Discrete Comput Geom 17:137–141 (1997)

Discrete & Computational
Geometry

Bernstein's Theorem in Affine Space*

B. Huber[1] and B. Sturmfels[2]

[1] Department of Mathematics, Cornell University,
Ithaca, NY 14853, USA
birk@math.cornell.edu

[2] Department of Mathematics, University of California,
Berkeley, CA 94720, USA
bernd@math.berkeley.edu

Abstract. The stable mixed volume of the Newton polytopes of a polynomial system is defined and shown to equal (generically) the number of zeros in affine space $\mathbf{C}^n$. This result refines earlier bounds by Rojas, Li, and Wang [5], [7], [8]. The homotopies in [4], [9], and [10] extend naturally to a computation of all isolated zeros in $\mathbf{C}^n$.

Our object of study is a system $F = (f_1, \dots, f_n)$ of polynomial equations of the form

$$f_i = \sum_{\mathbf{q}\in\mathcal{A}_i} c_{i,\mathbf{q}} \cdot \mathbf{x}^{\mathbf{q}}, \qquad \text{where} \quad c_{i,\mathbf{q}} \in \mathbf{C}^* \quad \text{and} \quad \mathbf{x}^{\mathbf{q}} = x_1^{\mathbf{q}_1} \cdots x_n^{\mathbf{q}_n}. \tag{1}$$

Here $\mathcal{A}_i$ is a finite subset of $\mathbf{N}^n$, called the *support* of f_i, and $Q_i = \operatorname{conv}(\mathcal{A}_i)$ is the *Newton polytope* of f_i. The *mixed volume* $\mathcal{M}(\mathcal{A}_1, \dots, \mathcal{A}_n)$ is the coefficient of $l_1 l_2 \cdots l_n$ in the homogeneous polynomial $\operatorname{Vol}(l_1 Q_1 + \cdots + l_n Q_n)$, where Vol is the Euclidean volume, and

$$Q_1 + \cdots + Q_n := \{x_1 + \cdots + x_n \in \mathbf{R}^n \colon x_i \in Q_i \text{ for } i = 1, \dots, n\} \tag{2}$$

denotes the Minkowski sum of polytopes [2]. The following toric root count is well known.

Theorem 1 (Bernstein's Theorem [1]). *The number of isolated zeros of F in $(\mathbf{C}^*)^n$ is bounded above by $\mathcal{M}(\mathcal{A}_1, \dots, \mathcal{A}_n)$. This bound is exact for generic choices of the coefficients $c_{i,\mathbf{q}}$.*

* This research was supported by the David and Lucile Packard Foundation and the National Science Foundation.

In many situations, studying all zeros of F in affine space $\mathbf{C}^n$, not just those in the algebraic torus $(\mathbf{C}^*)^n$, is preferred. Li and Wang [5] have shown that the number of isolated roots in $\mathbf{C}^n$ is bounded above by $\mathcal{M}(\mathcal{A}_1 \cup \{0\}, \ldots, \mathcal{A}_n \cup \{0\})$. Rojas [7] has given an alternative bound on the number of roots in $\mathbf{C}_I = \{\mathbf{x} \in \mathbf{C}^n\colon x_i = 0 \text{ only if } i \in I\}$, where $I \subseteq \{1, \ldots, n\}$. Note that $\mathbf{C}_I \cong (\mathbf{C}^*)^{n-\#I} \times \mathbf{C}^{\#I}$. Our result sharpens the bounds given in [5], [7], and [8].

Theorem 2. *The number of isolated zeros of F in $\mathbf{C}_I$ is bounded above by the I-stable mixed volume $\mathcal{SM}_I(\mathcal{A}_1, \ldots, \mathcal{A}_n)$. This bound is exact for generic choices of coefficients $c_{i,\mathbf{q}}$, provided F has only finitely many roots in $\mathbf{C}_I$ (see Lemma 5).*

To define the *I-stable mixed volume* we modify the process of computing the Li–Wang bound $\mathcal{M}(\mathcal{A}_1 \cup \{0\}, \ldots, \mathcal{A}_n \cup \{0\})$ by subdivisions as in [4]. Let $P_i = \operatorname{conv}(\mathcal{A}_i \cup \{0\})$ and $\hat{P}_i = \operatorname{conv}\big(\{(\mathbf{q}, \omega_i(\mathbf{q})) \in \mathbf{N}^{n+1}\colon \mathbf{q} \in \mathcal{A}_i \cup \{0\}\}\big)$, where ω_i is the function which maps each point of $\mathcal{A}_i$ to zero and, if $0 \notin \mathcal{A}_i$, lifts the zero vector 0 to one. A *lower face* of a polytope in $\mathbf{R}^{n+1}$ is a face which has an inner normal with positive $(n+1)$st coordinate. The lower facets $\hat{C}$ of the Minkowski sum $\hat{P}_1 + \cdots + \hat{P}_n$ are themselves sums $\hat{C} = \hat{C}_1 + \cdots + \hat{C}_n$, where each $\hat{C}_i$ is a lower face of $\hat{P}_i$. Let $(\gamma^C, 1) = (\gamma_1^C, \ldots, \gamma_n^C, 1)$ be the unique inner normal of $\hat{C}$ whose last coordinate is equal to one, and set $C_i := \pi(\hat{C}_i)$, where π is the projection from $\mathbf{R}^{n+1}$ onto $\mathbf{R}^n$ deleting the last coordinate. The collection

$$\Delta_\omega = \{C_1 + \cdots + C_n\colon \hat{C} \text{ is a lower } \textit{facet} \text{ of } \hat{P}_1 + \cdots + \hat{P}_n\} \tag{3}$$

is the polyhedral subdivision of $P_1 + \cdots + P_n$ induced by the *lifting function* ω. An element of Δ_ω is called a *cell*. A cell C of Δ_ω is called *I-stable* if the vector γ^C is nonnegative, and in addition $\gamma_i^C > 0$ only if $i \in I$. We define the *I-stable mixed volume* $\mathcal{SM}_I(\mathcal{A}_1, \ldots, \mathcal{A}_n)$ to be the sum of the mixed volumes $\mathcal{M}(C_1, \ldots, C_n)$ where $C = C_1 + \cdots + C_n$ runs over all I-stable cells of Δ_ω.

Since the points of $\mathcal{A}_i$ remain unlifted under ω, the sum $\operatorname{conv}(\mathcal{A}_1) + \cdots + \operatorname{conv}(\mathcal{A}_n)$ appears as a cell C in the subdivision Δ_ω. In fact, it is the unique cell C with $\gamma^C = 0$. Thus the $\emptyset$-stable mixed volume $\mathcal{SM}_\emptyset(\mathcal{A}_1, \ldots, \mathcal{A}_n)$ is just the mixed volume $\mathcal{M}(\mathcal{A}_1, \ldots, \mathcal{A}_n)$ in Theorem 1. On the other extreme, summing the mixed volumes $\mathcal{M}(C_1, \ldots, C_n)$ over all cells of Δ_ω yields $\mathcal{M}(\mathcal{A}_1 \cup \{0\}, \ldots, \mathcal{A}_n \cup \{0\})$. It follows that, for all I,

$$\mathcal{M}(\mathcal{A}_1, \ldots, \mathcal{A}_n) \le \mathcal{SM}_I(\mathcal{A}_1, \ldots, \mathcal{A}_n) \le \mathcal{M}(\mathcal{A}_1 \cup \{0\}, \ldots, \mathcal{A}_n \cup \{0\}). \tag{4}$$

Example 3. The inequalities in (4) are generally strict. Consider the bivariate system

$$ay + by^2 + cxy^3 = dx + ex^2 + fx^3y = 0, \tag{5}$$

whose support sets (solid points) are pictured in Fig. 1 along with the subdivision Δ_ω tabulated in Table 1. There are, in fact, $SM_{\{1,2\}}(\mathcal{A}_1, \mathcal{A}_2) = 6$ isolated roots in $\mathbf{C}^n$, while the Li–Wang bound, $\mathcal{M}(\mathcal{A}_1 \cup \{0\}, \mathcal{A}_2 \cup \{0\}) = 8$, overcounts by two roots. Finally the $\{1\}$- and $\{2\}$-stable mixed volumes are both 4, and the $\emptyset$-stable mixed volume $\mathcal{M}(\mathcal{A}_1, \mathcal{A}_2) = 3$. The geometric process of inducing the mixed subdivision in Fig. 1 is depicted in Fig. 2.

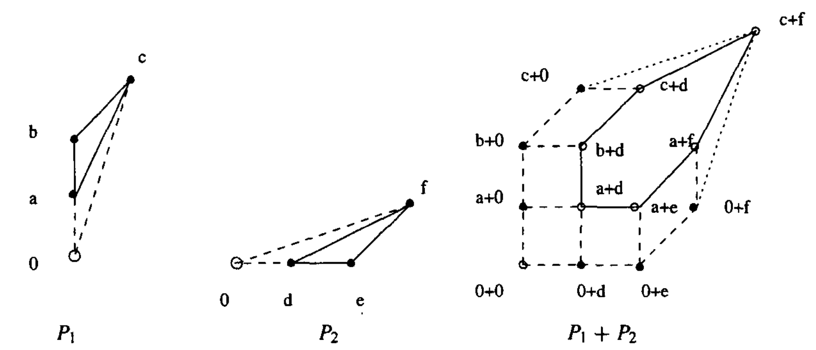

Fig. 1. An example in two dimensions.

Proof of Theorem 2. We deform the given system $F = (f_1, \dots, f_n)$ by a homotopy

$$h_i(\mathbf{x}, t) := \begin{cases} c_{i,0} \cdot t + f_i(\mathbf{x}) & \text{if } 0 \notin \mathcal{A}_i \\ f_i(\mathbf{x}) & \text{if } 0 \in \mathcal{A}_i \end{cases} \qquad (i = 1, 2, \dots, n). \tag{6}$$

All coefficients $c_{i,0}$ and $c_{i,\mathbf{q}}$ are assumed to be sufficiently generic in the sense of Theorem 1. By Bernstein's theorem, for all but finitely many t, the system (6) has $\mathcal{M}(\mathcal{A}_1 \cup \{0\}, \dots, \mathcal{A}_n \cup \{0\})$ zeros in the torus $(\mathbf{C}^*)^n$. For $t \neq 0$ it has no zeros in $\mathbf{C}^n \backslash (\mathbf{C}^*)^n$. We study the zeros of (6) as algebraic functions $\mathbf{x}(t)$ as the parameter t tends to zero [6]. As was shown in Lemma 2.2 of [5], every isolated zero $\mathbf{x}$ of F in $\mathbf{C}^n$ is the limit $\mathbf{x} = \lim_{t \to 0} \mathbf{x}(t)$ of one of the branches $\mathbf{x}(t)$. Hence to prove Theorem 2, we must count how many of the branches $\mathbf{x}(t)$ converge as $t \to 0$.

In Lemma 3.1 of [4] it was shown that the Puiseux expansion about $t = 0$ for each of the branches of the algebraic function $\mathbf{x}(t)$ has the form

$$\mathbf{x}(t) = (z_1 \cdot t^{\gamma_1^C}, \dots, z_n \cdot t^{\gamma_n^C}) + \text{higher-order terms in } t, \tag{7}$$

where $\gamma^C \in \mathbf{Q}^n$ is the normal vector for some cell C of Δ_ω, and $\mathbf{z} = (z_1, \dots, z_n) \in (\mathbf{C}^*)^n$ is a solution of the restriction of (6) to C. In other words, the vector $\mathbf{z}$ is a root of

$$\sum_{\mathbf{q} \in C_i \cap \mathcal{A}_i} c_{i,\mathbf{q}} \cdot \mathbf{z}^{\mathbf{q}} = 0 \qquad (i = 1, 2, \dots, n) \tag{8}$$

Table 1. Cells of Δ_ω.

C	γ^C	$\mathcal{M}(C)$	$\{1, 2\}$-stable
$(\{a, c, 0\}, \{f\})$	$(-2, 1)$	0	No
$(\{a, 0\}, \{d, e\})$	$(0, 1)$	1	Yes
$(\{a, 0\}, \{e, f\})$	$(-1, 1)$	1	No
$(\{b, c\}, \{d, 0\})$	$(1, -1)$	1	No
$(\{a, b\}, \{d, 0\})$	$(1, 0)$	1	Yes
$(\{a, 0\}, \{d, 0\})$	$(1, 1)$	1	Yes
$(\{c\}, \{d, f, 0\})$	$(1, -2)$	0	No
$(\{a, b, c\}, \{d, e, f\})$	$(0, 0)$	3	Yes

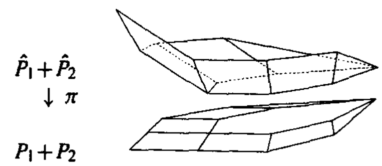

Fig. 2. Inducing the polyhedral subdivision Δ_ω.

By Bernstein's theorem, each cell C contributes $\mathcal{M}(C)$ branches of the form (7). A branch converges to an affine solution as $t \to 0$ precisely when all the exponents γ_i^C are nonnegative, while the ith coordinate of such a solution can only vanish when $\gamma_i^C > 0$. The rest of the theorem now follows by a simple deformation argument. □

The construction in the proof of Theorem 2 gives rise to the following algorithm.

Algorithm 4 (Homotopy method for finding all roots of a sparse system F in $\mathbf{C}_I$).

(1) Find the I-stable mixed cells of Δ_ω and their normals γ^C (using the methods in [4] and [10]).
(2) For each I-stable mixed cell C:
  (a) Compute all solutions $\mathbf{z}$ of (8) (using Algorithm 4.1 of [4]).
  (b) For each solution $\mathbf{z}$ in (a) set z_i to zero if $\gamma_i^C > 0$.

We close with a sufficient (but not necessary) condition for the hypothesis in the second part of Theorem 2. Lemma 5 appears in a different guise in Proposition 1.4 of [3]. The containment "$f_i \in \langle x_j\colon\ j \in J\rangle$" is equivalent to the combinatorial condition "$\mathrm{supp}(\mathbf{q}) \cap J \neq \emptyset$ for each $\mathbf{q} \in \mathcal{A}$." A more complicated but necessary and sufficient condition is presented in Lemma 3 of [8].

Lemma 5. The system F has only finitely many zeros in $\mathbf{C}_I$ if, for each subset J of I,

$$\#J \geq \#\{i \in \{1, \ldots, n\}\colon\ f_i \in \langle x_j\colon\ j \in J\rangle\}. \tag{9}$$

Proof. We abbreviate $O_J := \{\mathbf{x} \in \mathbf{C}^n\colon\ x_j = 0 \text{ if and only if } j \in J\}$. Note that $O_J \simeq (\mathbf{C}^*)^{n-\#J}$ and $\mathbf{C}_I = \cup_{J \subseteq I} O_J$. Let n_J be the cardinality on the right-hand side of (9). The restriction of f_i to O_J is zero precisely when f_i lies in the ideal $\langle x_j\colon\ j \in J\rangle$. Thus the restriction of F to O_J is a system of $n-n_J$ nonzero Laurent polynomials in $n-\#J \leq n-n_J$ variables. Theorem 1 ensures that it has at most finitely many zeros in O_J. □

References

1. D. N. Bernstein: The number of roots of a system of equations, *Functional Analysis and Its Applications* **9** (1975), 1–4.
2. T. Bonnesen and W. Fenchel: *Theory of Convex Bodies*, BCS Associates, Moscow, ID, 1987.

3. D. Eisenbud and B. Sturmfels: Finding sparse systems of parameters, *Journal of Pure and Applied Algebra* **94** (1994), 143–157.
4. B. Huber and B. Sturmfels: A polyhedral method for solving sparse polynomial systems, *Mathematics of Computation*, **64** (1995), 1541–1555.
5. T. Y. Li and X. Wang: The BKK root count in $\mathbf{C}^n$, *Mathematics of Computation*, to appear.
6. J. Maurer: Puiseux expansions for space curves, *Manuscripta Mathematica* **32** (1980), 91–100.
7. M. Rojas: A convex geometrical approach to counting the roots of a polynomial system, *Theoretical Computer Science* **133** (1994), 105–140.
8. M. Rojas and X. Wang: Counting affine roots of polynomial systems via pointed Newton polytopes, *Journal of Complexity*, to appear.
9. J. Verschelde, P. Verlinden, and R. Cools: Homotopies exploiting Newton polytopes for solving sparse polynomial systems, *SIAM Journal on Numerical Analysis*, **31** (1994), 915–930.
10. J. Verschelde, K. Gatermann, and R. Cools: Mixed volume computation by dynamic lifting applied to polynomial system solving, *Discrete and Computational Geometry*, to appear.

Received February 21, 1995, *and in revised form August* 10, 1995.