau\in\mathrm{Gal}\,K$.*

(a) *We have $(\widetilde{K}|K,v\iota)^d=\iota^{-1}K^d$.*
(b) *If $v\sigma=v\tau$ on $\tau^{-1}K^d$, then $v\sigma=v\tau$ on $\widetilde{K}$ and $\sigma\tau^{-1}\in G^d$.*
(c) *The restriction $\mathrm{res}_{K^h}(\iota^{-1})$ is the unique isomorphism over K sending K^d onto $(\widetilde{K}|K,v\iota)^d$.*

Proof. Observe that $G^d(\widetilde{K}|K,v\iota)=\iota^{-1}\left(G^d(\widetilde{K}|K,v)\right)\iota$ ([10, Proposition 9.4], [4, (15.2)]). Assertion (a) thus follows since $\mathrm{Fix}(\iota^{-1}G\iota)=\iota^{-1}\,\mathrm{Fix}(G)$ for each automorphism group G.

Consider now σ and τ as in (b), then $v\sigma\tau^{-1}=v$ on K^d. Since the extension of v from K^d to $\widetilde{K}$ is unique, $v\sigma\tau^{-1}=v$ also holds on $\widetilde{K}$. Assertion (b) then follows from the definition of G^d.

It follows from part (a) that the restriction of ι^{-1} is the required isomorphism. If there were a second isomorphism, say σ^{-1}, then $v\sigma=v\iota$ on $\iota^{-1}K^d$, so by part (b), ι^{-1} and σ^{-1} must coincide on K^d. This proves part (c). $\square$

If w is another extension of v from K to $\widetilde{K}$, then we will denote by $K^{h(w)}$ the henselization of (K,v) in $(\widetilde{K},w)$. The above lemma allows us to represent extensions of v from K to K^h by means of the automorphism ι.

Lemma 45. *For every* $\iota \in \operatorname{Gal} K$*, the field* $(\iota^{-1}K^h, v\iota)$ *is the henselization* $(K^{h(v\iota)}, v\iota)$ *of* (K, v) *in* $(\widetilde{K}, v\iota)$*, and* (K^h, v) *is isomorphic over* K *to* $(\iota^{-1}K^h, v\iota)$ *via the uniquely determined isomorphism* $\operatorname{res}_{K^h}(\iota^{-1})$*.*

Proof. The assertion follows from the definition of the henselization together with part (a) of Lemma 44. The uniqueness of $\operatorname{res}_{K^h}(\iota^{-1})$ comes from part (c) of that lemma. □

Lemma 46. *Let* $\iota_1, \ldots, \iota_s$ *and* L *be as in Notation 38, and write* $v_i := v\iota_i$*. Then* $(L.\iota_i^{-1}K^h, v_i)$ *is the henselization of* (L, v_i) *in* $(\widetilde{K}, v_i)$*, and it is isomorphic over* K *to* $(\iota_i L.K^h, v)$ *via* ι_i*. Further, the distinct extensions of* v *from* K *to* L *are precisely the restrictions of the valuations* v_i *to* L*,* $1 \le i \le s$*.*

Proof. By virtue of Lemma 45, $(\iota_i^{-1}K^h, v_i)$ is the henselization of (K, v) in $(\widetilde{K}, v_i)$. From Equation (29) it follows that $(L.\iota_i^{-1}K^h, v_i)$ is the henselization of (L, v_i) in $(\widetilde{K}, v_i)$. The restriction of ι_i is an isomorphism from $(L.\iota_i^{-1}K^h, v\iota_i)$ onto $(\iota_i L.K^h, v)$ over K.

We turn to the second assertion of the lemma. Assume that for some $\iota \in \operatorname{Gal} K$ we have $v\iota = v\iota_i$ on L. Then $v\iota$ and $v\iota_i$ are both extensions of the same valuation from L to $\widetilde{K}$. Since those extensions are conjugate, there exists $\tau \in \operatorname{Gal} L$ such that $v\iota_i\tau = v\iota$ on $\widetilde{K}$. By Lemma 45, $\iota^{-1}K^h = \tau^{-1}\iota_i^{-1}K^h$ is the henselization of K in $(\widetilde{K}, v\iota) = (\widetilde{K}, v\iota_i\tau)$, so the restrictions of ι^{-1} and $\tau^{-1}\iota_i^{-1}$ to K^h must be equal. Hence, $\sigma := \iota\tau^{-1}\iota_i^{-1} \in \operatorname{Gal} K^h$ and thus $\iota = \sigma\iota_i\tau \in (\operatorname{Gal} K^h)\iota_i(\operatorname{Gal} L)$.

For the converse, assume that $\iota \in (\operatorname{Gal} K^h)\iota_i(\operatorname{Gal} L)$. Write $\iota = \sigma\iota_i\tau$ with $\sigma \in \operatorname{Gal} K^h$ and $\tau \in \operatorname{Gal} L$. Since $\operatorname{Gal} K^h = G^d(\widetilde{K}|K, v)$, we have $v\iota a = v\sigma\iota_i\tau a = v\iota_i a$ for all $a \in L$, that is, $v\iota = v\iota_i$ on L. □

The above lemma allows us to describe all extensions of v from K to L using the representatives $\iota_1, \ldots, \iota_s$ of the respective double cosets. In particular, the number of such distinct extensions is precisely s.

Note that the field $K^{h(v_i)} = \iota_i^{-1}K^h$ lies in the henselization $L^{h(v_i)} = L.\iota_i^{-1}K^h$ (the last equality follows from Equation (29)). Since ι_i sends $\iota_i^{-1}K^h$ onto K^h and $L.\iota_i^{-1}K^h$ onto $\iota_i L.K^h$, we find that

$$[L^{h(v_i)} : K^{h(v_i)}] = [\iota_i L.K^h : K^h] = [(\iota_i L)^h : K^h]. \tag{30}$$

We can then apply Equation (27) to obtain:

$$[L : K] = \sum_{1 \le i \le s} [L^{h(v_i)} : K^{h(v_i)}]. \tag{31}$$

The degrees $[L^{h(v_i)} : K^{h(v_i)}]$ are called *local degrees*. Hence the equation says that the degree $[L : K]$ is the sum of the associated local degrees.

Finally, we will present a result on the behavior of the following ramification theoretical invariants related to polynomials that are close to each other. The *ramification index* of $(L|K, v)$ is $\mathrm{e}(L|K, v) = (vL : vK)$, and the *inertia degree* is $\mathrm{f}\,(L|K,v) := [Lv : Kv]$. The *Fundamental Inequality* (which can be found e.g. in [5]) states that

$$[L:K] \geq \sum_{1\leq i\leq s} \mathrm{e}(L|K,v_i)\cdot \mathrm{f}\,(L|K,v_i).$$

If v extends uniquely from K to L, then by *the Lemma of Ostrowski* (see e.g. [11]) we have

$$[L:K] = p^{\nu}\cdot \mathrm{e}(L|K,v)\cdot \mathrm{f}\,(L|K,v),$$

where $p = \operatorname{charexp} Kv$. The factor p^{ν} is called the *defect* of the extension $(L|K,v)$ and denoted by $\mathrm{d}(L|K,v)$.

In general, for a finite extension L of a valued field (K,v), we can define defects in a similar manner, using the fact that the valuation in each extension $L^{h(v_i)}|K^{h(v_i)}$ extends uniquely and that henselizations are immediate extensions:

$$\mathrm{d}(L|K,v_i) := \frac{[L^{h(v_i)} : K^{h(v_i)}]}{\mathrm{e}(L|K,v_i)\cdot \mathrm{f}\,(L|K,v_i)}. \tag{32}$$

This is often called the *Henselian defect*. We can now combine Equations (31) and (32), together with the definition of the defect, to obtain the following version of the Fundamental Inequality:

$$[L:K] = \sum_{1\leq i\leq s} \mathrm{d}(L|K,v_i)\cdot \mathrm{e}(L|K,v_i)\cdot \mathrm{f}\,(L|K,v_i). \tag{33}$$

Lemma 47 ([8], Lemma 11.2)**.** *Let $\iota_1,\dots,\iota_s$ and L be as in Notation 38. Then for each $1\leq i\leq s$ we have:*

$$\begin{array}{rcccl} \mathrm{d}(L|K,v_i) & = & \mathrm{d}((\iota_i L).K^h|K^h,v) & = & \mathrm{d}(\iota_i L|K,v),\\ \mathrm{e}(L|K,v_i) & = & \mathrm{e}((\iota_i L).K^h|K^h,v) & = & \mathrm{e}(\iota_i L|K,v),\\ \mathrm{f}\,(L|K,v_i) & = & \mathrm{f}\,((\iota_i L).K^h|K^h,v) & = & \mathrm{f}\,(\iota_i L|K,v). \end{array}$$

Moreover, the following equality holds:

$$[L:K] = \sum_{1\leq i\leq s} \mathrm{d}(\iota_i L|K,v)\cdot \mathrm{e}(\iota_i L|K,v)\cdot \mathrm{f}\,(\iota_i L|K,v). \tag{34}$$

Proof. Since henselizations are immediate extensions, we obtain:

$$\mathrm{f}\,(L|K,v_i) = [Lv_i : Kv_i] = [L^{h(v_i)}v_i : K^{h(v_i)}v_i] = [(L.\iota_i^{-1}K^h)v_i : (\iota_i^{-1}K^h)v_i].$$

As observed before, ι_i sends $\iota_i^{-1}K^h$ onto K^h and $L.\iota_i^{-1}K^h$ onto $\iota_i L.K^h$. Therefore, the above number is equal to

$$[(\iota_i L.K^h)v : K^h v] = [(\iota_i L)^h v : K^h v] = [\iota_i L v : Kv] = \mathrm{f}\,(\iota_i L|K,v).$$

The result for $\mathrm{e}(L|K,v_i)$ is proved analogously from the same observations. The result for $\mathrm{d}(L|K,v_i)$ then follows by Equations (32) and (30). Those equalities together with Equation (33) imply Equation (34). □

The notions and results presented in the above lemmas now allow us to formulate the following root continuity theorem.

Theorem 48. *Let (K, v) be an arbitrary valued field, $f \in K[x]$ an irreducible monic polynomial over K and $\alpha \in \widetilde{K}$ a root of f. Further, let $v_1, \dots, v_s$ be all extensions of v from K to $K(\alpha)$. Then there is some $\delta \in vK$ such that the following holds: If g is any irreducible monic polynomial over K satisfying $\operatorname{ins} g \geq \operatorname{ins} f$ and $v(f-g) > \delta$, and if $\beta \in \widetilde{K}$ is a root of g and $w_1, \dots, w_t$ are all extensions of v from K to $K(\beta)$, then $s = t$ and after a suitable renumbering of the w_i, we have:*

$$\begin{array}{rcl} \mathrm{d}(K(\alpha)|K, v_i) & = & \mathrm{d}(K(\beta)|K, w_i) \\ \mathrm{e}(K(\alpha)|K, v_i) & = & \mathrm{e}(K(\beta)|K, w_i) \\ \mathrm{f}\,(K(\alpha)|K, v_i) & = & \mathrm{f}\,(K(\beta)|K, w_i)\ . \end{array}$$

Proof. Observe that by Lemma 46 the extensions of v from K to $K(\alpha)$ are in correspondence with the double cosets as in Notation 38 with $L := K(\alpha)$ via $v_i := v\iota_i$. By virtue of Lemma 41 we can choose the indices of the irreducible factors $f_1, \dots, f_s$ of f over K^h in such a way that $\iota_i\alpha$ is a root of f_i. We do the same for g and its irreducible factors $g_1, \dots, g_t$, taking $L := K(\beta)$ and choosing the automorphisms $\iota'_1, \dots, \iota'_t$.

Take δ as in Theorem 42. Then the factors f_i and g_i satisfy the assertions of that theorem; in particular, we have $s = r = t$. After a suitable renumbering, we can assume that $\iota'_i\beta$ is a root of g_i if and only if $\iota_i\alpha$ is a root of f_i. By Lemma 47 we have:

$$\mathrm{f}\,(K(\alpha)|K, v_i) = \mathrm{f}\,(\iota_i K(\alpha).K^h|K^h, v) = [(\iota_i K(\alpha).K^h)v : K^h v].$$

By Theorem 42, $\iota_i K(\alpha).K^h = K^h(\iota_i\alpha)$ and $\iota'_i K(\beta).K^h = K^h(\iota'_i\beta)$ are isomorphic over K^h. Therefore, the degree above is equal to

$$[K^h(\iota_i\alpha)v : K^h v] = [K^h(\iota'_i\beta)v : K^h v] = \mathrm{f}\,(\iota'_i K(\beta).K^h|K^h, v),$$

which in turn is equal to $\mathrm{f}\,(K(\beta)|K, w_i)$ by Lemma 47. The equations for the inertia degree are analogous. The result for the defect then follows from these equations, together with (30) and (32). □

8. Appendix: Induction on the degree of the polynomial

In this section we present a theorem whose essential feature is that its proof employs induction on the degree of the polynomials. It serves as a demonstration of what can be achieved through this method. The bound for the value $v(f-g)$ will be larger than the ones in Section 5, which makes it a less optimal method than those already presented in the paper.

The method can be described as follows: first set $f_1 := f$, $g_1 := g$ and choose any root α_1 of f. Then use Lemma 20 to find a root β_1 of g such that $v(\alpha_1 - \beta_1) > \varepsilon$. Then set $f_2 := \frac{f_1}{x-\alpha_1}$ and $g_2 := \frac{g_1}{x-\beta_1}$ to repeat the procedure and continue the process until we arrive at linear polynomials.

To prove the main theorem, we first need the following lemma.

Lemma 49. *Take $f, g \in K[x]$, assume that f is monic and let α be a root of f. If $v(f-g) > 0$ and if β is a root of g such that $v(\alpha-\beta) \geq vf + \frac{v(f-g)}{n}$, then*

$$v\left(\frac{f(x)}{x-\alpha} - \frac{g(x)}{x-\beta}\right) \geq 2vf + \frac{v(f-g)}{n}.$$

Proof. Since $(x-\alpha)(x-\beta)$ is monic, we have $v((x-\alpha)(x-\beta)) \leq 0$. Therefore, we obtain:

$$\begin{aligned} v\left(\frac{f(x)}{x-\alpha} - \frac{g(x)}{x-\beta}\right) &= v(f(x)(x-\beta) - g(x)(x-\alpha)) - v((x-\alpha)(x-\beta)) \\ &\geq v(f(x)(x-\beta) - g(x)(x-\alpha)) \\ &\geq \min\left\{v\left((f(x)-g(x))x\right), v\left(f(x)\beta - g(x)\alpha\right)\right\} \\ &= \min\left\{v(f-g), v\left(f(x)\beta - g(x)\alpha\right)\right\}. \end{aligned}$$

We wish to find a lower bound for $v(f(x)\beta - g(x)\alpha)$. We use the assumption of the lemma and the facts that $vf \leq 0$ and $vf \leq v\alpha$ (since f is monic) to obtain:

$$\begin{aligned} v\left(f(x)\beta - g(x)\alpha\right) &= v\left(f(x)(\beta-\alpha) + (f(x)-g(x))\alpha\right) \\ &\geq \min\left\{vf + v(\beta-\alpha), v(f-g) + v\alpha\right\} \\ &\geq \min\left\{2vf + \frac{v(f-g)}{n}, v(f-g) + vf\right\} \\ &= 2vf + \frac{v(f-g)}{n}. \end{aligned}$$

Going back to the initial inequality, we obtain that

$$v\left(\frac{f(x)}{x-\alpha} - \frac{g(x)}{x-\beta}\right) \geq \min\left\{v(f-g), 2vf + \frac{v(f-g)}{n}\right\} = 2vf + \frac{v(f-g)}{n}.$$

□

Theorem 50. *Take polynomials f, g as in (1) and $\varepsilon > 0$. Assume that*

$$v(f-g) > n!\,\varepsilon - (n+1)!\,(vf - va_n) + va_n. \tag{35}$$

Then there is an enumeration of the roots of g such that $v(\alpha_i - \beta_i) \geq \varepsilon$ for $1 \leq i \leq n$.

Proof. Condition (35) can be written in a simpler way, with f replaced by the monic polynomial $a_n^{-1}f$:

$$v(a_n^{-1}f - a_n^{-1}g) > n!\varepsilon - (n+1)!(va_n^{-1}f).$$

Since f and $a_n^{-1}f$ have the same roots and the same is true for g and $a_n^{-1}g$, we may assume that f is monic. In this case, $va_n = 0$ and $vf \leq 0$.

We will proceed by reverse induction on the degree n of f as long as it is larger than 1. We set $f_1 := f$ and $g_1 := g$ and choose any root α_1 of f. We use Lemma 20 to find a root β_1 of g such that

$$\begin{aligned} v(\alpha_1-\beta_1) &\geq vf_1+\frac{v(f_1-g_1)}{n} > vf_1+\frac{n!\,\varepsilon-(n+1)!\,vf_1}{n} \\ &= vf+(n-1)!\,\varepsilon-(n+1)(n-1)!\,vf \\ &= (n-1)!\,\varepsilon-n!\,vf-((n-1)!\,-1)vf \geq \varepsilon, \end{aligned}$$

where we have used that $n-1\geq 1$ and $vf_1 = vf \leq 0$.

Now we assume that $i<n$ and that for $1\leq j\leq i$, we have already found roots α_j,β_j such that $v(\alpha_j-\beta_j)\geq\varepsilon$, and polynomials f_j,g_j such that

$$\deg f_j = \deg g_j = n-j+1,$$

as well as $vf_j \geq vf$ and

$$v(f_j-g_j)\geq (n-j+1)!\,\varepsilon-(n-j+2)!\,vf. \tag{36}$$

We define

$$f_{i+1} := \frac{f_i}{x-\alpha_i} \quad \text{and} \quad g_{i+1} := \frac{g_i}{x-\beta_i}.$$

Observe that

$$\deg f_{i+1} = \deg g_{i+1} = n-(i+1)+1$$

and that $vf_{i+1}\geq vf_i \geq vf$ because $v(x-\alpha_i)\leq 0$. By Lemma 49 we have:

$$\begin{aligned} v(f_{i+1}-g_{i+1}) &\geq 2vf_i+\frac{v(f_i-g_i)}{n-i+1} \\ &\geq 2vf+\frac{(n-i+1)!\,\varepsilon-(n-i+2)!\,vf}{n-i+1} \\ &= 2vf+(n-i)!\,\varepsilon-(n-i+2)(n-i)!\,vf \\ &= (n-i)!\,\varepsilon-(n-i+1)!\,vf-((n-i)!\,-2)vf\,. \end{aligned}$$

If $i+1<n$, then $((n-i)!\,-2)\geq 0$ and we obtain that (36) also holds for $j=i+1$. If $i+1=n$, then $-(n-i+1)!-((n-i)!-2)=-1$, thus

$$v(f_n-g_n)\geq \varepsilon - vf\,. \tag{37}$$

Assume that $i+1<n$. Then $\deg f_{i+1} = \deg g_{i+1} > 1$ and we have to continue our induction. We choose a root α_{i+1} of f_{i+1}. Then by Lemma 20 there is a root β_{i+1} of g_{i+1} such that

$$\begin{aligned} v(\alpha_{i+1}-\beta_{i+1}) &\geq vf_{i+1}+\frac{v(f_{i+1}-g_{i+1})}{n-i} \\ &\geq vf+(n-i-1)!\,\varepsilon-(n-i+1)(n-i-1)!\,vf\geq\varepsilon, \end{aligned}$$

where we use that $n-i-1\geq 1$. This completes our induction step.

Finally, we deal with the case of $i+1=n$. Then both f_{i+1} and g_{i+1} are linear polynomials, say, $x-\alpha$ and $b_n(x-\beta)$. We set $\alpha_n:=\alpha$. In view of

(37), Lemma 20 shows the existence of a root β_n of g_n, which consequently must be equal to β, such that

$$v(\alpha_n - \beta_n) \geq v f_n + v(f_n - g_n) \geq \varepsilon.$$

This completes the proof of our theorem. □

References

[1] Bosch, S., Güntzer, U., Remmert, R. (1984). *Non-Archimedean analysis.* Vol. 261 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin. ISBN:978-3-540-12546-4

[2] Brink, D. (2006). *New light on Hensel's lemma.* Expo. Math. 24(4):291-306. DOI: http://dx.doi.org/10.1016/j.exmath.2006.01.002

[3] Cohn, P. M. (2003). *Basic Algebra: Groups, Rings and Fields.* Springer-Verlag, London. DOI: http://dx.doi.org/10.1007/978-0-85729-428-9

[4] Endler, O. (1972). *Valuation theory.* Springer-Verlag, Berlin. DOI: http://dx.doi.org/10.1007/978-3-642-65505-0

[5] Engler, A. J., Prestel, A. (2005). *Valued fields.* Springer Monographs in Mathematics, Springer-Verlag, Berlin. DOI: https://doi.org/10.1007/3-540-30035-X

[6] Ershov, Yu. L. (2007). *A theorem on values of roots and coefficients.* Dokl. Math. 76(3):913-915. DOI: http://dx.doi.org/10.1134/S1064562407060269

[7] Ershov, Yu. L. (2008). *Root continuity theorems in valued fields. II.* Sib. Math. J. 49(5):852-856. DOI: http://dx.doi.org/10.1007/s11202-008-0084-y

[8] Kuhlmann, F.-V. *Valuation Theory*, book in preparation. Preliminary versions of several chapters are available on the website:
http://math.usask.ca/~fvk/Fvkbook.htm

[9] Kuhlmann, F.-V. (2010). *A classification of Artin–Schreier defect extensions and characterizations of defectless fields.* Illinois J. Math. 54(2):397-448. DOI: http://dx.doi.org/10.1215/ijm/1318598666

[10] Neukirch, J. (1999). *Algebraic Number Theory.* vol. 322 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag Berlin Heidelberg. DOI: http://dx.doi.org/10.1007/978-3-662-03983-0

[11] Ribenboim, P. (1968). *Théorie des valuations.* Les Presses de l'Université de Montréal

[12] Shell, N. (1990). *Topological fields and near valuations.* Vol. 135 of Monographs and Textbooks in Pure and Applied Mathematics, Marcel Dekker, Inc., New York. DOI: http://dx.doi.org/10.1201/9781003210177

[13] Warner, S. (1989). *Topological fields.* Vol. 157 of North-Holland Mathematics Studies, North-Holland Publishing Co., Amsterdam. ISBN:978-0-444-87429-0

Institute of Mathematics, University of Szczecin
ul. Wielkopolska 15, 70-451 Szczecin, Poland

Institut für Algebra, Technische Universität Dresden, Zellescher Weg 12-14, 01062 Dresden, Germany

Email address: hannacmielmath@gmail.com

Email address: fvk@usz.edu.pl

Email address: piotr.szewczyk@tu-dresden.de