# BERNSTEIN'S THEOREM OVER FIELDS WITH DISCRETE VALUATION

VICTOR REINER AND STEVEN SPERBER

Abstract. For fields complete with respect to a discrete valuation, we prove a refinement of Bernstein's theorem counting the generic number of solutions to a system of $n$ polynomial equations in $n$ unknowns. The refinement predicts the number of solutions whose coordinates have given valuations, generalizing to several variables the classical use of Newton polygons for determining the valuations of the roots of a polynomial in one variable.

with an appendix by William Messing

## 1. Introduction

This paper is about a common generalization of two well-known results related to root-counting of polynomial systems and Newton polytopes. One of these results is Bernstein's theorem and the other is the classical theorem on the Newton polygon of a polynomial (or Laurent polynomial) in one variable over a completely valued field. The generalization has the same spirit as (and also generalizes) Khovanskiĭ's "Curve Theorem" [8, Theorem 27.7.11].

We first recall Bernstein's Theorem [2]. Let $\mathcal{A}_1, \ldots, \mathcal{A}_n$ be finite subsets of $\mathbb{Z}^d$. Consider a collection $\mathbf{f}$ of Laurent polynomials $f_1, \ldots, f_n$ such that for each $i$, the exponent vectors of monomials occurring in $f_i$ lie in $\mathcal{A}_i$. Let $P_i$ denote the polytope which is the convex hull of $\mathcal{A}_i$. Bernstein's Theorem states that, for suitably general collections $\mathbf{f}$, the number of common zeros (counting multiplicities) of $f_1, \ldots, f_n$ in the complex torus $(\mathbb{C}^\times)^n$ is the *Minkowski mixed volume* $\mathcal{M}(P_1, \ldots, P_n)$. Recall [5, §5.4] that the mixed volume may be defined as follows: when one dilates each $P_i$ by an independent positive scalar $\lambda_i$ and forms the Minkowski sum $\sum_{i=1}^n \lambda_i P_i$, the $n$-dimensional volume of $\sum_{i=1}^n \lambda_i P_i$ turns out to grow as a homogeneous polynomial function of degree $n$ in the $\lambda_i$'s, and $\mathcal{M}(P_1, \ldots, P_n)$ is defined to be the coefficient of the monomial $\lambda_1 \cdots \lambda_n$ in this polynomial.

For our purposes, the most convenient version of Bernstein's Theorem is the following special case of a result of Rojas [11] (see also [9]), which is valid over arbitrary algebraically closed fields. Let $\mathbf{x} = (x_1, \ldots, x_n)$ denote a vector of variables. We use the exponential notation $\mathbf{x}^{\mathbf{a}} := x_1^{a_1} \cdots x_n^{a_n}$ with $\mathbf{a} \in \mathbb{Z}^n$ for monomials in the Laurent polynomial ring $k[\mathbf{x}, \mathbf{x}^{-1}] := k[x_1, x_1^{-1}, \ldots, x_n, x_n^{-1}]$.

*Key words and phrases.* Bernstein's theorem, mixed volume, Newton polygon, discrete valuation.

First author partially supported by NSF grant DMS-9877047.

**Theorem 1.** *Let $k$ be an algebraically closed field, $\mathbf{A} = \{\mathcal{A}_i\}_{i=1}^n$ a collection of finite subsets in $\mathbb{Z}^n$, with $P_i$ the convex hull of $\mathcal{A}_i$. For $i = 1, \ldots, n$, let*

$$f_i := \sum_{\mathbf{a} \in \mathcal{A}_i} c_{i,\mathbf{a}} \mathbf{x}^{\mathbf{a}} \in k[\mathbf{x}, \mathbf{x}^{-1}].$$

*Then on a non-empty Zariski-open subset $U$ of the moduli space of such polynomials (that is, the affine space $\mathbb{A}_k^{|\mathcal{A}_1|} \times \cdots \times \mathbb{A}_k^{|\mathcal{A}_n|}$ of coefficients $(c_{i,\mathbf{a}})$), the system*

$$f_1(\mathbf{x}) = \cdots = f_n(\mathbf{x}) = 0$$

*has only isolated solutions $\mathbf{x}^*$ in the torus $(k^\times)^n$, and there are $\mathcal{M}(P_1, \ldots, P_n)$ such solutions, counted with multiplicity.*

*The set $U$ is defined by the following genericity conditions: for every choice of $\gamma \in \mathbb{Q}^n - \{0\}$, there are no solutions in $(k^\times)^n$ for the initial system*

$$\mathrm{in}_\gamma(f_1) = \ldots = \mathrm{in}_\gamma(f_n) = 0,$$

*where*

$$\mathrm{in}_\gamma(f_i(\mathbf{x})) := \sum_{\substack{\mathbf{a} \in \mathcal{A}_i: \\ \gamma \cdot \mathbf{a} \text{ is minimal}}} c_{i,\mathbf{a}} \mathbf{x}^{\mathbf{a}}. \quad \square$$

See [2] for a discussion of why the genericity conditions stated in the theorem actually are generic, that is, why they define a subset $U$ which is non-empty and Zariski-open. For convenience, we will abbreviate by $\mathbf{f} = 0$ and $\mathrm{in}_\gamma(\mathbf{f}) = 0$ the polynomial systems which appear in the above theorem.

The second result to be generalized is the classical theorem on valuations of roots of a (Laurent) polynomial in one variable over a field complete with respect to an ultrametric valuation. Let $K$ be the fraction field of a discrete valuation ring $\mathfrak{o}_K$, having unique maximal ideal $\mathfrak{m}_K$. It will be convenient for us to use the additive form of this valuation

$$\mathrm{ord} : K \to \mathbb{Z} \cup \{\infty\}$$

defined on $\mathfrak{o}_K$ by

$$\mathrm{ord}(x) = \begin{cases} \min\{N |\ x \in (\mathfrak{m}_K)^N\} \text{ for } x \in \mathfrak{o}_K - \{0\} \\ \infty \text{ for } x = 0 \end{cases}$$

then extended to $K$ by $\mathrm{ord}(\frac{x}{y}) = \mathrm{ord}(x) - \mathrm{ord}(y)$. This valuation satisfies the *ultrametric* (or *non-Archimedean*) inequality

$$\mathrm{ord}(x+y) \geq \min(\mathrm{ord}(x), \mathrm{ord}(y)). \tag{1}$$

The logarithm of ord defines an absolute value and a metric on $K$ making it a topological field, and we assume that $K$ is complete with respect to this metric. A standard result [1] asserts that ord has a unique extension to the algebraic closure of $K$ (taking values in $\mathbb{Q} \cup \{\infty\}$) and hence also to the completion $L$ of this algebraic closure.

For our purposes, we state the classical theorem on Newton polygons for Laurent polynomials, and in terms most amenable to multivariate generalization.

**Theorem 2.** [1] *Given a Laurent polynomial $f(x) = \sum_{a \in \mathbb{Z}} c_a x^a \in K[x, x^{-1}]$, define*

$$\hat{A} := \{(a, \mathrm{ord}(c_a))\}_{c_a \neq 0}$$

*a finite collection of points in $\mathbb{Z} \times \mathbb{Z}$.*

*Then for any $\gamma \in \mathbb{Q}$, the number of solutions $x^* \in L^\times$ to $f(x) = 0$ having $\mathrm{ord}(x^*) = \gamma$, counted with multiplicity, is the length of the projection to the first coordinate axis of the face in the (lower) convex hull of $\hat{A}$ on which the dot product with $(\gamma, 1)$ is minimized.* □

Theorem 4 below is a generalization of Theorem 1 (when we view a field $K$ as endowed with the trivial valuation $\mathrm{ord}(x) = 0$ for all $x \in K^\times$) and Theorem 2 (when $n = 1$). Its statement and proof bear a close resemblance to certain proofs of Bernstein's Theorem (e.g. [2, 6]) and in particular to Khovanskiĭ's "Curve Theorem" [8, Theorem 27.7.11], which it also generalizes. Roughly speaking, these proofs deform the system $\mathbf{f}(\mathbf{x}) = 0$ by introducing a new parameter $t$, and then work with Puiseux expansions for the solutions $\mathbf{x}(t)$ of $\mathbf{f}(\mathbf{x}, t) = 0$, considered as algebraic functions in the new variable $t$. In Theorem 4, the roles played by $K$ and its ord function are similar to the roles in these proofs played by the field of formal Laurent series $\mathbb{C}((t))$ and the function $\mathrm{ord} : \mathbb{C}((t)) \to \mathbb{R}$ which measures the order of vanishing at $t = 0$ (see also Remark 13 below).

Although our proof has the disadvantage that it relies on Theorem 1 and hence does *not* give an independent proof of Bernstein's Theorem, it has other advantages. It holds for discretely valued fields in all characteristics, not just for Laurent series over $\mathbb{C}$, and it uses a weaker genericity hypothesis than that of [8, Theorem 27.7.11].

We are indebted to W. Messing and J. Roberts for many helpful discussions regarding the proof of a crucial lemma from commutative algebra (Lemma 17). In particular, the Appendix contains a proof of this due to Messing.

## 2. The theorem

We begin with some notation and preliminaries. As in the introduction, $K$ is a complete field under the metric induced by its valuation ord. Denote by $L$ the completion of its algebraic closure. We denote the residue class field $\bar{K} = \mathfrak{o}_K/\mathfrak{m}_K$ and the reduction map $\mathfrak{o}_K \to \bar{K}$ by $x \mapsto \bar{x}$. Because $\mathfrak{m}_K = \mathfrak{m}_L \cap \mathfrak{o}_K$, the inclusion $\mathfrak{o}_K \hookrightarrow \mathfrak{o}_L$ induces an inclusion $\bar{K} \hookrightarrow \bar{L}$. We recall [1] that this identifies $\bar{L}$ with the algebraic closure of $\bar{K}$.

Extend the ord function to a map on vectors as follows: for $\mathbf{x} = (x_1, \ldots, x_n) \in (L^\times)^n$, define $\mathrm{ord}(\mathbf{x}) := (\mathrm{ord}(x_1), \ldots, \mathrm{ord}(x_n)) \in \mathbb{Q}^n$. For vectors $\mathbf{x} = (x_1, \ldots, x_n)$ in $\mathfrak{o}_L^n$, define $\bar{\mathbf{x}} := (\bar{x}_1, \ldots, \bar{x}_n)$.

We fix once and for all a uniformizing parameter for $K$, that is, a generator $\pi$ for the maximal ideal $\mathfrak{m}_K$ of $\mathfrak{o}_K$. None of our assertions will depend in an essential way on the particular choice of $\pi$.

Fix a collection $\mathbf{A} = \{\mathcal{A}_i\}_{i=1}^n$ of finite subsets of $\mathbb{Z}^n$, and a sequence of integers $\omega = \{\omega_{i,\mathbf{a}} : i = 1, \ldots, n, \mathbf{a} \in \mathcal{A}_i\}$. For each $i = 1, \ldots, n$ define $\hat{\mathcal{A}}_i := \{(\mathbf{a}, \omega_{i,\mathbf{a}})\}_{\mathbf{a} \in \mathcal{A}_i}$ a finite subset of points in $\mathbb{Z}^{n+1}$, and denote the whole collection $\hat{\mathbf{A}} = \{\hat{\mathcal{A}}_i\}_{i=1}^n$. We define an *$\hat{\mathbf{A}}$-system of Laurent polynomials* to be a collection $\mathbf{f} = \{f_i\}_{i=1}^n \in K[\mathbf{x}, \mathbf{x}^{-1}]^n$ of the form

$$f_i := \sum_{\mathbf{a} \in \mathcal{A}_i} c_{i,\mathbf{a}} \pi^{\omega_{i,\mathbf{a}}} \mathbf{x}^{\mathbf{a}} \in K[\mathbf{x}, \mathbf{x}^{-1}],$$

where $c_{i,\mathbf{a}} \in \mathfrak{o}_K$. In other words, it is assumed that the Newton polytope of $f_i$ is contained in $\mathcal{A}_i$, and that each of its coefficients has a fixed lower bound on its ord. We will make frequent reference to the *reduced* coefficients $\bar{c}_{i,\mathbf{a}} \in \bar{K}$.

We also need to define for each $\gamma$ in $\mathbb{Q}^n - \{0\}$ a certain reduced initial system of $\mathbf{f} = 0$. Let

$$\overline{\mathrm{in}}_\gamma(f_i(\mathbf{x})) := \sum_{\substack{\mathbf{a}\in\mathcal{A}_i: \\ (\gamma,1)\cdot(\mathbf{a},\omega_{i,\mathbf{a}}) \text{ is minimal}}} \bar{c}_{i,\mathbf{a}} \mathbf{x}^{\mathbf{a}} \in \bar{K}[\mathbf{x}, \mathbf{x}^{-1}].$$

Let $\overline{\mathrm{in}}_\gamma(\mathbf{f}) = 0$ denote the system $\overline{\mathrm{in}}_\gamma(f_1) = \cdots = \overline{\mathrm{in}}_\gamma(f_n) = 0$. Note that $\overline{\mathrm{in}}_\gamma(f_i)$ depends implicitly on the choice of $\hat{\mathbf{A}}$, although we have chosen to omit this dependence from the notation.

**Proposition 3.** *Let $\mathbf{f}$ be an $\hat{\mathbf{A}}$-system of Laurent polynomials. If $\mathbf{x}^* \in (L^\times)^n$ is a solution to $\mathbf{f} = 0$ having $\mathrm{ord}(\mathbf{x}) = \gamma \in \mathbb{Q}^n$, then there exists a solution in $(\bar{L}^\times)^n$ to the system $\overline{\mathrm{in}}_\gamma(\mathbf{f}) = 0$.*

*Proof.* Let $d$ be a least common denominator for $\gamma$, i.e. $\gamma \in \frac{1}{d}\mathbb{Z}^n$. Fix a $d$-th root $\pi^{\frac{1}{d}}$ of $\pi$ in $L$, so that one can define $\pi^\alpha$ for any $\alpha$ in $\frac{1}{d}\mathbb{Z}$. Since $\mathrm{ord}(\mathbf{x}) = \gamma$, this allows us to write $\mathbf{x} = (\pi^{\gamma_1} y_1, \ldots, \pi^{\gamma_n} y_n)$ for some $y_i \in \mathfrak{o}_L^\times$. One can then check that $\bar{\mathbf{y}} := (\bar{y}_1, \ldots, \bar{y}_n)$ gives the desired solution to $\overline{\mathrm{in}}_\gamma(\mathbf{f}) = 0$ as follows:

$$\begin{aligned} 0 &= f_i(\mathbf{x}^*) \\ &= \sum_{\mathbf{a}\in\mathcal{A}_i} c_{i,\mathbf{a}} \pi^{\omega_{i,\mathbf{a}}} (\mathbf{x}^*)^{\mathbf{a}} \\ &= \sum_{\mathbf{a}\in\mathcal{A}_i} \pi^{(\gamma,1)\cdot(\mathbf{a},\omega_{i,\mathbf{a}})} c_{i,\mathbf{a}} \mathbf{y}^{\mathbf{a}} \end{aligned}$$

Let $\mu_i$ be the minimal value of $(\gamma, 1) \cdot (\mathbf{a}, \omega_{i,\mathbf{a}})$ for $\mathbf{a} \in \mathcal{A}_i$. Then dividing the previous equation by $\pi^{\mu_i}$ gives an equation in $\mathfrak{o}_K$ which reduces in $\bar{K}$ to

$$\begin{aligned} 0 &= \sum_{\substack{\mathbf{a}\in\mathcal{A}_i: \\ (\gamma,1)\cdot(\mathbf{a},\omega_{i,\mathbf{a}}) \text{ is minimal}}} \bar{c}_{i,\mathbf{a}} \bar{\mathbf{y}}^{\mathbf{a}} \\ &= \overline{\mathrm{in}}_\gamma(f)(\bar{\mathbf{y}}) \end{aligned}$$

□

As before, let $P_i$ denote the convex hull of $\mathcal{A}_i$ for $i = 1, \ldots, n$, so that Theorem 1 tells us that (generically) the system $\mathbf{f} = 0$ will have exactly $\mathcal{M}(P_1, \ldots, P_n)$ isolated solutions (counted with multiplicity) in $(L^\times)^n$. Let $p : \mathbb{R}^{n+1} \to \mathbb{R}^n$ be the projection onto the first $n$ coordinates. For $\gamma \in \mathbb{Q}^n$ and $i = 1, \ldots, n$, let $\hat{P}_i^\gamma$ be the face of the convex hull of $\hat{\mathcal{A}}_i$ on which the dot product in $\mathbb{R}^{n+1}$ with $(\gamma, 1)$ is minimized, and let $P_i^\gamma = p(\hat{P}_i^\gamma)$. Note that $P_i^\gamma$ need not be a face of $P_i$, and in particular, $P_i^\gamma$ should not be confused with the face of $P_i$ on which the dot product with $\gamma$ is maximized. We also note that, by definition of $\overline{\mathrm{in}}_\gamma(\mathbf{f})$ and by Theorem 1, generically the system $\overline{\mathrm{in}}_\gamma(\mathbf{f}) = 0$ will have exactly $\mathcal{M}(P_1^\gamma, \ldots, P_n^\gamma)$ isolated solutions (counted with multiplicity) in $(\bar{L}^\times)^n$.

We can now state our main result.

**Theorem 4.** *Let $\mathbf{f}$ be an $\hat{\mathbf{A}}$-system of Laurent polynomials over $K$. On a non-empty Zariski-open subset $U$ in the moduli space $\mathbb{A}_{\bar{K}}^{|\mathcal{A}_1|} \times \cdots \times \mathbb{A}_{\bar{K}}^{|\mathcal{A}_n|}$ of reduced coefficients in $\bar{K}$, the system $\mathbf{f} = 0$ has only isolated solutions in the torus $(L^\times)^n$. Furthermore, for each $\gamma \in \mathbb{Q}^n$, the number of such solutions $\mathbf{x}^*$ having $\mathrm{ord}(\mathbf{x}^*) = \gamma$ is, counted with multiplicity, the mixed volume $\mathcal{M}(P_1^\gamma, \ldots, P_n^\gamma)$.*

*The set $U$ may be defined by the following genericity conditions: for every $\gamma \in \mathbb{Q}^n$, the system $\overline{\mathrm{in}}_\gamma(\mathbf{f}) = 0$ has no solutions in $(\bar{L}^\times)^n$ whenever $\mathcal{M}(P_1^\gamma, \ldots, P_n^\gamma) = 0$.*

Before proving the theorem, a few remarks are in order about the genericity conditions assumed in the theorem.

**Remark 5.** *(The case of trivial valuation – Bernstein's Theorem)*
In the case that the valuation on $K$ is trivial, the genericity conditions above reduce to those of Theorem 1, as we now explain. For each $\gamma \neq 0$, we have that $P_i^\gamma$ is the face of $P_i$ on which the dot product with $\gamma$ is minimized and consequently $\overline{\mathrm{in}}_\gamma(\mathbf{f}) = \mathrm{in}_\gamma(\mathbf{f})$. Because each of the faces $P_i^\gamma$ lies in a hyperplane perpendicular to $\gamma$, the mixed volume $\mathcal{M}(P_1^\gamma, \ldots, P_n^\gamma)$ vanishes by an argument similar to the one given in the proof of Proposition 9 below.

**Remark 6.** *(Dependence upon uniformizing parameter $\pi$)*
Assume we had made a different choice of uniformizing parameter $\pi$, say $\pi' = u\pi$ for $u \in \mathfrak{o}_K^\times$. This has the effect of scaling each reduced coefficient $\bar{c}_{i,\mathbf{a}}$ by the unit factor $\bar{u}^{\omega_{i,\mathbf{a}}}$, and hence alters the Zariski open set $U$ by an invertible diagonal transformation.

Also, the genericity hypotheses of the theorem are unchanged: if $\mu_i$ is the minimum value of the dot product $(\gamma, 1) \cdot (\mathbf{a}, \omega_{i,\mathbf{a}})$ for $\mathbf{a}$ in $\hat{\mathcal{A}}_i$, one can check that the invertible transformation $x_i \mapsto \bar{u}^{\gamma_i} x_i$ sends $\overline{\mathrm{in}}_{\gamma,\pi}(f_i)$ to $\bar{u}^{\mu_i}\overline{\mathrm{in}}_{\gamma,\pi'}(f_i)$ (where $\overline{\mathrm{in}}_{\gamma,\pi}(-)$ means $\overline{\mathrm{in}}_\gamma(-)$ defined with respect to the uniformizer $\pi$).

**Remark 7.** *(Zariski-openness of $U$)*
To see that $U$ *is* non-empty and Zariski-open, first note that even though there are infinitely many $\gamma \in \mathbb{Q}^n$, there are only finitely many different systems $\overline{\mathrm{in}}_\gamma(\mathbf{f}) = 0$, depending on the convex hulls of the finite sets $\hat{\mathcal{A}}_i$ which occur in the system $\hat{\mathbf{A}}$.

Now fix one such system $\overline{\mathrm{in}}_\gamma(\mathbf{f}) = 0$. Note that its coefficients are a subset of all of the reduced coefficients $\{\bar{c}_{i,\mathbf{a}}\}$ (namely those for which $(\gamma, 1) \cdot (\mathbf{a}, \omega_{i,\mathbf{a}})$ is maximized). Hence a non-empty Zariski-open condition on this subset of coefficients gives rise to a non-empty Zariski-open condition on $\mathbb{A}_{\bar{K}}^{|\mathcal{A}_1|} \times \cdots \times \mathbb{A}_{\bar{K}}^{|\mathcal{A}_n|}$. The conditions stated in Theorem 1 under which $\overline{\mathrm{in}}_\gamma(\mathbf{f}) = 0$ will have exactly $\mathcal{M}(P_1^\gamma, \ldots, P_n^\gamma)$ solutions are known to be non-empty and Zariski-open (see [2, 9, 11]). Hence the conditions under which there will be *no* solutions to $\overline{\mathrm{in}}_\gamma(\mathbf{f}) = 0$ if $\mathcal{M}(P_1^\gamma, \ldots, P_n^\gamma) = 0$ are non-empty and Zariski-open.

The set $U$ in the theorem is then the intersection of these finitely many non-empty Zariski-open subsets indexed by the different systems $\overline{\mathrm{in}}_\gamma(\mathbf{f}) = 0$, and hence is non-empty and Zariski-open.

**Remark 8.** *(Alternate genericity conditions)*
One might be tempted to alter the genericity conditions in the above theorem by adding in the extra hypothesis that the initial systems $\overline{\mathrm{in}}_\gamma(\mathbf{f}) = 0$, when they have roots, have only *simple* roots (cf. the Curve Theorem [8, Theorem 27.7.11]). In particular, one might have in mind to use things like multivariate Newton's method/Hensel's lemma to approximate roots of $\mathbf{f} = 0$ starting with the roots of $\overline{\mathrm{in}}_\gamma(\mathbf{f}) = 0$. However, note that in characteristic $p$, for certain choices of $\hat{\mathbf{A}}$ these roots for the initial system will *never be simple.* As an example, when $n = 1$ and $K$ has characteristic $p$, the equation $c_1 x^p + c_2 = 0$ will always encounter this difficulty. Nevertheless, our theorem as stated still applies (regardless of the choice of the minimum ord values $\omega_1, \omega_2$ for $c_1, c_2$).

The proof of Theorem 4 occupies the remainder of this section. We begin with some simple facts about the commutation of the operations $\mathrm{in}_\gamma$ and $\overline{\mathrm{in}}_\gamma$:

**Proposition 9.** *For any $\gamma, \gamma'$ in $\mathbb{Q}^n$ and for $\epsilon > 0$ sufficiently small,*

(a)
$$\overline{\mathrm{in}}_{\gamma'}(\mathrm{in}_\gamma(\mathbf{f})) = \overline{\mathrm{in}}_{\epsilon^{-1}\gamma+\gamma'}(\mathbf{f})$$
*and* $\mathcal{M}(P_1^{\epsilon^{-1}\gamma+\gamma'}, \ldots, P_n^{\epsilon^{-1}\gamma+\gamma'}) = 0$ *whenever* $\gamma \neq 0$.

(b)
$$\mathrm{in}_{\gamma'}(\overline{\mathrm{in}}_\gamma(\mathbf{f})) = \overline{\mathrm{in}}_{\gamma+\epsilon\gamma'}(\mathbf{f})$$
*and* $\mathcal{M}(P_1^{\gamma+\epsilon\gamma'}, \ldots, P_n^{\gamma+\epsilon\gamma'}) = 0$ *whenever* $\gamma' \neq 0$.

*Proof.* Both (a), (b) follow from a basic fact in polyhedral geometry (see e.g. [12, Equation (2.3)]): for any finite subset $\mathcal{A}$ of $\mathbb{R}^n$, if $\mathrm{in}_\delta(\mathcal{A})$ for $\delta \in \mathbb{R}^n$ denotes the subset of $\mathcal{A}$ on which the dot product with $\delta$ is minimized, then for $\epsilon > 0$ sufficiently small,
$$\mathrm{in}_{\delta'}(\mathrm{in}_\delta(\mathcal{A})) = \mathrm{in}_{\delta+\epsilon\delta'}(\mathcal{A}).$$
Using this new notation, we can rephrase the definitions of $\mathrm{in}, \overline{\mathrm{in}}$:
$$\mathrm{in}_\gamma(f_i) = \sum_{\mathbf{a}\in\mathrm{in}_\gamma(\mathcal{A}_i)} c_{i,\mathbf{a}}\pi^{\omega_{i,\mathbf{a}}}\mathbf{x}^{\mathbf{a}} = \sum_{(\mathbf{a},\omega_{i,\mathbf{a}})\in\mathrm{in}_{(\gamma,0)}(\hat{\mathcal{A}}_i)} c_{i,\mathbf{a}}\pi^{\omega_{i,\mathbf{a}}}\mathbf{x}^{\mathbf{a}}$$
$$\overline{\mathrm{in}}_\gamma(f_i) = \sum_{(\mathbf{a},\omega_{i,\mathbf{a}})\in\mathrm{in}_{(\gamma,1)}(\hat{\mathcal{A}}_i)} \bar{c}_{i,\mathbf{a}}\mathbf{x}^{\mathbf{a}}$$

We note here a basic fact which will be used below: for any Laurent polynomial $f$, the initial form $\mathrm{in}_\gamma(f)$ only contains monomials which lie on an affine hyperplane on which the dot product with $\gamma$ is constant (but the same need *not* be true for $\overline{\mathrm{in}}_\gamma(f)$ !).

To show (a), note that from the above rephrasings one has
$$\overline{\mathrm{in}}_{\gamma'}(\mathrm{in}_\gamma(f_i)) = \sum_{(\mathbf{a},\omega_{i,\mathbf{a}})\in\mathrm{in}_{(\gamma',1)}(\mathrm{in}_{(\gamma,0)}(\hat{\mathcal{A}}_i))} \bar{c}_{i,\mathbf{a}}\mathbf{x}^{\mathbf{a}}$$
and since
$$\mathrm{in}_{(\gamma',1)}(\mathrm{in}_{(\gamma,0)}(\hat{\mathcal{A}}_i)) = \mathrm{in}_{(\gamma,0)+\epsilon(\gamma',1)}(\hat{\mathcal{A}}_i) = \mathrm{in}_{\epsilon^{-1}(\gamma,0)+(\gamma',1)}(\hat{\mathcal{A}}_i) = \mathrm{in}_{(\epsilon^{-1}\gamma+\gamma',1)}(\hat{\mathcal{A}}_i)$$
the first assertion in (a) follows.

To see the second assertion in (a), note that if $\gamma \neq 0$, then for each $i = 1, \ldots, n$, by the first assertion in (a), the polytope $P_i^{\epsilon^{-1}\gamma+\gamma'}$ lies in an affine hyperplane in $\mathbb{R}^n$ on which the dot product with $\gamma$ is constant. Consequently, any Minkowski sum $\sum_{i=1}^n \lambda_i P_i^{\epsilon^{-1}\gamma+\gamma'}$ is a sum of polytopes lying in a family of parallel affine hyperplanes, and thus lies in such an affine hyperplane, implying that it will have zero $n$-dimensional volume. This implies the vanishing of the the mixed volume $\mathcal{M}(P_1^{\epsilon^{-1}\gamma+\gamma'}, \ldots, P_n^{\epsilon^{-1}\gamma+\gamma'})$.

The proof of (b) is similar. From the above rephrasings, one has
$$\mathrm{in}_{\gamma'}(\overline{\mathrm{in}}_\gamma(f_i)) = \sum_{(\mathbf{a},\omega_{i,\mathbf{a}})\in\mathrm{in}_{(\gamma',0)}(\mathrm{in}_{(\gamma,1)}(\hat{\mathcal{A}}_i))} \bar{c}_{i,\mathbf{a}}\mathbf{x}^{\mathbf{a}}$$
and since
$$\mathrm{in}_{(\gamma',0)}(\mathrm{in}_{(\gamma,1)}(\hat{\mathcal{A}}_i)) = \mathrm{in}_{(\gamma,1)+\epsilon(\gamma',0)}(\hat{\mathcal{A}}_i) = \mathrm{in}_{(\gamma+\epsilon\gamma',1)}(\hat{\mathcal{A}}_i)$$

the first assertion in (b) follows. The second assertion in (b) follows because when $\gamma' \neq 0$, each $P_i^{\gamma+\epsilon\gamma'}$ lies in an affine hyperplane of $\mathbb{R}^n$ on which the dot product with $\gamma'$ is constant. $\square$

The next lemma explains the reason behind the phrasing of the genericity conditions in Theorem 4.

**Lemma 10.** *The genericity conditions assumed in Theorem 4 imply both*

(a) *the genericity needed to apply Theorem 1 with $k = L$ and conclude that $\mathbf{f} = 0$ has exactly $\mathcal{M}(P_1, \ldots, P_n)$ isolated solutions, counted with multiplicities, in $(L^\times)^n$.*

(b) *the genericity needed to apply Theorem 1 with $k = \bar{L}$ and conclude that $\overline{\text{in}}_\gamma(\mathbf{f}) = 0$ has exactly $\mathcal{M}(P_1^\gamma, \ldots, P_n^\gamma)$ isolated solutions, counted with multiplicities, in $(\bar{L}^\times)^n$, for each $\gamma$ in $\mathbb{Q}^n - \{0\}$.*

*Proof.* For (a), one argues that for any $\gamma \neq 0$ in $\mathbb{Q}^n$, there can be no solution in $(L^\times)^n$ to the system $\text{in}_\gamma(\mathbf{f}) = 0$. If $\mathbf{x}^*$ were such a solution, and if $\gamma' = \text{ord}(\mathbf{x}^*)$, then Proposition 3 would give rise to a solution in $(\bar{L}^\times)^n$ for the system $\overline{\text{in}}_{\gamma'}(\text{in}_\gamma(\mathbf{f})) = 0$, which is the same as the system $\overline{\text{in}}_{\epsilon^{-1}\gamma+\gamma'}(\mathbf{f}) = 0$ by Proposition 9(a). This contradicts the genericity assumption of Theorem 4 by the second assertion of Proposition 9(a).

For (b), one argues that for any $\gamma, \gamma'$ in $\mathbb{Q}^n$ with $\gamma' \neq 0$, there can be no solution in $(\bar{L}^\times)^n$ to the system $\text{in}_{\gamma'}(\overline{\text{in}}_\gamma(\mathbf{f})) = 0$: this is the same as the system $\overline{\text{in}}_{\gamma+\epsilon\gamma'}(\mathbf{f}) = 0$ by Proposition 9(b), and it should have no solutions by the second assertion of Proposition 9(b). $\square$

The assertion of Theorem 4 can now be rephrased as the following lemma, after introducing some terminology. Let $\hat{\mu}_\gamma$ denote the (finite) number of solutions $\mathbf{x}^*$ to $\mathbf{f} = 0$ in $(L^\times)^n$ having $\text{ord}(\mathbf{x}^*) = \gamma$, counted with multiplicity. Let $\mu_\gamma$ denote the (finite) number of solutions $\bar{\mathbf{x}}^*$ to $\overline{\text{in}}_\gamma(\mathbf{f}) = 0$ in $(\bar{L}^\times)^n$, counted with multiplicity. Recall that by Lemma 10 part (b), we have $\mu_\gamma = \mathcal{M}(P_1^\gamma, \ldots, P_n^\gamma)$.

**Lemma 11.** *Under the assumptions of Theorem 4,*

$$\hat{\mu}_\gamma = \mu_\gamma (= \mathcal{M}(P_1^\gamma, \ldots, P_n^\gamma))$$

*for all $\gamma \in \mathbb{Q}^n - \{0\}$.*

*Proof.* We first explain why there only finitely many $\gamma \in \mathbb{Q}^n - \{0\}$ for which either $\hat{\mu}_\gamma$ or $\mu_\gamma$ are non-zero. Since $\hat{\mu}_\gamma \neq 0$ implies $\mu_\gamma \neq 0$ by Proposition 3, we need only consider when $\mu_\gamma$ is non-zero.

For $\mu_\gamma$, this follows from the method of *coherent mixed subdivisons* for calculating the mixed volume $\mathcal{M}(P_1, \ldots, P_n)$ which is used in [6]. This method involves lifting each point $\mathbf{a}$ in $\mathcal{A}_i \subset \mathbb{R}^n$ to a generic height $w_{i,\mathbf{a}}$ in $\mathbb{R}^{n+1}$, and then computing the volumes of certain projected faces of the Minkowski sum of the convex hulls of these lifted sets $\mathcal{A}_i$. If the heights $w_{i,\mathbf{a}}$ are chosen to be generic perturbations of our given integers $\omega_{i,\mathbf{a}}$, the following equality of mixed volumes results:

$$\mathcal{M}(P_1, \ldots, P_n) = \sum_\gamma \mathcal{M}(P_1^\gamma, \ldots, P_n^\gamma), \tag{2}$$

where $\gamma$ ranges over the finite subset $\Gamma \subset \mathbb{Q}^n - \{0\}$ consisting of those $\gamma$ for which $(\gamma, 1)$ is a normal vector of some $n$-dimensional face of the Minkowski sum

$\mathrm{conv}(\hat{\mathcal{A}}_1) + \cdots + \mathrm{conv}(\hat{\mathcal{A}}_n)$. This theory also predicts that $\mathcal{M}(P_1^\gamma, \ldots, P_n^\gamma) = 0$ for any $\gamma$ outside of this finite set $\Gamma$, and hence $\mu_\gamma = 0$ for such $\gamma$.

We next claim that it suffices to show the seemingly weaker inequality

$$\hat{\mu}_\gamma \leq \mu_\gamma \text{ for every } \gamma \in \Gamma. \tag{3}$$

To see this claim, note that assuming (3), we would have

$$\mathcal{M}(P_1, \ldots, P_n) = \sum_{\gamma \in \Gamma} \hat{\mu}_\gamma \leq \sum_{\gamma \in \Gamma} \mu_\gamma = \sum_{\gamma \in \Gamma} \mathcal{M}(P_1^\gamma, \ldots, P_n^\gamma).$$

But then because of Equation (2), the inequality (3) must actually be an equality, as desired.

We now work on proving this inequality (3) for each $\gamma \in \mathbb{Q}^n$. We can reduce to the case where $\gamma = 0$ and each $f_i$ has coefficients in $\mathfrak{o}_L$ by making a change of variables, as follows. As in the proof of Proposition 3, fix a choice of $\pi^{\frac{1}{d}}$ where $d$ is the least common denominator of $\gamma$, so that $\pi^{\gamma_i}$ is well-defined for each $i$. Let $\mu_i$ denote the minimal value of the dot product $(\gamma, 1) \cdot (\mathbf{a}, \omega_{i,\mathbf{a}})$ for $\mathbf{a}$ in $\mathcal{A}_i$. Then one can check that $g_i(\mathbf{y}) := \pi^{-\mu_i} f(\pi^{\gamma_1} y_1, \ldots, \pi^{\gamma_n} y_n)$ has the following properties:

- It lies in $\mathfrak{o}_L[\mathbf{y}, \mathbf{y}^{-1}]$.
- Solutions $\mathbf{x}^*$ to $\mathbf{f}(\mathbf{x}) = 0$ in $(L^\times)^n$ having $\mathrm{ord}(\mathbf{x}^*) = \gamma$ biject via the invertible map $x_i = \pi^{\gamma_i} y_i$ with solutions $\mathbf{y}^*$ to $\mathbf{g}(\mathbf{y}) = 0$ having $\mathrm{ord}(\mathbf{y}^*) = 0$.
- $\overline{\mathrm{in}}_\gamma(\mathbf{f}) = \overline{\mathrm{in}}_0(\mathbf{g})$.

Note that $\overline{\mathrm{in}}_0(-)$ is simply the reduction map

$$\begin{array}{ccc} \mathfrak{o}_L[\mathbf{x}, \mathbf{x}^{-1}] & \to & \bar{L}[\mathbf{x}, \mathbf{x}^{-1}] \\ f & \mapsto & \bar{f} \end{array}$$

induced from the reduction map on the coefficients $\mathfrak{o}_L \to \bar{L}$.

As a further reduction, we can replace $L$ with any finite algebraic extension of $K$ that contains the coordinates of each of the (finitely many) solutions $\mathbf{x}^*$ in $(L^\times)^n$ of the system $\mathbf{f} = 0$. Having done this, the new field $L$ will be discretely valued, so we can also assume without loss of generality that $K = L$.

It only remains to show then that for a complete field $K$ with discrete valuation, given $\mathbf{f} \in \mathfrak{o}_K[\mathbf{x}, \mathbf{x}^{-1}]^n$ with the property that

- the system $\mathbf{f} = 0$ has only a finite number of isolated solutions in $(K^\times)^n$,
- the system $\bar{\mathbf{f}} = 0$ in $\bar{K}[\mathbf{x}, \mathbf{x}^{-1}]$ has only a finite number of isolated solutions in $(\bar{K}^\times)^n$,

then the number of solutions to $\mathbf{f} = 0$ lying in $(\mathfrak{o}_K^\times)^n$ is at most the number of solutions to $\bar{\mathbf{f}} = 0$ in $(\bar{K}^\times)^n$, counted with multiplicities.

It turns out that a stronger statement holds for each solution in $(\bar{K}^\times)^n$: its multiplicity as a solution to $\bar{\mathbf{f}} = 0$ gives an upper bound for the sum of the multiplicities as solutions to $\mathbf{f} = 0$ of all points in $(\mathfrak{o}_K^\times)^n$ which reduce to it. This follows from Lemma 17 in the Appendix, applied with $A = \mathfrak{o}[\mathbf{x}, \mathbf{x}^{-1}]/(\mathbf{f})$. $\square$

This completes the proof of Theorem 4.

## 3. Examples, remarks, questions

**Example 12.**
We give an easy example involving a system of linear equations to illuminate the

nature of the genericity conditions and the conclusion in Theorem 4. Consider the system of equations

$$\begin{aligned} f_1(x_1, x_2) &= A_{11}x_1 + A_{12}x_2 - B_1 = 0 \\ f_2(x_1, x_2) &= A_{21}x_1 + A_{22}x_2 - B_2 = 0 \end{aligned}$$

equivalent to the matrix system

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}.$$

In our terminology, this implies that the Newton polygons $P_1, P_2$ coincide, with both being equal to the triangle in $\mathbb{R}^2$ having vertices $\{(0,0),(1,0),(0,1)\}$. Since the mixed volume $\mathcal{M}(P_1, P_2) = 1$, Bernstein's Theorem predicts that this system has exactly 1 solution in $(K^\times)^2$ for generic choices of the coefficients $A_{ij}, B_i$.

Now assume that the coefficients $A_{ij}, B_i$ lie in a complete discretely valued field $K$, and to further specify $\hat{\mathbf{A}}$, assume that

$$\begin{aligned} A_{i1} &= \pi^{\alpha_1} a_{i1} \\ A_{i2} &= \pi^{\alpha_2} a_{i2} \\ B_i &= \pi^{\beta} b_i \end{aligned}$$

where $a_{ij}, b_i$ lie in the valuation ring $\mathfrak{o}$.

The lifted Newton polygons $\mathrm{conv}(\hat{\mathcal{A}}_1), \mathrm{conv}(\hat{\mathcal{A}}_2)$ again coincide, both being equal to the triangle in $\mathbb{R}^3$ with vertices $\{(0,0,\beta),(1,0,\alpha_1),(0,1,\alpha_2)\}$. Their Minkowski sum $\mathrm{conv}(\hat{\mathcal{A}}_1) + \mathrm{conv}(\hat{\mathcal{A}}_2)$ is the dilation by 2 of the same triangle, having only one 2-dimensional face. This 2-dimensional face has normal vector in $\mathbb{R}^3$ of the form $(\gamma_0, 1)$ with $\gamma_0 = (\beta - \alpha_1, \beta - \alpha_2)$ in $\mathbb{R}^2$, and

$$\mathcal{M}(P_1^{\gamma_0}, P_2^{\gamma_0}) (= \mathcal{M}(P_1, P_2)) = 1.$$

Consequently, Theorem 4 predicts that generically the unique solution $(x_1, x_2)$ to the system will have $\mathrm{ord}(x_1, x_2) = \gamma_0 = (\beta - \alpha_1, \beta - \alpha_2)$.

Of course, we know how to solve this system explicitly via Cramer's rule, giving $(x_1, x_2) = (\frac{D_1}{D}, \frac{D_2}{D})$ where

$$D = \det \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad D_1 = \det \begin{bmatrix} B_1 & A_{12} \\ B_2 & A_{22} \end{bmatrix}, \quad D_2 = \det \begin{bmatrix} A_{11} & B_1 \\ A_{21} & B_2 \end{bmatrix}.$$

Here we can be very explicit about the genericity conditions assumed in the theorem, and check if they suffice to force the coordinates of this unique solution to have the predicted valuations.

If one chooses a generic vector $\gamma$ in $\mathbb{R}^2$, then the dot product of $(\gamma, 1)$ with points in the lifted triangle in $\mathbb{R}^3$ will be minimized on one of its vertices. However, somewhat less generic choices of $\gamma$ in $\mathbb{R}^2$ give rise to vectors $(\gamma, 1)$ in $\mathbb{R}^3$ whose dot product minimizes on various *edges* of the lifted triangle, producing initial systems $\overline{\mathrm{in}}_\gamma(\mathbf{f}) = 0$ of the form

$$\begin{bmatrix} \bar{a}_{11} \\ \bar{a}_{21} \end{bmatrix} x_1 = \begin{bmatrix} \bar{b}_1 \\ \bar{b}_2 \end{bmatrix}, \quad \begin{bmatrix} \bar{a}_{12} \\ \bar{a}_{22} \end{bmatrix} x_2 = \begin{bmatrix} \bar{b}_1 \\ \bar{b}_2 \end{bmatrix}, \quad \begin{bmatrix} \bar{a}_{11} & \bar{a}_{12} \\ \bar{a}_{21} & \bar{a}_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Requiring that *all* of these initial systems have no solutions in $(x_1, x_2)$ in $(\bar{K}^\times)^2$ is then seen to be equivalent to non-vanishing for the following three determinants in

$\bar{K}$:

$$\bar{D} := \det \begin{bmatrix} \bar{a}_{11} & \bar{a}_{12} \\ \bar{a}_{21} & \bar{a}_{22} \end{bmatrix}, \quad \bar{D}_1 := \det \begin{bmatrix} \bar{b}_1 & \bar{a}_{12} \\ \bar{b}_2 & \bar{a}_{22} \end{bmatrix}, \quad \bar{D}_2 := \det \begin{bmatrix} \bar{a}_{11} & \bar{b}_1 \\ \bar{a}_{21} & \bar{b}_2 \end{bmatrix}.$$

Equivalently, these conditions assert that

$$\begin{aligned} \mathrm{ord}(D) &= \alpha_1 + \alpha_2 \\ \mathrm{ord}(D_1) &= \beta + \alpha_2 \\ \mathrm{ord}(D_2) &= \alpha_1 + \beta, \end{aligned}$$

One can then see that indeed these genericity conditions suffice to imply the weaker genericity assumptions needed to apply Bernstein (cf. Lemma 10), and also to conclude that

$$\mathrm{ord}(x_1, x_2) = \left(\mathrm{ord}\left(\frac{D_1}{D}\right), \mathrm{ord}\left(\frac{D_2}{D}\right)\right) = (\beta - \alpha_1, \beta - \alpha_2) = \gamma_0.$$

**Remark 13.**
Consider the special case of Theorem 2 where $K$ is the ring of formal Laurent series $k((t))$ with $t$-adic valuation, and $k$ has characteristic zero. This relates to classic results in Walker [13], Maurer [10], as well as Khovanskiĭ's Curve Theorem on the Puiseux expansions $x(t)$ about $t = 0$ for the branches of an algebraic curve defined by $f_i(\mathbf{x}, t) \in k[x_1, \ldots, x_n, t]$ for $i = 1, \ldots, n$ with prescribed $(n+1)$-dimensional Newton polytopes $\hat{\mathcal{A}}_i$. The Puiseux expansions are vectors in $L^n$ where $L$ is the algebraic closure of $K = k((t))$. Theorem 4 or the Curve Theorem predicts (generically) that the variety $V(f_1, \ldots, f_n)$ is a curve having exactly $\mathcal{M}(P_1^\gamma, \ldots, P_n^\gamma)$ branches with Puiseux expansions about $t = 0$ of the form

$$\mathbf{x} = \begin{bmatrix} c_{\gamma_1} t^{\gamma_1} + o(t^{\gamma_1 + 1}) \\ \vdots \\ c_{\gamma_n} t^{\gamma_n} + o(t^{\gamma_n + 1})) \end{bmatrix}$$

**Remark 14.**
For some applications where $f_1, \ldots, f_n$ are *polynomials* rather than Laurent polynomials, it is more appropriate to count solutions in affine space $\mathbb{A}_L^n$ rather than the torus $(L^\times)^n$. For example, building on previous work in this direction, Huber and Sturmfels [7] calculate the generic number of solutions in complex affine space $\mathbb{A}_{\mathbb{C}}^n$ for $\mathbf{f} = 0$ where $\mathbf{f} \in \mathbb{C}[\mathbf{x}]^n$. In fact, the paper of Rojas [11] proves a much more general result than Theorem 1, which as a special case, extends the results of [7] from $\mathbb{C}$ to an arbitrary algebraically closed field $k$. For a field $K$ with discrete valution, one can (although we omit the details here) adapt the proof of Theorem 4 to prove the following generalization.

Given $f_1, \ldots, f_n \in K[\mathbf{x}]$, choose a large positive number $N$, and define *pointed* versions $\mathcal{A}_i^+, \hat{\mathcal{A}}_i^+$ of $\mathcal{A}_i, \hat{\mathcal{A}}_i$ as follows:

$$\begin{aligned} \mathcal{A}_i^+ &:= \mathcal{A}_i \cup \{0\} \\ \omega_{i,0}^+ &:= \begin{cases} \omega_{i,0} & \text{if } 0 \in \mathcal{A}_i \\ N & \text{if } 0 \notin \mathcal{A}_i \end{cases} \\ \hat{\mathcal{A}}_i^+ &:= \{(\mathbf{a}, \omega_{i,a}^+) : \mathbf{a} \in \mathcal{A}_i^+. \end{aligned} \tag{4}$$

Let $\mathrm{conv}(\cdot)$ denote the convex hull of a set of points. Following [7], it can be shown that for $N$ sufficiently large, the combinatorial structure of the lower convex hull of the Minkowski sum $\sum_{i=1}^{n} \operatorname{conv}(\mathcal{A}_i^+)$ is constant in the following sense: there is a finite collection of vectors $\Gamma = \{\gamma_i\} \subset \mathbb{Q}^n$, having entries which are affine functions $\gamma_i = m_i N + b_i$ of $N$, such that $\{(\gamma_i, 1)\}_{\gamma_i \in \Gamma}$ is exactly the set of inward-pointing normal vectors for $n$-dimensional faces of the lower convex hull (note that if the point $(0, N)$ is not a vertex of the face corresponding to $\gamma_i$, then $m_i = 0$). Say that $\gamma$ in $\Gamma$ is *stable* if all of the slopes $m_i$ are non-negative, and for a stable $\gamma$ in $\Gamma$ say that a vector $x \in L^n$ has $\operatorname{ord}(x)$ *agreeing with* $\gamma$ if

- $x_i = 0$ if and only if $m_i > 0$.
- $\operatorname{ord}(x_i) = \gamma_i (= b_i)$ if $m_i = 0$.

Roughly speaking, this convention means we are treating $N$ as if $N = +\infty = \operatorname{ord}(0)$.

Given a stable $\gamma$ in $\Gamma$, let $\mathcal{M}(P_{1,+}^{\gamma}, \ldots, P_{n,+}^{\gamma})$ be the mixed volume of the polytopes $P_{1,+}^{\gamma}, \ldots, P_{n,+}^{\gamma}$, where $P_{i,+}^{\gamma}$ is the image under the projection $\mathbb{R}^{n+1} \to \mathbb{R}^n$ of the face of the lower hull of $\hat{\mathcal{A}}_i^+$ on which the dot product with $(\gamma, 1)$ is minimized. We remark that for $N$ sufficiently large, the polytopes $\{(P_{1,+}^{\gamma}, \ldots, P_{n,+}^{\gamma})\}_{\gamma \in \Gamma}$ give a subdivision of $(\mathcal{A}_1^+, \ldots, \mathcal{A}_n^+)$ in the sense of [6], which refines the subdivision introduced in [7] for the sake of stating their affine root-count.

**Theorem 15.** *With the above notation, on a non-empty Zariski-open subset $U$ of such $f_i$'s in the moduli space $\mathbb{A}_{\bar{K}}^{|\mathcal{A}_1|} \times \cdots \times \mathbb{A}_{\bar{K}}^{|\mathcal{A}_n|}$ of reduced coefficients $(\bar{c}_{i,\mathbf{a}})$, the system $\mathbf{f} = 0$ has only a finite number of isolated solutions in the affine space $\mathbb{A}_L^n$. Furthermore, every such solution $\mathbf{x}^*$ has $\operatorname{ord}(\mathbf{x}^*)$ agreeing with some stable $\gamma \in \Gamma$, and the number which agree with $\gamma$, counted with multiplicity, is $\mathcal{M}(P_{1,+}^{\gamma}, \ldots, P_{n,+}^{\gamma})$.* □

**Example 16.**
We give an example to illustrate Theorems 4 and 15. Consider the system $\mathbf{f} = 0$

$$f_1 = ay^3 + bx^2y^2 = 0$$
$$f_2 = cx + dx^2 + ex^3 = 0$$

with $f_1, f_2 \in K[x, y]$ for some complete discretely valued field $K$, and assume that $a, b, c, d, e$ are generically chosen with ord-values $\alpha, \beta, \gamma, \delta, \epsilon \in \mathbb{Z}$, respectively. In this case it happens that one can enumerate solutions in the affine plane by factoring out powers of variables from $f_1, f_2$. Letting $r_1, r_2$ denote the two roots of the quadratic equation $c + dx + ex^2 = 0$, if we assume for definiteness that $\delta < \frac{1}{2}(\gamma + \epsilon)$, then the quadratic formula shows

$$\operatorname{ord}(r_1) = \gamma - \delta$$
$$\operatorname{ord}(r_2) = \delta - \epsilon$$

and $\mathbf{f} = 0$ has the following solutions:

| solution | ord | multiplicity |
|---|---|---|
| $(0, 0)$ | $(+\infty, +\infty)$ | 3 |
| $(r_1, 0)$ | $(\gamma - \delta, +\infty)$ | 2 |
| $(r_2, 0)$ | $(\delta - \epsilon, +\infty)$ | 2 |
| $(r_1, -\frac{b}{a} r_1)$ | $(\gamma - \delta, 2(\gamma - \delta) + \beta - \alpha)$ | 1 |
| $(r_2, -\frac{b}{a} r_2)$ | $(\delta - \epsilon, 2(\delta - \epsilon) + \beta - \alpha)$ | 1 |

The convex geometry of this example is illustrated in Figure 1 for the particular values

$$N = 5, \ \alpha = 1, \ \beta = 2, \ \gamma = 3, \ \delta = 3, \ \epsilon = 4.$$

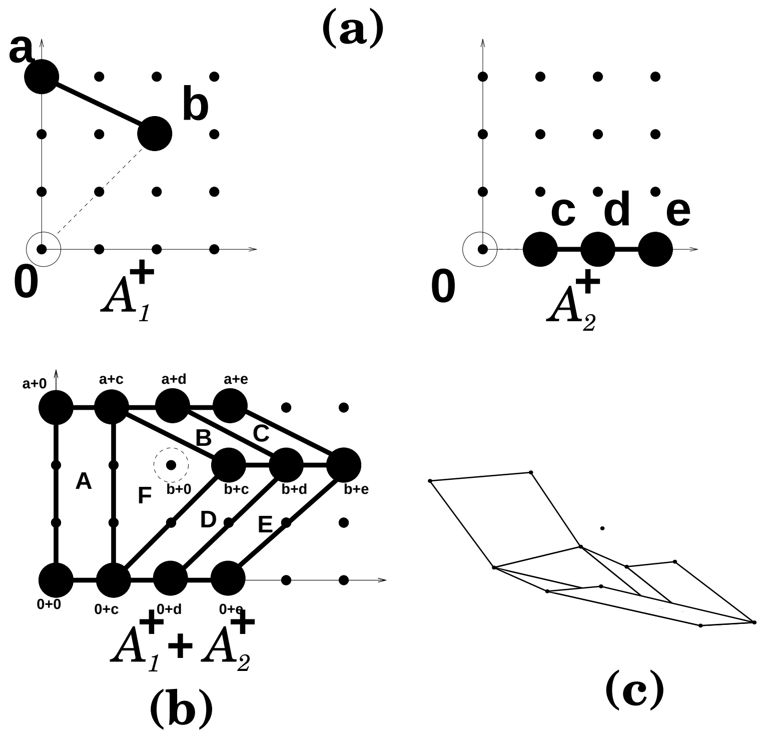


FIGURE 1. An example of Theorems 4 and 15.

Figure 1(a) shows the point sets $\mathcal{A}_1^+, \mathcal{A}_2^+$ in $\mathbb{Z}^2$, with the points labelled by the coefficient of their corresponding monomial in $f_1$ or $f_2$. Figure 1(b) depicts their Minkowski sum $\mathcal{A}_1^+ + \mathcal{A}_2^+$ along with its subdivision induced by the lower convex hull of the lifted points $\hat{\mathcal{A}}_1^+ + \hat{\mathcal{A}}_2^+$, as shown in Figure 1(c). Note that the $b+0$ does not participate in the subdivision, as it lifts to a point above the lower convex hull. The set $\Gamma$ contains six vectors $\{\gamma_A, \gamma_B, \gamma_C, \gamma_D, \gamma_E, \gamma_F\}$, corresponding to the labelled 2-dimensional faces in Figure 1(b). The corresponding normal vectors $(\gamma, 1)$ to the 2-dimensional faces in Figure 1(c), and the mixed volumes $\mathcal{M}(P_{1,+}^{\gamma}, \ldots, P_{n,+}^{\gamma})$ may then be computed to be

| | normal vector $(\gamma, 1)$ | stable? | $\mathcal{M}(P_{1,+}^{\gamma}, \ldots, P_{n,+}^{\gamma})$ |
|---|---|---|---|
| $A$ | $\big(N-\gamma,\ \ \frac{1}{3}(N-\alpha),\ \ 1\big)$ | yes | 3 |
| $B$ | $\big(\gamma-\delta,\ \ N+\delta-\beta-\gamma,\ \ 1\big)$ | yes | 2 |
| $C$ | $\big(\delta-\epsilon,\ \ N+\epsilon-\beta-\delta,\ \ 1\big)$ | yes | 2 |
| $D$ | $\big(\gamma-\delta,\ \ 2(\gamma-\delta)+\beta-\alpha,\ \ 1\big)$ | yes | 1 |
| $E$ | $\big(\delta-\epsilon,\ \ 2(\delta-\epsilon)+\beta-\alpha,\ \ 1\big)$ | yes | 1 |
| $F$ | $\big(\frac{1}{6}(N+2\alpha-3\beta),\ \ \frac{1}{3}(N-\alpha),\ \ 1\big)$ | yes | 0 |

In this example, every $\gamma$ in $\Gamma$ is stable, but $\gamma_F$ contributes no solutions because the associated mixed volume is 0. Note that the multiplicities and valuations of non-zero coordinates of the roots of $\mathbf{f} = 0$ agree with the mixed volumes and the first and second coordinates of the entries in the above table, as predicted by Theorem 15. Theorem 4 deals only with the two solutions lying in the torus $(L^\times)^2$, whose coordinates' valuations are predicted by the data in the above table corresponding to the faces labelled $D, E$ that give a subdivision of $\mathcal{A}_1 + \mathcal{A}_2$.

## 4. Appendix by William Messing

Our goal is to state and prove Lemma 17 below on multiplicities, which was used at the end of the proof of Theorem 4. We begin by recalling the definition of multiplicity [4, IV 4.7.12] which we employ.

Assume $C$ is an Artin local ring with residue field $K$, and that $C$ is a finite dimensional $k$-algebra for some field $k$. Then clearly

$$\dim_k C = \ell(C) \cdot [K : k]$$

where $\ell(C)$ denotes the length of $C$ (as a $C$-module). In the case $C = \mathcal{O}_{Y,y}$ for $y$ a point on the $k$-scheme $Y$, we define $\mathrm{mult}_y Y = \dim_k C$.

**Lemma 17.** *Let $\mathfrak{o}$ be a complete discrete valuation ring with residue field $\bar{K}$, fraction field $K$, and maximal ideal generated by the uniformizing parameter $\pi$.*

*Let $X = \mathrm{Spec}(A)$ with $A$ a finitely generated $\mathfrak{o}$-algebra. Assume $X_{\bar{K}}$ is a zero-dimensional scheme. If $x \in X_{\bar{K}}$ and $z_1, \ldots, z_N$ are the points of $X_K$ which specialize to $x$, then*

$$\sum_{i=1}^{N} \mathrm{mult}_{z_i} X_K \leq \mathrm{mult}_x X_{\bar{K}}.$$

The following is an easy application of Nakayama's Lemma.

**Lemma 18.** *Let $R$ be a local ring and integral domain, with fraction field $K$ and residue class field $\bar{K}$. Let $M$ be a finitely generated $R$-module. Then*

$$\dim_K(M \otimes_R K) \leq \dim_{\bar{K}}(M \otimes_R \bar{K}).$$

*Proof.* $M \otimes_R \bar{K}$ is a finite dimensional $\bar{K}$-vector space. Let $\{\alpha_1, \ldots, \alpha_r\}$ be a basis for $M \otimes_R \bar{K}$ over $\bar{K}$. Let $\{\hat{\alpha}_1, \ldots, \hat{\alpha}_r\}$ be an arbitrary set of liftings in $M$. By Nakayama's Lemma, $\{\hat{\alpha}_1, \ldots, \hat{\alpha}_r\}$ in fact generate $M$ over $R$. But then $\{\hat{\alpha}_1 \otimes 1, \ldots, \hat{\alpha}_r \otimes 1\}$ generate $M \otimes_R K$ over $K$. □

We now apply this lemma and some results of Grothendieck to prove Lemma 17. Let $S = \mathrm{Spec}(\mathfrak{o})$ and $X \xrightarrow{f} S$ a separated morphism of finite type. Assume $\mathfrak{o}$ is a henselian local ring. Let $f(x) = s$, the closed point of $S$. Assume the closed point $x$ is isolated in the special fiber (i.e. $\{x\}$ is open and closed in $X_{\bar{K}}$). Then $X = X' \amalg X''$, the disjoint union of open subschemes with $x \in X'$ and $X'/S$ finite and $X'' \subseteq X_K$ [4, IV 18.12.1]. In the case where $\mathfrak{o}$ is a complete Noetherian local ring (as in Lemma 17 when $\mathfrak{o}$ is a complete discrete valuation ring), the result is considerably more elementary and a simpler proof is given in [4, II 6.2.4].

By hypothesis, $X_{\bar{K}}$ is zero-dimensional, so the coordinate ring $A$ is an Artinian $\bar{K}$-algebra. As such,

$$A = A_1 \times \cdots \times A_m \tag{5}$$

where the $A_i$ are the local rings of $A$ with respect to its maximal ideals. These ideals correspond to the closed points of $X_{\bar{K}}$, and we fix notation so that $A_1 = \mathcal{O}_{X_{\bar{K}},x}$. If we set $X' = \mathrm{Spec}(A')$, then $A'$ is finite over $\mathfrak{o}$, and $\mathfrak{o}$ is henselian. Hence there is a one-to-one correspondence between the idempotents of $A$ and those of

$A'$ [3, III, Lemme 2]. Consequently the system of pairwise orthogonal idempotents corresponding to (5) can be lifted to $A'$ so that

$$A' = A'_1 \times \cdots \times A'_m$$

with $A'_i/\mathfrak{m}A'_i = A_i$, where $\mathfrak{m}$ is the maximal ideal of $\mathfrak{o}$. Any maximal ideal $M'$ of $A'_1$ must contain $\mathfrak{m}A'_1$ since otherwise $\mathfrak{m}A'_1 + M' = A'_1$, which contradicts Nakayama's Lemma.

But then clearly $M'$ is the pre-image (under reduction mod $\mathfrak{m}A'_1$) of the unique maximal ideal $M_1$ of $A_1$. Hence $M'$ is the unique maximal ideal of $A'_1$, and $A'_1$ is local. Since $A'_1$ is a factor of $A'$, clearly $A'_1 = \mathcal{O}_{X,x}$. The prime ideals of $A'$ which contain $\pi$ trivialize in $A' \otimes_{\mathfrak{o}} K$. Thus the points of $X_K$ which specialize to $x$ correspond to the height zero prime ideals of $A'_1$ which do not contain $\pi$. Let $\{\mathfrak{p}_i\}_{i=1}^N$ be those prime ideals of $A'_1$ not containing $\pi$, and let $\mathfrak{p}'_i = \mathfrak{p}_i \otimes_{\mathfrak{o}} K$ be the corresponding prime ideals in $A'_1 \otimes_{\mathfrak{o}} K$. Let $z_i$ denote point of $X_K$ specializing to $x$ which corresponds to $\mathfrak{p}'_i$. The direct product factorization of the Artinian ring $A'_1 \otimes_{\mathfrak{o}} K$ yields

$$A'_1 \otimes_{\mathfrak{o}} K = \prod_{i=1}^{N} (A'_1 \otimes_{\mathfrak{o}} K)_{\mathfrak{p}'_i}.$$

Then

$$\begin{aligned} \dim_K(A'_1 \otimes_{\mathfrak{o}} K) &= \sum_{i=1}^{N} \dim_K(A'_1 \otimes_{\mathfrak{o}} K)_{\mathfrak{p}'_i} \\ &= \sum_{i=1}^{N} \operatorname{mult}_{z_i} X_K. \end{aligned}$$

But by Lemma 18 above,

$$\begin{aligned} \dim_K(A'_1 \otimes_{\mathfrak{o}} K) &\le \dim_{\bar{K}}(A'_1 \otimes_{\mathfrak{o}} \bar{K}) \\ &= \dim_{\bar{K}} A_1 \\ &= \operatorname{mult}_x(X_{\bar{K}}). \end{aligned}$$

This completes the proof of Lemma 17.

## References


[1] E. Artin, *Theory of algebraic numbers*, (Notes by Gerhard Würges from lectures held at the Mathematisches Institut, Göttingen, Germany, in the Winter Semester, 1956/7), Göttingen 1959

[2] D.N. Bernstein, *The number of roots of a system of equations*, Func. Anal. Appl. **9** (1975), 1–4.

[3] N. Bourbaki, Éléments de mathématique. Fascicule XXVIII. Algèbre commutative. Chapitre 3 et 4, Hermann, Paris, 1961.

[4] A. Grothendieck, Éléments de géométrie algébrique, Inst. Hautes Études Sci. Publ. Math..

[5] W. Fulton, Introduction to toric varieties, *Annals of Mathematics Studies* **131**, Princeton University Press, Princeton, NJ, 1993.

[6] B. Huber and B. Sturmfels, *A polyhedral method for solving sparse polynomial systems*, Mathematics of Computation **64** (1995), 1541–1555.

[7] B. Huber and B. Sturmfels, *Bernstein's theorem in affine space*, **64** (1995), 1541–1555.

[8] A.G. Khovanskiĭ, *Addendum 3. Algebra and Mixed Volumes* in Yu.D. Burago and V.A. Zalgaller, *Geometric inequalities*, Translated from the Russian by A. B. Sosinskiĭ. Grundlehren der Mathematischen Wissenschaften **285**, Springer-Verlag, Berlin, 1988

[9] A.G. Kushnirenko, *Newton polytopes and the Bezout Theorem*, Functional Anal. Appl. **10** (1976), no. 3, 233–235 (1977).

[10] J. Maurer, *Puiseux expansions for space curves*, Manuscripta Mathematica **32** (1980), 91–100.
[11] J.M. Rojas, *Toric intersection theory for affine root counting*, J. Pure Appl. Algebra **136** (1999), 67–100.
[12] B. Sturmfels, Gröbner bases and convex polytopes, *University Lecture Series* **8**, American Math. Society, Providence RI, 1996.
[13] R. J. Walker, Algebraic curves Reprint of the 1950 edition. Springer-Verlag, New York-Heidelberg, 1978.

*E-mail address*: `messing@math.umn.edu, reiner@math.umn.edu, sperber@math.umn.edu`

School of Mathematics, University of Minnesota, Minneapolis, MN 55455