LAJM - LATIN AMERICAN JOURNAL OF MATHEMATICS
Proceedings: I Encontro Paulista da Pós-Graduação em Matemáticas
EPPM

The equality between $\epsilon(f)$ and $\delta(f)$ proved via Newton polygons*

Caio Henrique Silva de Souza[1]

[1]Department of Mathematics – Universidade Federal de São Carlos
Brazil

caiohss@estudante.ufscar.br

Abstract. *In this paper, we reproduce the proof given in [1] of the equality between $\epsilon(f)$ and $\delta(f)$, two important objects in Valuation Theory. This proof uses the notion of Newton polygons. We present some details that were omitted in [1] and illustrate a step-by-step construction of a Newton Polygon associated to a specific finite set.*

Keywords – *Key polynomials, Newton polygons, MacLane-Vaquié key polynomials, abstract key polynomials.*
MSC2020 – *Primary 13A18*

1. Acknowledgments

During the realization of this project the author and his advisor were supported by grants from Fundação de Amparo à Pesquisa do Estado de São Paulo (process numbers 2017/17835-9 and 2020/05148-0).

2. Introduction

In Valuation Theory, an important notion is the concept of *key polynomial*. This object showed to be important in some programs that intend to proof local uniformization and resolution of singularities in positive characteristic. These are significant problems in Algebraic Geometry (see [9]). In 1936, Mac Lane started in [6] the study of key polynomials in order to understand all possible extensions of a valuation from $\mathbb{K}$ to $\mathbb{K}[x]$. Years latter, Vaquié introduced a generalization of Mac Lane's key polynomials in [12]. After that, Novacoski and Spivakovsky in [8] and Decaup, Mahboub and Spivakovsky in [3] introduced a new version of key polynomial. This new definition depends on the following object, that will be in the center of our discussion in this paper.

Let $\mathbb{K}[x]$ be the ring of polynomials on one indeterminate over the field $\mathbb{K}$. Fix a valuation ν on $\mathbb{K}[x]$ with value group Γ_ν and let $f \in \mathbb{K}[x]$ be a non-zero polynomial.

*Received June 10th, 2022; Revised March 23rd, 2023.

For every $i \in \mathbb{N}$, we consider $\partial_i f$ the formal Hasse-derivative of order i of f. This is the uniquely determined polynomial such that, for all $a \in \mathbb{K}$, we have that $\partial_i f(a)$ is the coefficient of the degree i monomial of the Taylor expansion of f around a. If $f \notin \operatorname{supp}(\nu)$ and $\deg(f) > 0$, then we define

$$\epsilon(f) := \max_{1 \leq i \leq \deg(f)} \left\{ \frac{\nu(f) - \nu(\partial_i f)}{i} \,\middle|\, \nu(\partial_i f) \in \Gamma_\nu \right\} \in \Gamma_\nu \otimes_{\mathbb{Z}} \mathbb{Q}.$$

A key polynomial is a monic polynomial Q that satisfies the following property: if $f \in \mathbb{K}[x]$ is such that $\deg(f) < \deg(Q)$, then $\epsilon(f) < \epsilon(Q)$.

The definition of $\epsilon(f)$ is not natural at first. However, it allows us to prove all of the initial results about key polynomials and truncations in an explicit way (see [8]).

In [7], Novacoski introduced the notion of $\delta(f)$, an object that is easier to visualize than $\epsilon(f)$. Take μ an extension of ν to $\overline{\mathbb{K}}[x]$, where $\overline{\mathbb{K}}$ is a fixed algebraic closure of $\mathbb{K}$. Given a non-zero polynomial f, if $\deg(f) > 0$, then we define

$$\delta(f) := \max\{\mu(x - c) \mid c \in \overline{\mathbb{K}} \text{ and } f(c) = 0\}.$$

This object can be used, for example, to see the relations between key polynomials, minimal pairs and truncations (see [7] or [10]).

Novacoski proved in [7] that $\epsilon(f)$ is equal to $\delta(f)$. His proof is purely algebraic. In [1], Bengus-Lasnier gives another proof of this equality, using Newton polygons. This new proof gives a geometric approach to Valuation Theory.

In this paper, we reproduce the proof of this equality given in [1], which is based on a result of [5]. We present some details that were omitted in [1]. Also, we illustrate a step-by-step construction of the Newton Polygon associated to a specific finite set.

In Section 3, we define the Newton Polygon of a set X, together with the notions of line, slope, convex hull, among others. We also give a step-by-step construction for the Newton Polygon of a finite set of the form

$$X = \{(i, \gamma_i) \in \mathbb{N} \times \Phi_{\mathbb{Q}} \mid 0 \leq i \leq m,\, m \in \mathbb{N}, \}.$$

where $\Phi_{\mathbb{Q}}$ is the divisible hull of a totally ordered abelian group Φ.

LAJM - LATIN AMERICAN JOURNAL OF MATHEMATICS
Proceedings: I Encontro Paulista da Pós-Graduação em Matemáticas
EPPM

In Section 4, we begin by defining a valuation on a commutative ring with unit. Take a valuation ν on a field $\mathbb{K}$ and a polynomial $g(x) = a_0 + \ldots + a_n x^n$, with $a_0 \neq 0$. We study the Newton Polygon associated to the finite set of the form

$$X = \{(i, \nu(a_i)) \in \mathbb{N} \times \Gamma_\nu \mid 0 \leq i \leq n \text{ and } a_i \neq 0\}.$$

Using an adaptation of a lemma from [5] (our Theorem 4.2), we prove Corollary 4.3 that relates the geometric aspects of the Newton Polygon to the roots of g. Then, we prove Corollary 4.5 that deals with the Newton polygon associated to the set

$$X = \{(i, \nu(\partial_i f)) \in \mathbb{N} \times \Gamma_\nu \mid 1 \leq i \leq n \text{ and } \nu(\partial_i f) \neq \infty\},$$

where f is a polynomial of degree n.

Finally, in Section 5, we prove the main result of this paper, that is, the equality $\epsilon(f) = \delta(f)$ (Theorem 5.1). We also conclude that $\delta(f)$ does not depend on the choice of the extension μ nor on the algebraic closure $\overline{\mathbb{K}}$.

3. Newton polygons

The presentation of Newton polygons that we chose for this paper relates to the one in [1], which is based on Vaquié's presentation in [11]. In the following, we define the main concepts that are necessary to construct a Newton polygon.

Let Φ be a totally ordered abelian group. We set $\Phi_{\mathbb{Q}} := \Phi \otimes_{\mathbb{Z}} \mathbb{Q}$, which is the divisible hull of Φ (see [4]). All elements in $\Phi_{\mathbb{Q}}$ can be reduced to a simple tensor $\phi \otimes q$, with $\phi \in \Phi$ and $q \in \mathbb{Q}$. Moreover, there is an injective map $\Phi \hookrightarrow \Phi_{\mathbb{Q}}$, mapping ϕ to $\phi \otimes 1$. We denote $\phi \otimes q$ by $\phi q = q\phi$ and, for $a, b \in \mathbb{Q}$, we denote $\frac{a}{b}\phi$ by $\frac{a\phi}{b}$. In $\Phi_{\mathbb{Q}}$, we have a natural order induced from Φ (given by $\frac{a_1\phi_1}{b_1} \leq \frac{a_2\phi_2}{b_2} \Leftrightarrow a_1 b_2 \phi_1 \leq a_2 b_1 \phi_2$ in Φ).

Definition 3.1. *Take $q \in \mathbb{Q}$ and $\alpha, \beta \in \Phi_{\mathbb{Q}}$. A **line** $L \subseteq \mathbb{Q} \times \Phi_{\mathbb{Q}}$ is a subset of the form*

$$L = L_{q,\alpha,\beta} := \{(x, \phi) \in \mathbb{Q} \times \Phi_{\mathbb{Q}} \mid q\phi + \alpha x + \beta = 0\}.$$

*When $q \neq 0$, we call $s(L) := -\frac{\alpha}{q}$ the **slope** of L.*

Given distinct points $P_1 = (x_1, \phi_1)$ and $P_2 = (x_2, \phi_2)$ in $\mathbb{Q} \times \Phi_{\mathbb{Q}}$, there exists a line $L = L_{q,\alpha,\beta}$ containing this points (take $q = x_2 - x_1$, $\alpha = \phi_1 - \phi_2$ and $\beta = x_1\phi_2 - x_2\phi_1$).

In this situation, we denote L by $L_{P_1P_2}$. Moreover, one can prove that

$$s(L_{P_1P_2}) = -\frac{\alpha}{q} = \frac{\phi_2 - \phi_1}{x_2 - x_1} = \frac{\phi_1 - \phi_2}{x_1 - x_2}.$$

Let $m_x := \min\{x_1, x_2\}$, $M_x := \max\{x_1, x_2\}$, $m_\phi := \min\{\phi_1, \phi_2\}$ and $M_\phi := \max\{\phi_1, \phi_2\}$. The **segment** defined by P_1 and P_2 is the subset

$$\overline{P_1P_2} := \{(x', \phi') \in L_{P_1P_2} \mid m_x \leq x' \leq M_x \text{ and } m_\phi \leq \phi' \leq M_\phi\}.$$

For each line $L = L_{q,\alpha,\beta}$, we define the **half-spaces**

$$H^L_{\geq} := \{(x, \phi) \in \mathbb{Q} \times \Phi_{\mathbb{Q}} \mid q\phi + \alpha x + \beta \geq 0\}$$

and

$$H^L_{\leq} := \{(x, \phi) \in \mathbb{Q} \times \Phi_{\mathbb{Q}} \mid q\phi + \alpha x + \beta \leq 0\}.$$

Definition 3.2. *Given a subset $A \subseteq \mathbb{Q} \times \Phi_{\mathbb{Q}}$, the **convex hull** of A is the intersection of all half-spaces containing A, that is,*

$$\text{Conv}(A) := \bigcap_{\substack{H \text{ is a half-space} \\ A \subseteq H}} H.$$

A **face** of A is a subset $F = \text{Conv}(A) \cap L$, where $L \subset \mathbb{Q} \times \Phi_{\mathbb{Q}}$ is a line such that F contains at least two points and

$$\text{Conv}(A) \subset H^L_{\geq} \quad \text{or} \quad \text{Conv}(A) \subset H^L_{\leq}.$$

Definition 3.3. *For $X \subseteq \mathbb{Q} \times \Phi_{\mathbb{Q}}$, the **Newton polygon** associated to X is given by*

$$\text{PN}(X) := \text{Conv}(\{(x, \phi + \delta) \mid (x, \phi) \in X,\ \delta \in \Phi_{\mathbb{Q}} \text{ and } \delta \geq 0\}).$$

In this paper, we focus on Newton polygons given by a particular kind of set. Namely, we consider the case where X is a finite subset of $\mathbb{Q} \times \Phi_{\mathbb{Q}}$ of the form

$$X = \{(i, \gamma_i) \in \mathbb{N} \times \Phi_{\mathbb{Q}} \mid 0 \leq i \leq m, m \in \mathbb{N}\}.$$

An example of a subset X is illustrated in Figure 1. Let us call $P_i = (i, \gamma_i)$, $0 \leq i \leq m$. Take $PN(X)$ the Newton polygon associated to X. In the following, we present a geometric interpretation of $PN(X)$.

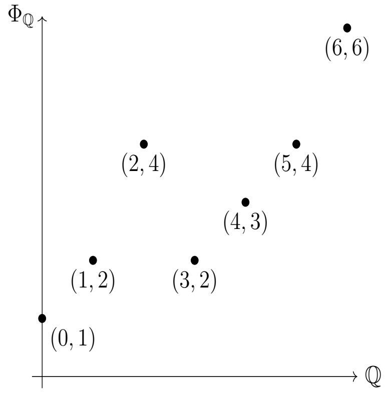

Figure 1. An example of subset X, for $\Phi = \mathbb{Z}$.

- We begin by taking the pair $P_0 = (0, \gamma_0)$ and defining $i_1 = 0$. Consider

$$S_{i_1} = \{L_{P_0P_i} \mid 1 \leq i \leq m\}.$$

Let P_{i_2} be such that $L_{P_0P_{i_2}}$ has the least slope among the lines in S_0, where i_2 is the greatest index among the ones for which the least slope is achieved. Take the segment $\overline{P_{i_1}P_{i_2}}$. We can see this first step in Figure 2.

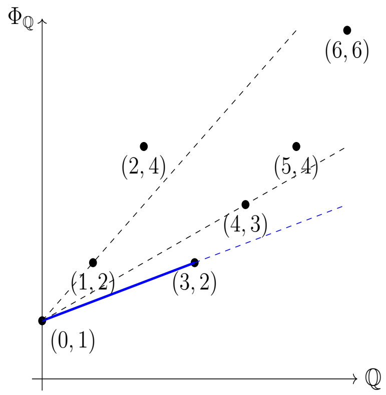

Figure 2. First step of the construction of the Newton polygon of X.

- For the second step, consider

$$S_{i_2} = \{L_{P_{i_2}P_i} \mid i_2 + 1 \leq i \leq m\}.$$

Let P_{i_3} be such that $L_{P_{i_2}P_{i_3}}$ has the least slope among the lines in S_{i_2}, where i_3 is the greatest index among the ones for which the least slope is achieved. Take the segment $\overline{P_{i_2}P_{i_3}}$. We can see this second step in Figure 3.

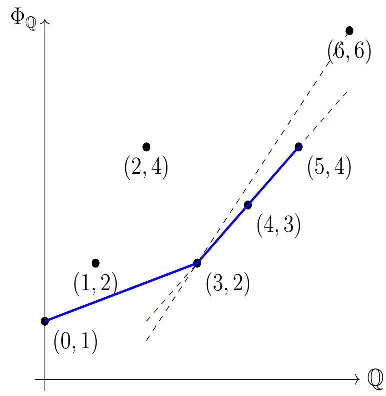

Figure 3. Second step of the construction of the Newton polygon of X.

- We repeat the above steps until we reach P_m. Let $i_1, i_2, i_3, \ldots, i_{k+1}$ be the highlighted indexes of the process, where $i_1 = 0$ and $i_{k+1} = m$. Note that this process gives us $i_1 < i_2 < \ldots < i_{k+1}$ and $\gamma_{i_1} < \gamma_{i_2} < \ldots < \gamma_{i_{k+1}}$.

- Take the segments $\overline{P_{i_l}P_{i_{l+1}}}$, with $1 \leq l \leq k$, and the subsets $\{(i, \gamma_i + \delta) \mid \delta \geq 0\}$ for $i = 0$ and $i = m$. We define $P \subset \mathbb{Q} \times \Phi_{\mathbb{Q}}$ by

$$P := \left(\bigcap_{l=1}^{k} H_{\geq}^{L_{P_{i_l}P_{i_{l+1}}}} \right) \cap H^{\gamma_0} \cap H^{\gamma_m},$$

where $H^{\gamma_0} = \{(x, \phi) \mid x \geq 0\}$ and $H^{\gamma_m} = \{(x, \phi) \mid x \leq m\}$. We illustrate the region P in Figure 4.

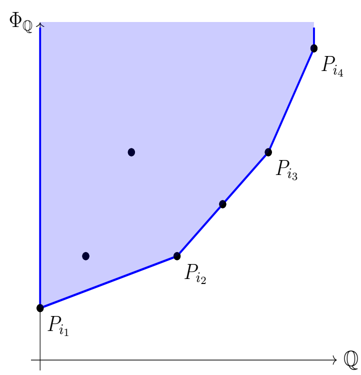

Figure 4. Region P.

We will prove that $P = PN(X)$. Let

$$Y = \{(x, \phi + \delta) \mid (x, \phi) \in X,\ \delta \in \Phi_{\mathbb{Q}} \text{ e } \delta \geq 0\},$$

so $PN(X) = \text{Conv}(Y)$. We initially see that $Y \subset P$. Indeed, take $(x, \phi) \in Y$. We have two options: either for some i, $0 \le i \le m$, we have $(x,\phi) = (i,\gamma_i)$ or $(x,\phi) = (i,\gamma_i+\delta)$ for some $\delta > 0$.

Let us suppose the first case, where $(x,\phi) = (i,\gamma_i)$. Consider the indexes $i_1,\dots,i_{k+1}$ from the above construction. To simplify the notation, we name $L_{P_{i_l}P_{i_{l+1}}} = L_{l,l+1}$. Let us show that $(i,\gamma_i) \in H_{\ge}^{L_{l,l+1}}$ for all l, $1 \le l \le k$. If $i = i_l$ for some l, then it is immediate that $(i,\gamma_i) \in H_{\ge}^{L_{l,l+1}}$ for that l. Suppose $i \ne i_l$ for all l. By the property that defines i_l and i_{l+1}, the slope of the line through P_{i_l} and (i,γ_i) is greater than or equal to the slope of the line through P_{i_l} and $P_{i_{l+1}}$. That is,

$$\frac{\gamma_i - \gamma_{i_l}}{i - i_l} \ge \frac{\gamma_{i_{l+1}} - \gamma_{i_l}}{i_{l+1} - i_l}.$$

Above, we can assume without lost of generality that $i > i_l$, since the case $i < i_l$ lead us to the same inequality. Manipulation of this inequalities lead us to

$$(i_{l+1} - i_l)(\gamma_i - \gamma_{i_l}) \ge (i - i_l)(\gamma_{i_{l+1}} - \gamma_{i_l})$$

if and only if

$$(i_{l+1} - i_l)(\gamma_i - \gamma_{i_l}) \ge i(\gamma_{i_{l+1}} - \gamma_{i_l}) - i_l(\gamma_{i_{l+1}} - \gamma_{i_l})$$

if and only if

$$(i_{l+1} - i_l)(\gamma_i - \gamma_{i_l}) + i_l(\gamma_{i_{l+1}} - \gamma_{i_l}) \ge i(\gamma_{i_{l+1}} - \gamma_{i_l})$$

if and only if

$$-(i_{l+1} - i_l)(\gamma_i - \gamma_{i_l}) - i_l(\gamma_{i_{l+1}} - \gamma_{i_l}) \le i(\gamma_{i_l} - \gamma_{i_{l+1}}).$$

Then, since $\alpha = \gamma_l - \gamma_{l+1}$, $q = i_{l+1} - i_l$ and $\beta = i_l\gamma_{i_{l+1}} - \gamma_{i_l} i_{l+1}$, we have

$$\begin{aligned}(i_{l+1} - i_l)\gamma_i + i(\gamma_l - \gamma_{l+1}) + \beta &\ge (i_{l+1} - i_l)\gamma_i - (i_{l+1} - i_l)(\gamma_i - \gamma_{i_l}) - i_l(\gamma_{i_{l+1}} - \gamma_{i_l}) + \beta \\ &= \gamma_{i_l}(i_{l+1} - i_l) - i_l(\gamma_{i_{l+1}} - \gamma_{i_l}) + \beta \\ &= \gamma_{i_l} i_{l+1} - i_l\gamma_{i_{l+1}} + \beta = 0.\end{aligned}$$

That is, $(i,\gamma_i) \in H_{\ge}^{L_{l,l+1}}$.

Now suppose the second case, where $(x,\phi) = (i,\gamma_i+\delta)$ with $\delta > 0$. Then,

$$(i_{l+1} - i_l)(\gamma_i + \delta) + i(\gamma_l - \gamma_{l+1}) + \beta = (i_{l+1} - i_l)\delta + (i_{l+1} - i_l)\gamma_i + i(\gamma_l - \gamma_{l+1}) + \beta \ge 0,$$

since $(i_{l+1} - i_l)\delta \ge 0$ and $(i_{l+1} - i_l)\gamma_i + i(\gamma_l - \gamma_{l+1}) + \beta \ge 0$. Hence, $(i,\gamma_i+\delta) \in H_{\ge}^{L_{l,l+1}}$.

Moreover, for every i, (i,γ_i) and $(i,\gamma_i+\delta)$ belong to H^{γ_0} and H^{γ_m}. Thus, we see that $Y \subset H_{\geq}^{L_{l,l+1}}$ for any l and $Y \subset H^{\gamma_0} \cap H^{\gamma_m}$. Therefore,

$$Y \subset P = \left(\bigcap_{l=1}^{k} H_{\geq}^{L_{l,l+1}} \right) \cap H^{\gamma_0} \cap H^{\gamma_m}.$$

Now we check $PN(X) = P$.

- $PN(X) \subseteq P$: by the definition of $PN(X)$, we have $PN(X) \subset H$ for every half-space H that contains Y. Hence, we consider the half-spaces $H_{\geq}^{L_{l,l+1}}$ for every l, $1 \leq l \leq k$, and the half-spaces H^{γ_0} and H^{γ_m}. By what we did above, $Y \subset H_{\geq}^{L_{l,l+1}}$ for every l, $1 \leq l \leq k$, and $Y \subset H^{\gamma_0} \cap H^{\gamma_m}$. Therefore, $PN(X) \subset H_{\geq}^{L_{l,l+1}}$ for every l and $PN(X) \subset H^{\gamma_0} \cap H^{\gamma_m}$. Thus, $PN(X) \subseteq P$.
- $P \subseteq PN(X)$: Take $H = H_{\geq}^{L}$ a half-space determined by a line $L = L_{\alpha,q,\beta}$ such that $Y \subset H$. We show that $P \subset H$. Take $(x, \phi) \in P$. If $(x, \phi) \in Y$, then $(x, \phi) \in H$.

  Suppose that (x, ϕ) belongs to some segment $\overline{P_{i_l} P_{i_{l+1}}}$. Then, $m_x \leq x \leq M_x$ and $m_\phi \leq \phi \leq M_\phi$, with $m_x = \min\{i_l, i_{l+1}\} = i_l$, $M_x = \max\{i_l, i_{l+1}\} = i_{l+1}$, $m_\phi = \min\{\gamma_{i_l}, \gamma_{i_{l+1}}\} = \gamma_{i_l}$ e $M_\phi = \max\{\gamma_{i_l}, \gamma_{i_{l+1}}\} = \gamma_{i_{l+1}}$. Then, since $(m_x, m_\phi) = (i_l, \gamma_{i_l}) \in Y \subset H$,

  $$q\phi + \alpha x + \beta \geq q m_\phi + \alpha m_x + \beta \geq 0,$$

  hence $(x, \phi) \in H$.

  Suppose that $(x, \phi) \in P$ do not satisfies the preceding cases. Considering the indexes $i_1, i_2, \ldots, i_{k+1}$, since they are distinct, $i_1 = 0$ and $i_{k+1} = m$, we have that they form a partition of the interval $[0, m]$. Hence, there exists l, $1 \leq l \leq k$, such that $i_l \leq x \leq i_{l+1}$. Take in the segment $\overline{P_{i_l} P_{i_{l+1}}}$ a point (x, ϕ'), $m_\phi \leq \phi' \leq M_\phi$. Since (x, ϕ) do not belong to any segment, we have $\phi > \phi'$. Hence,

  $$q\phi + \alpha x + \beta > q\phi' + \alpha x + \beta \geq 0,$$

  since (x, ϕ') belongs to a segment, thus by the preceding case it belongs to H. Hence, $P \subset H$. Since H is an arbitrary half-space that contains Y, we conclude that $P \subseteq PN(X)$.

The points P_{i_l}, with $1 \leq l \leq k+1$, are called the **vertices** of the polygon. The segments $\overline{P_{i_l} P_{i_{l+1}}}$ are the faces of $PN(X)$. The slope of a face $\overline{P_{i_l} P_{i_{l+1}}}$ is the slope of the line $L_{P_{i_l}P_{i_{l+1}}}$. We will denote this slope by

$$\delta_l = \frac{\alpha_l}{q_l} = \frac{\gamma_{i_{l+1}} - \gamma_{i_l}}{i_{l+1} - i_l}, \text{ where } 1 \leq l \leq k.$$

We call $q_l = i_{l+1} - i_l$ the **length** of the slope δ_l.

4. Valuations and Newton polygons

In this section, we explore the Newton polygon of a finite set that will be defined by a fixed polynomial and a given valuation.

Given a totally ordered abelian group Γ, we extent it to the structure $\Gamma_\infty := \Gamma \cup \{\infty\}$. The extension of addition and order from Γ to Γ_∞ is done in the natural way.

Definition 4.1. *Take a commutative ring R with unity. A* ***valuation*** *on R is a mapping $\nu : R \longrightarrow \Gamma_\infty$, where Γ is a totally ordered abelian group, with the following properties.*

(V1) $\nu(ab) = \nu(a) + \nu(b)$ *for all* $a, b \in R$.

(V2) $\nu(a+b) \geq \min\{\nu(a), \nu(b)\}$ *for all* $a, b \in R$.

(V3) $\nu(1) = 0$ *and* $\nu(0) = \infty$.

Let $\nu : R \longrightarrow \Gamma_\infty$ be a valuation. The set $\text{supp}(\nu) = \{a \in R \mid \nu(a) = \infty\}$ is called the **support** of ν. The **value group** of ν is the subgroup of Γ generated by $\{\nu(a) \mid a \in R \setminus \text{supp}(\nu)\}$ and is denoted by Γ_ν.

Let $\mathbb{K}$ be a field and ν be a valuation on $\mathbb{K}$. We fix an algebraic closure $\overline{\mathbb{K}}$ of $\mathbb{K}$. Take μ a valuation that extends ν to $\overline{\mathbb{K}}$. We consider Γ_ν and Γ_μ to be the values groups of ν and μ, respectively. We know that $\Gamma_\nu \subseteq \Gamma_\mu$. Consider $\Phi_{\mathbb{Q}} = \Gamma_\mu \otimes_{\mathbb{Z}} \mathbb{Q}$. We see that this group contains Γ_ν, Γ_μ and $\Gamma_\nu \otimes_{\mathbb{Z}} \mathbb{Q}$. Consider $g(x) \in \mathbb{K}[x]$ with non-vanishing roots and such that $g(0) = 1$. Then we can write

$$g(x) = \sum_{i=0}^{n} a_i x^i = \prod_{i=1}^{n} \left(1 - \frac{x}{c_i}\right) \in \mathbb{K}[x] \tag{1}$$

such that $a_0 = 1$ and $c_1, \ldots, c_n \in \overline{\mathbb{K}}$ are the roots of g, listed with possible repetitions. We have $c_i \neq 0$ for all i, where $1 \leq i \leq n$. Take $\lambda_i = \mu(1/c_i)$. We reorganize the indexes i of the roots $c_1, \ldots, c_n$ such that $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$.

Let

$$X = \{(i, \nu(a_i)) \mid 0 \leq i \leq n \text{ and } \nu(a_i) \neq \infty\} \subset \mathbb{Q} \times \Phi_{\mathbb{Q}}.$$

Consider the Newton polygon $PN(X)$, together with the slopes δ_l and lengths q_l, where $1 \leq l \leq k$, as defined in Section 3. For a set S, we denote by $|S|$ the cardinality of S. The next theorem is an adaptation of a lemma proved in [5] (Lemma 4, p. 90).

Theorem 4.2. *The values $\lambda_i = \mu(1/c_i)$ are slopes for $PN(X)$. Moreover, for each slope δ_l of $PN(X)$, where $1 \leq l \leq k+1$, we have*

$$|\{i \mid \lambda_i = \delta_l\}| = q_l$$

and

$$\delta_1 < \delta_2 < \ldots < \delta_k.$$

Proof. Suppose $\lambda_1 = \lambda_2 = \ldots = \lambda_{r_1} < \lambda_{r_1+1}$. We will show initially that the first segment in $PN(X)$, $\overline{P_{i_1}P_{i_2}}$, is the segment $\overline{P_0P_{r_1}}$, with $P_0 = (0,0)$ and $P_{r_1} = (r_1, r_1\lambda_1)$.

For each j, $1 \leq j \leq n$, We deduce from Equation (1) that

$$a_j = (-1)^j \left[\sum_{1 \leq i_1 < i_2 < \ldots < i_j \leq n} \left(\prod_{t=1}^{j} \frac{1}{c_{i_t}} \right) \right].$$

Calculating the value of a_j, we see that

$$\begin{aligned} \nu(a_j) = \mu(a_j) &= j\mu(-1) + \mu \left(\sum_{1 \leq i_1 < i_2 < \ldots < i_j \leq n} \left(\prod_{t=1}^{j} \frac{1}{c_{i_t}} \right) \right) \\ &\geq \min_{1 \leq i_1 < i_2 < \ldots < i_j \leq n} \left\{ \mu \left(\prod_{t=1}^{j} \frac{1}{c_{i_t}} \right) \right\}. \end{aligned}$$

However,

$$\mu \left(\prod_{t=1}^{j} \frac{1}{c_{i_t}} \right) = \sum_{t=1}^{j} \mu \left(\frac{1}{c_{i_t}} \right) = \sum_{t=1}^{j} \lambda_{i_t} \geq j\lambda_1.$$

for any choice of $1 \leq i_1 < i_2 < \ldots < i_j \leq n$. Hence, $\nu(a_j) \geq j\lambda_1$. Therefore,

$$\frac{\nu(a_j) - \nu(a_0)}{j - 0} = \frac{\nu(a_j)}{j} \geq \frac{j\lambda_1}{j} = \lambda_1 \text{ for every } j,\ 1 \leq j \leq n.$$

That is, the slope of any line in the set $S_{i_1} = \{L_{P_0P_j} \mid 1 \leq j \leq n\}$ is greater or equal to λ_1, i.e., $\delta_1 \geq \lambda_1$.

Now we look at a_{r_1}. In the expression

$$a_{r_1} = (-1)^{r_1} \left[\sum_{1 \leq i_1 < i_2 < \ldots < i_{r_1} \leq n} \left(\prod_{t=1}^{r_1} \frac{1}{c_{i_t}} \right) \right], \tag{2}$$

we have

$$\mu\left(\frac{1}{c_1 \cdots c_{r_1}}\right) = \sum_{j=1}^{r_1} \mu\left(\frac{1}{c_j}\right) = \sum_{j=1}^{r_1} \lambda_1 = r_1\lambda_1.$$

More than that, this is the only summand present in Equation (2) that has the value $r_1\lambda_1$, since any other product that is a summand in Equation (2) uses at least one of the indexes $r_1 + 1, \ldots, n$. Therefore, when we take the value of this product, we achieve a sum in which appears at least one of the $\lambda_{r_1+1}, \ldots, \lambda_n$, which are all bigger than λ_1. For instance,

$$\mu\left(\frac{1}{c_1 \cdots c_{r_1-1}c_{r_1+1}}\right) = \sum_{j=1}^{r_1-1} \mu\left(\frac{1}{c_j}\right) + \mu\left(\frac{1}{c_{r_1+1}}\right) = \sum_{j=1}^{r_1-1} \lambda_1 + \lambda_{r_1+1} > r_1\lambda_1.$$

Hence, any other product, which is a summand in (2), has value strictly greater than $r_1\lambda_1$, i.e.,

$$\mu\left(\sum_{\substack{1 \leq i_1 < i_2 < \ldots < i_{r_1} \leq n \\ \exists i_t > r_1}} \left(\prod_{t=1}^{r_1} \frac{1}{c_{i_t}} \right) \right) > r_1\lambda_1 = \mu\left(\frac{1}{c_1 \cdots c_{r_1}}\right).$$

Therefore, since

$$a_{r_1} = (-1)^{r_1} \left[\frac{1}{c_1 \cdots c_{r_1}} + \sum_{\substack{1 \leq i_1 < i_2 < \ldots < i_{r_1} \leq n \\ \exists i_t > r_1}} \left(\prod_{t=1}^{r_1} \frac{1}{c_{i_t}} \right) \right],$$

we have

$$\nu(a_{r_1}) = \mu(a_{r_1}) = \mu\left(\frac{1}{c_1 \cdots c_{r_1}}\right) = r_1\lambda_1.$$

By looking at the slope of $L_{P_0P_{r_1}}$ we obtain

$$\frac{\nu(a_{r_1})}{r_1} = \frac{r_1\lambda_1}{r_1} = \lambda_1.$$

Now we take $j > r_1$. In the expression of a_j each product consist of j factors $\frac{1}{c_{j_t}}$. Hence, since $j > r_1$, each product has at least one factor among $\frac{1}{c_{r_1+1}}, \ldots \frac{1}{c_n}$. Thus, all products in the sum that defines a_j have value strictly greater than $j\lambda_1$. Hence, the slope of $L_{P_0P_j}$ is

$$\frac{\nu(a_j)}{j} > \frac{j\lambda_1}{j} = \lambda_1.$$

We saw that all the slopes of the lines in S_{i_1} are greater or equal to λ_1, that $L_{P_0P_{r_1}}$ has slope equal to λ_1 and r_1 is the biggest index with such slope, since for $i > r_1$ we have slope strictly greater than λ_1. Hence, the segment $\overline{P_0P_{r_1}}$ is the first face of $PN(X)$, with slope $\delta_1 = \lambda_1$ and $q_1 = r_1$. It also follows that

$$|\{j \mid \lambda_j = \delta_1\}| = |\{1, 2, \ldots, r_1\}| = r_1 = q_1.$$

Now suppose $\lambda_{r_1} < \lambda_{r_1+1} = \lambda_{r_1+2} = \ldots = \lambda_{r_1+r_2} < \lambda_{r_1+r_2+1}$. We repeat the same reasoning above to prove that $\overline{P_{i_2}P_{i_3}} = \overline{P_{r_1}P_{r_1+r_2}}$, with $P_{r_1} = (r_1, r_1\lambda_1)$ and $P_{r_1+r_2} = (r_1 + r_2, r_1\lambda_1 + r_2\lambda_{r_1+1})$.

Take $j > r_1$. We have

$$\begin{aligned} \mu\left(\prod_{t=1}^{j} \frac{1}{c_{i_t}}\right) &= \sum_{t=1}^{j} \mu\left(\frac{1}{c_{i_t}}\right) \\ &= \sum_{t=1}^{j} \lambda_{i_t} \geq r_1\lambda_1 + (j - r_1)\lambda_{r_1+1}, \end{aligned}$$

for any choice of $1 \leq j_1 < j_2 < \ldots < j_{p_j} \leq n$. Hence,

$$\frac{\nu(a_j) - \nu(a_{r_1})}{j - r_1} = \frac{\nu(a_j) - r_1\lambda_1}{j - r_1} \geq \frac{r_1\lambda_1 + (j - r_1)\lambda_{r_1+1} - r_1\lambda_1}{j - r_1} = \lambda_{r_1+1}.$$

That is, $\delta_2 \geq \lambda_{r_1+1}$. For $r_1 + r_2$, we have

$$\mu\left(\frac{1}{c_1 \cdots c_{r_1+r_2}}\right) = \sum_{j=1}^{r_1+r_2} \mu\left(\frac{1}{c_j}\right) = \sum_{j=1}^{r_1} \lambda_1 + \sum_{j=r_1+1}^{r_1+r_2} \lambda_{r_1+1} = r_1\lambda_1 + r_2\lambda_{r_1+1}$$

and this is the only summand in the expression of $a_{r_1+r_2}$ with such value. Moreover, the other summands have value strictly greater than $r_1\lambda_1 + r_2\lambda_{r_1+1}$. Then, $\nu(a_{r_1+r_2}) = r_1\lambda_1 + r_2\lambda_{r_1+1}$ and hence

$$\frac{\nu(a_{r_1+r_2}) - \nu(a_{r_1})}{r_1 + r_2 - r_1} = \frac{r_1\lambda_1 + r_2\lambda_{r_1+1} - r_1\lambda_1}{r_2} = \lambda_{r_1+1}.$$

For $j > r_1 + r_2$, by the same reasoning above, it follows that $\nu(a_j) > r_1\lambda_1 + r_2\lambda_{r_1+1}$, implying that the slope of $L_{P_{r_1}P_j}$ is strictly greater than λ_{r_1+1}.

Therefore, the second face of $PN(X)$ is $\overline{P_{r_1}P_{r_1+r_2}}$, with slope $\delta_2 = \lambda_{r_1+1}$ and $q_2 = r_2$. More than that,

$$|\{j \mid \lambda_j = \delta_2\}| = |\{r_1+1, r_1+2, \ldots, r_1+r_2\}| = r_2 = q_2.$$

In general, for some $m \in \mathbb{N}$ we must have

$$\begin{aligned}\lambda_1 = \lambda_2 = \ldots = \lambda_{r_1} &< \lambda_{r_1+1} = \lambda_{r_1+2} = \ldots = \lambda_{r_1+r_2} \\ &< \lambda_{r_1+r_2+1} = \lambda_{r_1+r_2+2} = \ldots = \lambda_{r_1+r_2+r_3} \\ &\vdots \\ &< \lambda_{r_1+r_2+\ldots+r_{m-1}+1} = \ldots = \lambda_{r_1+r_2+\ldots+r_m} = \lambda_n.\end{aligned}$$

If $s = r_1 + r_2 + \ldots + r_t$, then

$$\lambda_s < \lambda_{s+1} = \ldots = \lambda_{s+r_{t+1}} < \lambda_{s+r_{t+1}+1}$$

and the above construction tells us that the segment from

$$P_s = (s, r_1\lambda_1 + r_2\lambda_{r_1+1} + \ldots + r_t\lambda_{s-r_t+1})$$

to

$$P_{s+r_{t+1}} = (s + r_{t+1}, r_1\lambda_1 + r_2\lambda_{r_1+1} + \ldots + r_t\lambda_{s-r_t+1} + r_{t+1}\lambda_{s+1})$$

will be a face of the Newton polygon, with slope $\delta_l = \lambda_{s+1}$ and length $q_l = r_{t+1}$ for a certain l, satisfying

$$|\{j \mid \lambda_j = \delta_l\}| = r_{t+1} = q_l.$$

Since at some moment $s + r_{t+1} = n$, we will pass through all the faces of $PN(X)$ and obtain the result. □

The next corollary deals with the case where a_0 is not necessarily equal to 1. Take $g(x) \in \mathbb{K}[x]$ with non-vanishing roots. We write

$$g(x) = \sum_{i=0}^{n} a_i x^i \in \mathbb{K}[x]$$

such that $a_0 \neq 0$. Consider the Newton polygon associated to

$$X = \{(i, \nu(a_i)) \mid 0 \leq i \leq n \text{ and } \nu(a_i) \neq \infty\}.$$

Recall that k is the number of vertices in $PN(X)$.

Corollary 4.3. *For each l, where $1 \leq l \leq k$, there exists a root c of g such that $\mu(c) = -\delta_l$ and its multiplicity is at most q_l. Moreover, each root c of g is associated to a slope δ_l such that $\mu(c) = -\delta_l$.*

Proof. Since $a_0 \neq 0$, if we divide g by a_0, then we do not change its roots. We will apply Theorem 4.2 for $g' = \frac{1}{a_0}g$. Thus, although the vertices of the Newton polygons associated to g and g' are different, they have the same slopes and lengths. Let $c_1, \ldots, c_n$ be the roots of g. We define

$$\lambda_i = \mu\left(\frac{1}{c_i}\right)$$

for each i, where $1 \leq i \leq n$. Thus, for all l, where $1 \leq l \leq k$, there exist q_l indexes i such that $\lambda_i = \delta_l$. That is, there exists at least one root $c = c_j$, where $1 \leq j \leq n$, such that

$$\mu\left(\frac{1}{c}\right) = \delta_l,$$

which implies $\mu(c) = -\delta_l$. Moreover, there exist at most q_l roots c_i equal to c. Now, take $c = c_j$ any root of g, where $1 \leq j \leq n$. Then, by Theorem 4.2, $\lambda_j = \lambda_{r_1+r_2+\ldots+r_l}$ for some l, where $1 \leq l \leq k$. Hence, by the same proposition, $\lambda_j = \delta_l$. Also, $\mu(c) = -\delta_l$. □

Remark 4.4. As a consequence of the above results, we have that $\Gamma_\mu \otimes_{\mathbb{Z}} \mathbb{Q} \cong \Gamma_\nu \otimes_{\mathbb{Z}} \mathbb{Q}$. Indeed, for any $c \in \overline{\mathbb{K}}$, take $g(x) \in \mathbb{K}[x]$ its minimal polynomial over $\mathbb{K}$. Since $g(x)$ is irreducible, we must have $g(0) \neq 0$. The above corollary implies that $\mu(c)$ is a slope of the Newton polygon. That is, $\mu(c) \in \Gamma_\nu \otimes_{\mathbb{Z}} \mathbb{Q}$. Hence,

$$\Gamma_\mu \subseteq \Gamma_\nu \otimes_{\mathbb{Z}} \mathbb{Q} \subseteq \Gamma_\mu \otimes_{\mathbb{Z}} \mathbb{Q}$$

and then

$$\Gamma_\mu \otimes_{\mathbb{Z}} \mathbb{Q} \subseteq \Gamma_\nu \otimes_{\mathbb{Z}} \mathbb{Q} \subseteq \Gamma_\mu \otimes_{\mathbb{Z}} \mathbb{Q}.$$

Thus, $\Gamma_\mu \otimes_{\mathbb{Z}} \mathbb{Q} \cong \Gamma_\nu \otimes_{\mathbb{Z}} \mathbb{Q}$. More than that, Γ_μ is a divisible group. Indeed, given $\gamma = \mu(c)$ and $d \in \mathbb{N}$, since $\overline{\mathbb{K}}$ is algebraically closed, there exists $b \in \overline{\mathbb{K}}$ such that $c = b^d$, hence $\mu(c) = d\mu(b)$. Therefore, $\Gamma_\mu \cong \Gamma_\mu \otimes_{\mathbb{Z}} \mathbb{Q} \cong \Gamma_\nu \otimes_{\mathbb{Z}} \mathbb{Q}$.

Let $\mathbb{K}(x)$ be the field of rational functions on one indeterminate over the field $\mathbb{K}$. Assume that $\mathbb{K}$ is an algebraically closed field. Let μ be a valuation on $\mathbb{K}(x)$ and fix $f \in \mathbb{K}[x]$ a non zero polynomial with degree n. For every $i \in \mathbb{N}$, we consider $\partial_i f$ the formal Hasse-derivative of order i of f. That is, $\partial_1 f, \ldots, \partial_n f$ are the uniquely determined polynomials for which the Taylor expansion

$$f(x) - f(a) = \sum_{i=1}^{n} \partial_i f(a)(x-a)^i$$

is satisfied for every $a \in \mathbb{K}$. Let

$$X = \{(i, \mu(\partial_i f)) \mid 0 \leq i \leq n \text{ and } \mu(\partial_i f) \neq \infty\}.$$

Consider $PN(X)$ to be the Newton polygon associated to X, together with the slopes δ_l and lengths q_l, where $1 \leq l \leq k$. The presentation of the next corollary is due F.-V. Kuhlmann and Hanna Ćmiel. For more relations between roots of polynomials and slopes of Newton polygons, we recommend the work [2] of the mentioned authors.

Corollary 4.5. *For each l, where $1 \leq l \leq k$, we have that f has a root c of multiplicity at most q_l and such that $\mu(x-c) = -\delta_l$. Moreover, each root c of f is associated to a slope δ_l such that $\mu(x-c) = -\delta_l$.*

Proof. Consider

$$g(z) := \sum_{i=0}^{n} \partial_i f(x) z^i = \sum_{i=0}^{n} a_i z^i \in \mathbb{K}(x)[z],$$

where z is an indeterminate over $\mathbb{K}(x)$. We initially see that c is a root of $f(x)$ if and only if $c - x$ is a root of $g(z)$. In fact, we have

$$\sum_{i=0}^{n} \partial_i x^n (c-x)^i = \sum_{i=0}^{n} \binom{n}{i} x^{n-i}(c-x)^i = (x + (c-x))^n = c^n.$$

Thus, by the linearity of the Hasse derivative, for any $c \in \mathbb{K}$ we have

$$f(c) = \sum_{i=0}^{n} \partial_i f(x)(c-x)^i = g(c-x).$$

Thus, we obtain $f(c) = 0$ if and only if $g(c-x) = 0$.

Since $a_0 = f \neq 0$, we can apply Corollary 4.3 to $g(z)$ and the Newton polygon associated to

$$X' = \{(i, \mu(a_i)) \mid 0 \leq i \leq n \text{ and } \mu(\partial_i f(x)) \neq \infty\} = X.$$

Then, for each l, where $1 \leq l \leq k$, g has a root $c - x$ with multiplicity at most q_l and such that $\mu(c-x) = \mu(x-c) = -\delta_l$. That is, for each l, we have that f has a root c

with multiplicity at most q_l and such that $\mu(x-c) = -\delta_l$. By Corollary 4.3 and the same reasoning with the roots of g, we conclude that each root c of f is associated to a slope δ_l such that $\mu(x-c) = -\delta_l$. □

5. Main result

Let $\mathbb{K}$ be a field and ν a valuation on $\mathbb{K}[x]$. Fix an algebraic closure $\overline{\mathbb{K}}$ of $\mathbb{K}$. Suppose that there exists a valuation μ extending ν to $\overline{\mathbb{K}}(x)$. Let Γ_ν and Γ_μ be the value groups of ν and μ, respectively. We know that $\Gamma_\nu \subseteq \Gamma_\mu$ and $\Gamma_\mu \cong \Gamma_\nu \otimes_{\mathbb{Z}} \mathbb{Q}$.

Let us remember the definitions of $\epsilon(f)$ and $\delta(f)$ presented in the Introduction. For each $f \in \mathbb{K}[x]$ with $\nu(f) \in \Gamma_\nu$ and $\deg(f) = n > 0$, we have

$$\epsilon(f) = \max_{1\leq b\leq n}\left\{ \frac{\nu(f)-\nu(\partial_b f)}{b} \;\middle|\; \nu(\partial_b f) \in \Gamma_\nu \right\} \in \Gamma_\nu \otimes_{\mathbb{Z}} \mathbb{Q}$$

and

$$\delta(f) = \max\{\mu(x-c) \mid c \in \overline{\mathbb{K}} \text{ and } f(c) = 0\} \in \Gamma_\mu.$$

In the following, we prove our main result.

Theorem 5.1. *We have $\epsilon(f) = \delta(f)$. Moreover, $\delta(f)$ does not depend on the choice of the extension μ nor on the algebraic closure $\overline{\mathbb{K}}$.*

Proof. Consider the Newton polygon associated to

$$X = \{(i, \mu(\partial_i f)) \mid 0 \leq i \leq n \text{ and } \mu(\partial_i f) \neq \infty\}.$$

By Corollary 4.5, each root c of f is associated to a slope δ_l such that $\mu(x-c) = -\delta_l$. Moreover, each slope is associated to a root. Hence, there exists a root c' such that $\mu(x-c') = -\delta_1$. By Theorem 4.2, we have $-\delta_1 > -\delta_l$ for all l, where $2 \leq l \leq k$. Therefore,

$$\delta(f) = \max\{\mu(x-c) \mid c \in \overline{\mathbb{K}} \text{ and } f(c) = 0\} = -\delta_1.$$

Now, by the definition of the slope of a face of $PN(X)$, we have

$$\begin{aligned}\delta_1 = & \min_{1\le i\le n}\left\{\frac{\mu(\partial_i f)-\mu(f)}{i}\;\middle|\;\mu(\partial_b f)\in\Gamma_\mu\right\}\\ = & -\max_{1\le i\le n}\left\{\frac{\nu(f)-\nu(\partial_i f)}{i}\;\middle|\;\nu(\partial_b f)\in\Gamma_\nu\right\}\\ = & -\epsilon(f).\end{aligned}$$

Hence, $\epsilon(f) = -\delta_1 = \delta(f)$. We also see that $\delta(f)$ does not depend on the choice of the extension μ nor on the algebraic closure $\overline{\mathbb{K}}$, since $\epsilon(f)$ depends only on ν. □

References

[1] Bengus-Lasnier A. Minimal Pairs, Truncation and Diskoids. J. Algebra. 2021; 579: 388–427.

[2] Ćmiel H, Kuhlmann F-V, Szewczyk P. Continuity of roots for polynomials over valued fields. Comm. Algebra. 2023; 51 (4): 1383–1412.

[3] Decaup J, Spivakovsky M, Mahboub W. Abstract key polynomials and comparison theorems with the key polynomials of MacLane – Vaquie. Illinois J. Math. 2018; 62(1-4): 253 – 270.

[4] Engler A, Prestel A. Valued Fields. New York: Springer-Verlag; 2005. 205 p.

[5] Koblitz N. p-adic numbers, p-adic analysis and zeta-functions. New York: Springer-Verlag; 1977. 122 p.

[6] Mac Lane S. A construction for absolute values in polynomial rings. Trans. Amer. Math. Soc. 1936; 40: 363–395.

[7] Novacoski J. Key polynomials and minimal pairs. J. Algebra. 2019; 523: 1–14.

[8] Novacoski J, Spivakovsky M. Key polynomials and pseudo-convergent sequences. J. Algebra. 2018; 495: 199–219.

[9] Novacoski J, Spivakovsky M. On the local uniformization problem. Algebra, Logic and Number Theory, Banach Center Publ. 2016; 108: 231–238.

[10] Silva de Souza CH. Um estudo de valorizações transcendentes e algébricas via polinômios-chaves e pares minimais [Thesis (Master's degree)]. São Carlos: Universidade Federal de São Carlos; 2022 [cited 2022 Jun 10]. 244 s. Available from: https://repositorio.ufscar.br/handle/ufscar/15679.

[11] Vaquié M. Extension de valuation et polygone de Newton. Ann. Inst. Fourier. 2008; 58 (7): 2503–2541.

[12] Vaquié M. Extension d'une valuation. Trans. Amer. Math. Soc. 2007; 359(7): 3439–3481.