Why Polyhedra Matter in Non-Linear Equation Solving

J. Maurice Rojas

To my sister, Clarissa Amelia, on her $12^{\underline{\text{th}}}$ birthday.

Abstract. We give an elementary introduction to some recent polyhedral techniques for understanding and solving systems of multivariate polynomial equations. We provide numerous concrete examples and assume no background in algebraic geometry. Highlights include the following:

(1) A completely self-contained proof of an extension of Bernstein's Theorem. Our extension relates volumes of polytopes with the number of connected components of the complex zero set of a polynomial system, and allows any number of polynomials and/or variables.
(2) A near optimal complexity bound for computing mixed area — a quantity intimately related to counting complex roots in the plane.
(3) Illustration of the connection between polyhedral methods, amoeba theory, toric varieties, and discriminants.

We thus cover most of the theory preceding polyhedral homotopy and toric (a.k.a. sparse) resultant-based methods for solving systems of multivariate polynomial equations

1. Introduction

In a perfect world, a scientist or engineer who wishes to solve a system of polynomial equations arising from some important application would simply pick up a book on algebraic geometry, look through the table of contents, and find a well-explained, provably fast algorithm which solves his or her problem. (Algebraic geometry began 2000 years ago as the study of polynomial equations, didn't it?) He or she would then surf the web to download a good (free) implementation which would run quickly enough to be useful.

Once one stops laughing at how the real world compares, one realizes what is missing: the standard classical algebraic geometry texts (e.g., [**EGA1, Mum95,**

2000 *Mathematics Subject Classification.* Primary 12Y05; Secondary: 14M25, 52B20, 68W30.

Key words and phrases. Bernstein's Theorem, Kushnirenko's Theorem, A-discriminant, resultant, mixed subdivision, mixed volume, toric varieties, amoeba, polyhedral homotopy.

This research was partially supported by a grant from the Texas A&M College of Science and NSF Grants DMS-0138446 and DMS-0211458.

Har77, Sha94, GH94][1]) rarely contain algorithms and none contains a complexity analysis of any algorithm. Furthermore, one soon learns from experience that the specific structure underlying one's equations is rarely if ever exploited by a general purpose computational algebra package.

Considering the ubiquity of polynomial equations in applications such as geometric modeling [**Man98, Gol03**], control theory [**Sus98, NM99**], cryptography, radar imaging [**FH95**], learning theory [**VR02**], chemistry [**Gat01**], game theory [**McL97, Roj97**], and kinematics [**Emi94**] (just to mention a few applications), it then becomes clear that we need an algorithmic theory of algebraic geometry that is practical as well as rigourous. One need only look at the active research in numerical linear algebra (e.g., eigenvalue problems for large sparse matrices) to see how far we are from a completely satisfactory theory for the numerical solution of general systems of multivariate polynomial equations.

Recently, however, the introduction of algorithmic and combinatorial ideas has invigorated computational algebraic geometry. Here we give an elementary introduction to one recent aspect of computational algebraic geometry: polyhedral methods for solving systems of multivariate polynomial equations. The buzz-word for the cognicenti is **toric varieties** [**Ful93, Cox03, Sot03**]. However, rather than deriving algorithms from toric variety theory as an afterthought, we will begin directly with concrete examples and see how convex geometry naturally arises from solving equations.

EXAMPLE 1.0.1. *Suppose one has the following* 3 *equations in* 3 *unknowns x, y, and z:*

$$c_{1,1}+c_{1,2}x+c_{1,3}y^2+c_{1,4}z^3+c_{1,5}x^5y^6z^7+c_{1,6}x^6y^7x^5+c_{1,7}x^7y^5z^6+c_{1,8}x^8y^9z^9+c_{1,9}x^{10}y^9z^9+c_{1,10}x^9y^8z^9+c_{1,11}x^9y^{10}z^9+c_{1,12}x^9y^9z^{10}=0$$
$$c_{2,1}+c_{2,2}x+c_{2,3}y^2+c_{2,4}z^3+c_{2,5}x^5y^6z^7+c_{2,6}x^6y^7x^5+c_{2,7}x^7y^5z^6+c_{2,8}x^8y^9z^9+c_{2,9}x^{10}y^9z^9+c_{2,10}x^9y^8z^9+c_{2,11}x^9y^{10}z^9+c_{2,12}x^9y^9z^{10}=0$$
$$c_{3,1}+c_{3,2}x+c_{3,3}y^2+c_{3,4}z^3+c_{3,5}x^5y^6z^7+c_{3,6}x^6y^7x^5+c_{3,7}x^7y^5z^6+c_{3,8}x^8y^9z^9+c_{3,9}x^{10}y^9z^9+c_{3,10}x^9y^8z^9+c_{3,11}x^9y^{10}z^9+c_{3,12}x^9y^9z^{10}=0,$$

where the coefficients $c_{i,a}$ are any given complex numbers.

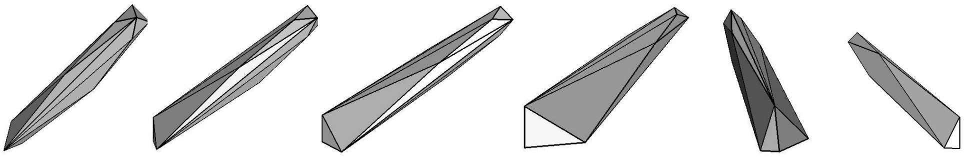


FIGURE 1. Several views of the Newton polytope shared by our three equations

One may reasonably guess that such a system of equations, being neither over-determined or under-determined, will have only finitely many roots $(x, y, z) \in \mathbb{C}^3$ with probability 1, *for any continuous probability distribution on the coefficient space $\mathbb{C}^{36}$. In fact, with probability* 1, *the number of roots will always be the same (cf. Theorems 4.2.4 and 4.2.9 of Section 4.2). What then is this "generic" number of roots?*

Noting that the maximum of the sum of the exponents in any summand of the first, second, or third equation is 28 *(i.e., our polynomials each have* **total degree** 28*), an 18th century theorem of Bézout (see, e.g.,* [**Sha94**, Ex. 1, Pg. 198]*) gives us an upper bound of* $21952 = 28^3$. *Noting that every polynomial above is of degree* 10 *with respect to x, y, or z, we can alternatively employ a* **multi-graded** *version of Bézout's Theorem (see, e.g.,* [**MS87**]*) to obtain a sharper upper bound of* $6000 = 6 \cdot 10^3$.

[1] In fairness, it should be noted that the major thrust of 20th century algebraic geometry was understanding the **topological** nature of zero sets of polynomials, rather than efficiently approximating the location of these zeros.

However, the true generic number of roots is **321**. *This number was calculated by using the correct concept in our setting: the convex hulls*[2] *of the exponent vectors (also known as the* **Newton polytopes***) of our polynomials. Here, all the Newton polytopes of our system are identical, and the volume (suitably normalized) of any one serves as the correct generic number of complex roots. This is a very special case of Main Theorem 1 below.* ◇

A natural question, especially relevant to geometric modeling, then arises: Is there an analogous theory for systems of equations expressed in other bases? In particular, the systems of equations arising from B-splines are expressed in the so-called Bernstein-Bezier basis which uses sums of terms like $\prod_i (1-x_i)^{j_i} x_i^{k_i}$. The short answer is that an analogous theory for such bases does not yet exist, and this is especially apparent when we want to find just the **real** solutions quickly. However, the philosophies of fewnomial theory [**Kho91, LRW03, Roj03**], straight-line programs [**Roj02, JKSS03**], not to mention polyhedral methods [**Roj94, HS95, Li97, Roj99b, Ver00, EP02, McD02, MR03**], are bringing us closer to a theory that can handle such questions much more efficiently than previously possible.

We now outline the main results explained in this paper.

Notation 1.0.2. *Let* $\mathbf{O}$ *denote the origin in* $\mathbb{R}^n$ *and let* $e_1,\ldots,e_n$ *denote the standard basis vectors* $\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix},\ldots,\begin{bmatrix}0\\\vdots\\0\\1\end{bmatrix}\in\mathbb{R}^n$. *Also, for any* $B\subseteq\mathbb{R}^n$, *let* $\mathrm{Conv}(B)$ *denote the smallest convex set containing* B. *Also, we let* $\mathrm{Vol}(\cdot)$ *denote the usual* n*-dimensional volume in* $\mathbb{R}^n$, *renormalized so that* $\mathrm{Vol}(\mathrm{Conv}(\{\mathbf{O},e_1,\ldots,e_n\})=1$. *Finally, we will let* # *denote the operation of taking set cardinality, and we will abuse notation slightly by setting* $\mathrm{Vol}(A):=\mathrm{Vol}(\mathrm{Conv}(A))$ *whenever* A *is a finite subset of* $\mathbb{R}^n$. ◇

Notation 1.0.3. *For any* $c\in\mathbb{C}^*:=\mathbb{C}\setminus\{0\}$ *and* $a=(a_1,\ldots,a_n)\in\mathbb{Z}^n$, *let* $x^a:=x_1^{a_1}\cdots x_n^{a_n}$ *and call* cx^a *a* **monomial term**. *Also, for any polynomial*[3] *of the form* $f(x):=\sum_{a\in A}c_a x^a$, *we call* $\mathrm{Supp}(f):=\{a \mid c_a\neq 0\}$ *the* **support** *of* f, *and define* $\mathrm{Newt}(f):=\mathrm{Conv}(\mathrm{Supp}(f))$ *to be the* **Newton polytope** *of* f. *We will frequently assume* $F:=(f_1,\ldots,f_k)$ *where, for all* i, $f_i\in\mathbb{C}[x_1,\ldots,x_n]$ *and* $\mathrm{Supp}(f_i)=A_i$. *We call such an* F *a* $\boldsymbol{k\times n}$ **polynomial system** *(over* $\mathbb{C}$*)* **with support** $\boldsymbol{(A_1,\ldots,A_k)}$. *Finally, we let* $Z_{\mathbb{C}}(F)$ *denote the set of* $x\in\mathbb{C}^n$ *with* $f_1(x)=\cdots=f_k(x)=0$. ◇

⋆ Main Theorem 1. (Special Case (full version in Sec. 8)) *Following the notation above,* $Z_{\mathbb{C}}(F)$ *has no more than* $\mathrm{Vol}(B)$ *connected components, where* $B:=\{\mathbf{O},e_1,\ldots,e_n\}\cup\bigcup_{i=1}^k \mathrm{Supp}(f_i)$. *In particular, if* F *has only finitely many complex roots, then there are no more than* $\mathrm{Vol}(B)$ *of them. Furthermore, for* $n\times n$ *polynomial systems with* $\{\mathbf{O},e_1,\ldots,e_n\}\subseteq\mathrm{Supp}(f_1)=\cdots=\mathrm{Supp}(f_n)$, *both bounds are tight.*

Although the resulting sharper bound is less trivial to evaluate than multiplying polynomial degrees, there are already freely downloadable software packages (independently by Ioannis Emiris, Birk Huber, Tien-Yien Li, and Jan Verschelde)[4] for practically computing these bounds in arbitrarily high dimensions.

In lower dimensions, one can even get a near-optimal complexity bound.

⋆ Main Theorem 2. *Following the notation of Main Theorem 1, suppose*

[2] Recall that a set $B\subseteq\mathbb{R}^n$ is **convex** iff for all $x,y\in B$, the line segment connecting x and y is also contained in B. The **convex hull** of B, $\mathrm{Conv}(B)$, is then simply the smallest convex set containing B, and the computational complexity of convex hulls of finite point sets is fairly well-understood [**PS85**].

[3] Polynomials with negative exponents are sometimes called **Laurent** polynomials.

[4] A quick search at `http://www.google.com` under any of these names quickly leads to web sites where these packages can be downloaded, along with accompanying research articles.

$k = n = 2$. Then the generic number of complex roots of a polynomial system $F = (f_1, f_2)$ can be computed within[5] *$O(\bar{N} \log \bar{N})$ arithmetic operations, involving integers with $O(b)$ bits, where $\bar{N} := \#\mathrm{Supp}(f_1) + \#\mathrm{Supp}(f_2)$ and b is the maximum bit-length of any coordinate of any A_i. Furthermore, $\Omega(\bar{N})$ arithmetic operations are needed in the worst case.*

Main Theorem 1 is proved in Section 8, preceded by ample background discussion on simpler special cases. Main Theorem 2 is proved in Section 7 as a simple consequence of mixed subdivisions in the plane. We also describe earlier related results along the way.

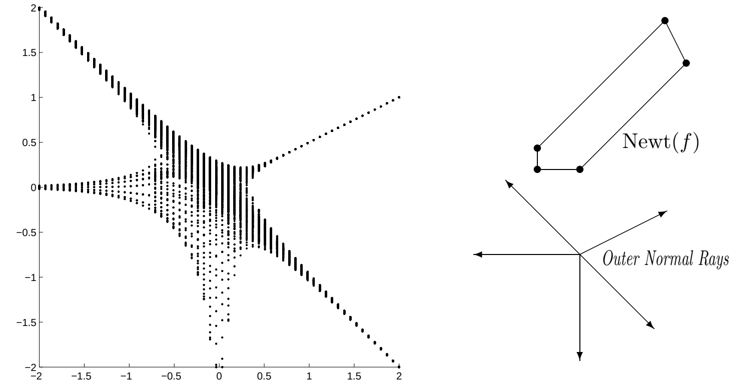
2. Zero Sets on Logarithmic Paper

Admittedly, visualizing complex zero sets in higher dimensions can be rather difficult. However, a simple and elegant approach is to look at absolute values instead to reduce the dimension, and then take a log to bring the asymptotics into view.

Example 2.0.4. *Suppose we'd like to visualize the complex zeroes of the polynomial $f(x,y) := 1 - x^2 + x - x^7y^5 + x^6y^7$. The set*

$$\mathrm{Amoeba}(f) := \{(\mathrm{Log}|x|, \mathrm{Log}|y|) \mid x, y \in \mathbb{C} \setminus \{0\}\ ,\ f(x,y) = 0\}$$

can then be sampled and drawn easily by any modern computer algebra package.



Note in particular that the tentacles of the "amoeba" appear to tend to rays that are parallel to the outer normal rays of the Newton polygon of f. ◇

Our last example is a very special case of a beautiful result due to Gelfand, Kapranov, and Zelevinsky [**GKZ94**, Ch. 6, Sec. A, pp. 193–200] (see also [**Vir01, Kap00, Roj03, Mik03**] and [**Stu02**, Ch. 9]). We will not pursue amoebae further but it is worth noting that they give the most direct and compelling evidence that polynomials are intimately related to polytopes.

3. From Binomial Systems to Volumes of Pyramids

Another simple place to begin to understand the connection between polytopes and polynomials is the special case of **binomial** systems, i.e., polynomial systems

[5] Recall that the computer scientists' notations $O(g)$ and $\Omega(g)$ are respectively used for the family of functions bounded above (resp. bounded below) asymptotically by a positive multiple of g.

where each polynomial has exactly 2 monomial terms. For such systems, there is an immediate connection to linear algebra over the integers.

Example 3.0.5. *Suppose we need to find all the complex solutions of the left-hand* 4×4 *system, where the* $c_{i,j}$ *are given nonzero complex numbers. Note in particular that this implies that any root of our system must satisfy* $xyzw \neq 0$. *A particularly elegant trick we'll generalize shortly is the following:*

$$\begin{array}{rcl} xy^7z^7w^4 & = & c_1 \\ x^6y^4z^9w^6 & = & c_2 \\ x^2y^3z^2w^6 & = & c_3 \\ x^6y^4z^8w^5 & = & c_4 \end{array}$$

Consider the 4×4 *matrix* $E := \begin{bmatrix} 1 & 7 & 7 & 4 \\ 6 & 4 & 9 & 6 \\ 2 & 3 & 2 & 6 \\ 6 & 4 & 8 & 5 \end{bmatrix}$ *whose* $i^{\underline{\text{th}}}$ *row vector is the exponent vector of the* $i^{\underline{\text{th}}}$ *equation above. Then multiplying and dividing the equations above is easily seen to be equivalent to performing row operations on* E. *For example, doing a pivot operation to zero out all but the top entry of the first column of* E *is just the computation of the matrix factorization* $\begin{bmatrix} 1 & 0 & 0 & 0 \\ -6 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ -6 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 7 & 7 & 4 \\ 6 & 4 & 9 & 6 \\ 2 & 3 & 2 & 6 \\ 6 & 4 & 8 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 7 & 7 & 4 \\ 0 & -38 & -33 & -18 \\ 0 & -11 & -12 & -2 \\ 0 & -38 & -34 & -19 \end{bmatrix}$, *which is in turn equivalent to observing that*

$$\begin{array}{rlllll} \textit{Equation 1 is...} & x & y^7 & z^7 & w^4 & = c_1 \\ \textit{(Equation 2)/(Equation 1)}^6 \textit{ is...} & & y^{-38} & z^{-33} & w^{-18} & = c_1^{-6}c_2 \\ \textit{(Equation 3)/(Equation 1)}^2 \textit{ is...} & & y^{-11} & z^{-12} & w^{-2} & = c_1^{-2}c_3 \\ \textit{(Equation 4)/(Equation 1)}^6 \textit{ is...} & & y^{-38} & z^{-34} & w^{-19} & = c_1^{-6}c_4 \end{array}$$

Note also that this new binomial system has exactly the same roots as our original system. (This follows easily from the fact that our left-most matrix above is invertible, and the entries of the inverse are all **integers**.*) So we can solve the last* 3 *equations for* (y, z, w) *and then substitute into the first equation to solve for* x *and be done.* ◇

Note, however, that if we wish to complete the solution of our example above, we must continue to use row operations on E that are **invertible over the integers**.[6] This can be done by performing a simple variant of Gauss-Jordan elimination where one uses **no** divisions. In essence, one uses elementary **integer** row operations to **minimize the absolute value** of the entries in a given column, instead of reducing them to zero.

This motivates the following definition from $19^{\underline{\text{th}}}$ century algebra.

Definition 3.0.6. *(See, e.g.,* [**Smi61**], [**Jac85**, Ch. 3.7], *or* [**Ili89**].*) Let* $\mathbb{Z}^{m\times n}$ *denote the set of* $m\times n$ *matrices with all entries integral, and let* $\mathbb{GL}_m(\mathbb{Z})$ *denote the set of all matrices in* $\mathbb{Z}^{m\times m}$ *with determinant* ± 1 *(the set of* **unimodular** *matrices). Recall that any* $m \times n$ *matrix* $[u_{ij}]$ *with* $u_{ij} = 0$ *for all* $i > j$ *is called* **upper triangular**. *Then, given any* $M \in \mathbb{Z}^{m\times n}$, *we call any identity of the form* $UM = H$, *with* $H = [h_{ij}] \in \mathbb{Z}^{n\times n}$ *upper triangular and* $U \in \mathbb{GL}_m(\mathbb{Z})$, *a* **Hermite factorization** *of* M. *Also if, in addition, we have:*

(1) $h_{ij} \geq 0$ *for all* i, j.

(2) *for all* i, *if* j *is the smallest* j' *such that* $h_{ij'} \neq 0$ *then* $h_{ij} > h_{i'j}$ *for all* $i' \leq i$.

then we call H *the* **Hermite normal form** *of* M. ◇

[6] While one could simply use rational operations on E, and thus radicals on our equations, this quickly introduces some unpleasant ambiguities regarding choices of $d^{\underline{\text{th}}}$ roots. Hence the need for our integrality restriction.

THEOREM 3.0.7. [**vdK00**] *For any $\varepsilon > 0$ and $M \in \mathbb{Z}^{n\times n}$, a Hermite factorization can be computed within $O\left((m+n)^4 m^{1+\varepsilon} h_M^{2+\varepsilon}\right)$ bit operations, where $h_M := \log(m+n+\max\limits_{i,j}|m_{ij}|)$ and $M = [m_{ij}]$. Furthermore, the Hermite normal form exists* **uniquely** *for M, and can also be computed within the preceding bit complexity bound.* ■

By extending the tricks from our last example, we easily obtain the following lemma.

LEMMA 3.0.8. *Suppose $a_1, \ldots, a_n \in \mathbb{Z}^n$ and $c_1, \ldots, c_n \in \mathbb{C}^* := \mathbb{C} \setminus \{0\}$. Let E denote the $n \times n$ matrix whose $i^{\underline{\text{th}}}$ row is the vector a_i. Then the complex roots of the binomial system $F := (x^{a_1} - c_1, \ldots, x^{a_n} - c_n)$ are exactly the complex solutions of the binomial system*

$$\begin{array}{rcl} x_1^{h_{11}} \cdots x_n^{h_{1n}} & = & c_1^{u_{11}} \cdots \ c_1^{u_{1n}} \\ \ddots \ \vdots & \vdots & \vdots \\ x_n^{h_{nn}} & = & c_1^{u_{n1}} \cdots c_1^{u_{nn}}, \end{array}$$

where $[u_{ij}]E = [h_{ij}]$ is any Hermite factorization of E. In particular, the complex roots of F can be expressed explicitly as monomials in $\sqrt[h]{c_1}, \ldots, \sqrt[h]{c_n}$, where $h := \prod_{i=1}^n h_{ii}$. ■

Letting $(\mathbb{C}^*)^n := (\mathbb{C} \setminus \{0\})^n$, we then easily obtain the following corollary.

DEFINITION 3.0.9. *Given any $k \times n$ polynomial system $F = (f_1, \ldots, f_k)$, its* **Jacobian matrix** *is the $k\times n$ matrix* $\mathrm{Jac}(F) := \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1} & \cdots & \frac{\partial f_k}{\partial x_n} \end{bmatrix}$. *We then say that a root $\zeta \in \mathbb{C}^n$ of F is* **degenerate**[7] *iff* $\mathrm{rank}\ \mathrm{Jac}(F)|_{x=\zeta} < k$, *and* **non-degenerate** *(or* **smooth***) otherwise.* ◇

COROLLARY 3.0.10. *Suppose $F = (f_1, \ldots, f_n)$ is any $n \times n$ binomial system and, for all i, v_i is either vector defined by the difference of the exponent vectors of f_i. Then F having only finitely many roots in $(\mathbb{C}^*)^n$ implies that F has exactly $|\det M|$ many, where M is the $n \times n$ matrix whose $i^{\underline{\text{th}}}$ row is v_i. In particular, the last quantity is exactly* $\mathrm{Vol}(\{\mathbf{O}, v_1, \ldots, v_n\})$.

Also, every root of F in $(\mathbb{C}^*)^n$ is non-degenerate iff F has exactly $\mathrm{Vol}(\{\mathbf{O}, v_1, \ldots, v_n\})$ *roots in $(\mathbb{C}^*)^n$. Finally, fixing the support of F, there is an algebraic hypersurface Δ in the coefficient space $\mathbb{C}^{2n}$ such that for all coefficient specializations* **outside** *of Δ, F has* **exactly** $\mathrm{Vol}(\{\mathbf{O}, v_1, \ldots, v_n\})$ *roots in $(\mathbb{C}^*)^n$.* ■

We illustrate the last portion of our corollary with the following example.

EXAMPLE 3.0.11. *Let us find all $(c_{1,1}, c_{1,2}, c_{2,1}, c_{2,2}, c_{3,1}, c_{3,2}) \in (\mathbb{C}^*)^6$ such that the left-hand system has infinitely many solutions in $(\mathbb{C}^*)^3$. In particular, by Lemma 3.0.8, the roots of this system are exactly those of the system below:*

$$\begin{array}{rcl} c_{1,1}x^2y^7z^5 & = & c_{1,2} \\ c_{2,1}x^4y^{14}z^{10} & = & c_{2,2} \\ c_{3,1}x^8y^{10}z^{14} & = & c_{3,2} \end{array}$$

$$\begin{array}{r} x^2\ y^7\ z^5 = \frac{c_{1,2}}{c_{1,1}} \\ y^{18}\ z^6 = \left(\frac{c_{1,2}}{c_{1,1}}\right)^4 \left(\frac{c_{3,2}}{c_{3,1}}\right)^{-1} \\ 1 = \left(\frac{c_{1,2}}{c_{1,1}}\right)^{-2} \frac{c_{2,2}}{c_{2,1}} \end{array}$$

Clearly then, our original system has infinitely many roots in $(\mathbb{C}^*)^3$ iff $c_{1,2}^2 c_{2,1} = c_{2,2} c_{1,1}^2$ (and no roots whatsoever if $c_{1,2}^2 c_{2,1} \neq c_{2,2} c_{1,1}^2$). So in this example, we can take $\Delta = \left\{(c_{1,1}, c_{1,2}, c_{2,1}, c_{2,2}, c_{3,1}, c_{3,2}) \in \mathbb{C}^6 \mid c_{1,2}^2 c_{1,2} = c_{2,2} c_{1,1}^2\right\}$. ◇

[7] Our definition here is slightly non-standard: Traditionally, one works more intrinsically and considers the dimension of the zero set of F in place of the number of equations defining F (compare [**Har77**]). However, for our purposes, and for the sake of simplifying our exposition, our more stringent criterion for non-degeneracy is preferable.

We conclude this section with a similar result for a slightly more complicated class of polynomial systems.

Definition 3.0.12. *Let $F := (f_1, \ldots, f_k)$ be any $k \times n$ polynomial system with $\mathrm{Supp}(f_i) = A_i$ for all i. Then we say that F* **is of type $(m_1, \ldots, m_k)$** *iff $\#A_i = m_i$ for all i. Also, we say that F is* **unmixed** *iff $A_1 = \cdots = A_k$. Finally, writing $f_i(x) = \sum_{a \in A_i} c_{i,a} x^a$ for all i, we say a property $\mathcal{P}$ pertaining to F holds* **generically** *iff there is an algebraic hypersurface $\mathcal{H} \subset \mathbb{C}^{\sum_i m_i}$ such that*

$$(c_{i,a} \mid i \in \{1, \ldots, n\}\ ,\ a \in A_i) \in \mathbb{C}^{\sum_i m_i} \setminus \mathcal{H} \Longrightarrow \mathcal{P} \text{ holds.} \quad \diamond$$

Corollary 3.0.13 (The Simplex Case of Kushnirenko's Theorem). *Given any* **unmixed** *$n \times n$ polynomial system $F = (f_1, \ldots, f_n)$ of type $(m, \ldots, m)$ with $m \leq n + 1$, let A be the support of any f_i. Then F either has exactly $\mathrm{Vol}(A)$ roots in $(\mathbb{C}^*)^n$, no roots in $(\mathbb{C}^*)^n$, or infinitely many roots in $(\mathbb{C}^*)^n$. Furthermore, for fixed A, the first possibility holds generically and implies that all the roots of F in $(\mathbb{C}^*)^n$ are non-degenerate. Finally, however many roots F has in $(\mathbb{C}^*)^n$, they can always be expressed explicitly as monomials in $\mathrm{Vol}(A)^{\underline{\text{th}}}$-roots of linear combinations of the coefficients of F and possibly some additional free parameters.*

Proof of Corollary 3.0.13: The case where A consists of a single point is clear since such an F would just be a system of n monomials, and such systems clearly have no roots off the coordinate hyperplanes. So let us assume A has at least 2 points.

By Gauss-Jordan elimination, F is then equivalent to a binomial system. So by Corollary 3.0.10, and some additional care with the Hermite normal form when F has infinitely many roots, we are done. $\blacksquare$

Corollary 3.0.13 will be the cornerstone of our proof of the special case of Main Theorem 1 where $k = n$ and F is unmixed (also known as **Kushnirenko's Theorem**). Note in particular that any Newton polytope from a polynomial system as in Corollary 3.0.13, when $\mathrm{Vol}(A) > 0$, is an n-simplex in $\mathbb{R}^n$.

4. Subdividing Polyhedra and Kushnirenko's Theorem

Here we prove the following central result which gives a strong connection between polytope volumes and the number of complex roots of polynomial systems.

Theorem 4.0.14 (Kushnirenko's Theorem). *Suppose A is a finite subset of $\mathbb{Z}^n$ and $F = (f_1, \ldots, f_n)$ is any $n \times n$ polynomial system with $\mathrm{Supp}(f_i) = A$ for all i. Then F having only finitely many roots in $(\mathbb{C}^*)^n$ implies that F has at most $\mathrm{Vol}(A)$ roots in $(\mathbb{C}^*)^n$. Furthermore, for fixed A, F generically has exactly $\mathrm{Vol}(A)$ roots in $(\mathbb{C}^*)^n$.*

This result is originally due to Anatoly Georgievich Kushnirenko.[8] His proof in [**Kus77**] takes less than a page but uses some rather non-trivial commutative algebra. Our proof requires no commutative algebra, is more visualizable for the geometrically inclined reader, and naturally leads us to some of the fastest current algorithms for solving systems of multivariate polynomial equations (see, e.g., [**HS95, Li97, Roj99b, EC00, Ver00, Roj00a, McD02, MR03**]).

Before laying the technical foundations for our proof, let us first see a concrete illustration of the main ideas. In essence, one proves Kushnirenko's Theorem by

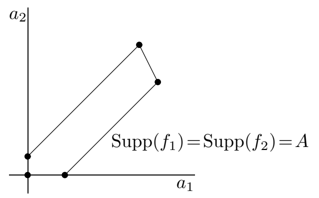
[8] His work in this area began no later than September 1974 with a question of Vladimir Arnold on Milnor numbers (multiplicities) of singular points of analytic functions.

deforming F (preserving the number of roots along the way) into a **collection** of simpler systems. Making this rigourous then provides a natural motivation for a new space (containing an embedded copy of $(\mathbb{C}^*)^n$) in which our roots will live.

EXAMPLE 4.0.15. *Consider the special case* $n=2$ *with*

$$f_1(x,y):=-2+x^2-3y+5x^7y^5+4x^6y^7$$
$$f_2(x,y):=3+2x^2+y+4x^7y^5+2x^6y^7.$$

The Newton polygon boundary and support are drawn to the right. According to Theorem 4.0.14, F either has ≤ 35 roots in $(\mathbb{C}^*)^2$ or infinitely many. (The standard and multi-graded Bézout bounds respectively reduce to $169=13^2$ and $98=2\cdot 7\cdot 7$.) The true number of roots for our example turns out to be **exactly** 35, *and these roots are all non-degenerate.* $\diamond$



To see why our last example has just 35 roots, let us start by defining a **toric deformation** [**HS95**] $\hat{F}_t:=(\hat{f}_1,\hat{f}_2)$ of $F:=(f_1,f_2)$ as follows.

EXAMPLE 4.0.16. *Let*

$$\hat{f}_1(x,y,t):=-2\boldsymbol{t}+x^2-3y+5x^7y^5+4x^6y^7\boldsymbol{t}$$
$$\hat{f}_2(x,y,t):=3\boldsymbol{t}+2x^2+y+4x^7y^5+2x^6y^7\boldsymbol{t},$$

and $\hat{A}:=\text{Supp}(\hat{f}_1)=\text{Supp}(\hat{f}_2)=\left\{\begin{bmatrix}0\\0\\1\end{bmatrix},\begin{bmatrix}2\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}7\\5\\0\end{bmatrix},\begin{bmatrix}6\\7\\1\end{bmatrix}\right\}$. *Note that $\hat{F}_1(x,y)$ is the polynomial system $F(x,y)$ from our last example. Intuitively, one would expect* 2 *equations in* 3 *unknowns to generically define a curve (cf. Theorem 4.2.4 below and the Implicit Function Theorem from calculus), and this turns out to be the case for our example. So we obtain a curve (not necessarily connected or irreducible) which contains our original finite zero set in one of its slices along the t-axis.* $\diamond$

More to the point, the number of roots of $\hat{F}_t$ in $(\mathbb{C}^*)^2$ is **constant** for all $t\in\mathbb{C}\setminus\Sigma$, where Σ is a finite set not containing 1.[9] So to show that F has exactly 35 roots in $(\mathbb{C}^*)^2$, it suffices to show that the number of roots of $\hat{F}_t$ in $(\mathbb{C}^*)^2$ is exactly 35 for some suitable fixed t outside of Σ. At least initially, this seems no easier than counting the roots of F.

The key trick then is to count something else which, for fixed t sufficiently close to 0, is easily provable to be the same as the number of roots of $\hat{F}_t$ in $(\mathbb{C}^*)^2$. This is where polyhedral subdivisions come into play almost magically.

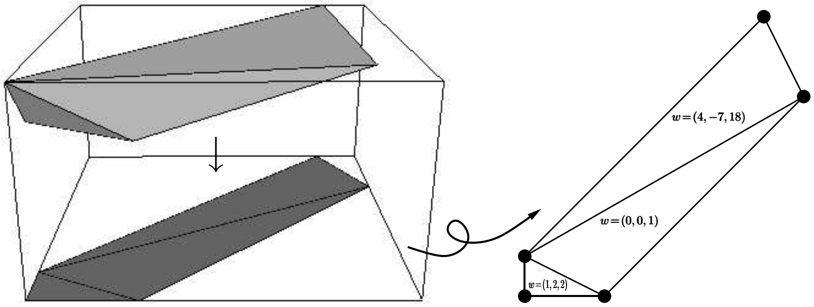
First, note that our new system is still unmixed but the equations now share a 3-dimensional Newton polytope: Next, note that any root $(x,y,t)\in(\mathbb{C}^*)^3$ of $\hat{F}$ lies on a parametric curve of the form $C_{(x_0,y_0,w)}(s):=(s^{w_1}x_0,s^{w_2}y_0,s^{w_3})$ for some $(x_0,y_0)\in(\mathbb{C}^*)^2$ and $w\in\mathbb{Z}^3$. In particular, abusing notation slightly by letting $C_{(x_0,y_0,w)}$ denote $C_{(x_0,y_0,w)}(\mathbb{C}^*)$, we will see momentarily that the set of $w\in\mathbb{Z}^3$ for which the roots of $\hat{F}_t$ in $(\mathbb{C}^*)^3$ **approach a $\boldsymbol{C_{(x_0,y_0,w)}}$ as $\boldsymbol{s\rightarrow 0}$** is dictated by the face structure of $\text{Conv}(\hat{A})$. Furthermore, **all** the roots of $\hat{F}_t$ in $(\mathbb{C}^*)^3$ approach a finite union of $C_{(x_0,y_0,w)}$ as $s\rightarrow 0$.

Let P^w denote[10] the **face of a polytope $\boldsymbol{P}$ with inner normal $\boldsymbol{w}$**.

[9] This crucial fact is elaborated a bit later in this section — specifically, Lemma 4.2.8.

[10] At this point, we will begin to use some more notions from convex geometry. This will pose no difficulty for the reader who works in geometric modeling but the reader who feels unfamiliar with these notions can take a look at, say, [**Zie95**] to see a beautiful exposition of the basics.

EXAMPLE 4.0.17. *Continuing Example 4.0.16, let us now examine the lower hull of* $\hat{A} = \left\{ \begin{bmatrix}0\\0\\1\end{bmatrix}, \begin{bmatrix}2\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}7\\5\\0\end{bmatrix}, \begin{bmatrix}6\\7\\1\end{bmatrix} \right\}$*, projected onto the* (x,y)*-plane, and its inner lower facet normals.*



In particular, the projections of the faces of the lower hull of $\hat{A}$ *onto* $\mathrm{Conv}(A)$ *induce a triangulation* $\{Q_i\}$ *of* $\mathrm{Conv}(A)$*.*

Picking $w=(1,2,2)$ *to examine the curves* $C_{(x_0,y_0,w)}$*, we see that* $\hat{F}(s^{w_1}x_0, s^{w_2}y_0, s^{w_3})$ *is exactly*

$$\begin{array}{l} s^2(-2+x_0^2-3y_0) + \textit{Higher Order Terms in } s \\ s^2(3+2x_0^2+y_0) \quad + \textit{Higher Order Terms in } s. \end{array}$$

In particular, the $(x_0,y_0)\in(\mathbb{C}^*)^2$ *which tend to a well-defined limit as* $s\to 0$ *while satisfying* $\hat{F}(s^1x_0, s^2y_0, s^2)=0$ *must also satisfy* $(-2+x_0^2-3y_0, 3+2x_0^2+y_0)=\mathbf{O}$ *in the limit. (This follows easily from the Implicit Function Theorem upon observing that the roots of* $(-2+x_0^2-3y_0, 3+2x_0^2+y_0)$ *are all non-degenerate.) So by Corollary 3.0.13 of the last section, there are exactly* $\mathrm{Vol}(\{(0,0),(2,0),(0,1)\})=2$ *such points. Put another way, the number of* $(x_0,y_0)\in(\mathbb{C}^*)^2$ *for which* $\hat{F}$ *has roots in* $(\mathbb{C}^*)^3$ *approaching* $C_{(x_0,y_0,(1,2,2))}$ *as* $s\to 0$ *is exactly* 2. ◇

Let us call the last system an **initial term** system and observe that its Newton polytopes are identical and equal to the cell $\mathrm{Conv}(\hat{A})^{(1,2,2)}$ of the subdivision of A induced by $\hat{A}$. Proceeding similarly with the other inner lower facet normals of $\hat{A}$, there are exactly $\mathrm{Vol}(\{(0,1),(7,5),(6,7)\})=18$ curves of the form $C_{(x_0,y_0,(4,-7,18))}$, and exactly $\mathrm{Vol}(\{(2,0),(0,1),(7,5)\})=15$ curves of the form $C_{(x_0,y_0,(0,0,1))}$, approached by roots of $\hat{F}$ in $(\mathbb{C}^*)^3$ as $s\to 0$. Also, the last two initial term systems have Newton polytope respectively equal to the cell of $\{Q_i\}$ with inner lower facet normal $(4,-7,18)$ or $(0,0,1)$.

To conclude, note that w not a multiple of $(1,2,2)$, $(4,-7,18)$, or $(0,0,1) \Longrightarrow$ the resulting initial term systems share Newton polytopes of dimension ≤ 1. Since $C_{(x_0,y_0,w)}=C_{(x_0,y_0,\alpha w)}$ for any $\alpha\in\mathbb{Z}$ and $w\in\mathbb{Z}^3$, another application of Corollary 3.0.13 then tells us that we have found **all** $C_{(x_0,y_0,w)}$ (with $(x_0,y_0)\in(\mathbb{C}^*)^2$ and $w\in\mathbb{Z}^3$) that are approached by roots of $\hat{F}$ in $(\mathbb{C}^*)^3$ as $t\to 0$. Since there are $35=\mathrm{Vol}(A)$ such curves, and since they don't intersect at any fixed t, this implies that $\hat{F}_t$ has exactly 35 roots in $(\mathbb{C}^*)^2$ for any $t\neq 0$ with $|t|$ sufficiently small. So, assuming every root of $\hat{F}$ in $(\mathbb{C}^*)^3$ converges to some $C_{(x_0,y_0,w)}$ as $t\to 0$, F has exactly 35 roots and we are done.

The preceding argument can be made completely general (not to mention rigourous) with just a little more work. In particular, we can prove our last assumption by constructing a space in which the roots of $\hat{F}$ all converge to well-defined limits as $t \to 0$. This is one of the main motivations behind **toric varieties**, which provide a useful and elegant way to compactify $(\mathbb{C}^*)^n$.

4.1. Polyhedral Aspects of Toric Compactifications.

Let us now give a more succinct and general definition of the initial term systems we met earlier, and formalize our constructions of $\hat{A}$ and $\hat{F}$.

DEFINITION 4.1.1. *For any $w \in \mathbb{R}^n$ and any $f \in \mathbb{C}[x_1, \ldots, x_n]$ of the form $\sum_{a\in A} c_a x^a$, let its* **initial term polynomial with respect to the weight $\boldsymbol{w}$** *be* $\mathrm{Init}_w(f) := \sum_{a\in A^w} c_a x^a$. $\diamond$

DEFINITION 4.1.2. *Given any finite subset $A \subset \mathbb{Z}^n$, a* **lifting function** *for A is any function $\omega : A \longrightarrow \mathbb{R}$ and we let $\hat{A} := \{(a, \omega(a)) \mid a \in A\}$. Also, letting $\pi : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}^n$ denote the natural projection which forgets the last coordinate, we call $A_\omega := \left\{ \mathrm{Conv}(\pi(\hat{A}^{(v,1)})) \mid v \in \mathbb{R}^n \right\}$ the* **subdivision of Conv($\boldsymbol{A}$) induced by $\boldsymbol{\omega}$**. *Finally, we say that ω is a* **generic lifting** *iff A_ω is a triangulation of* $\mathrm{Conv}(A)$. $\diamond$

DEFINITION 4.1.3. *Following the notation above, if we have in addition that $\omega(A) \subset \mathbb{Z}^n$, then for any polynomial $f(x) = \sum_{a\in A} c_a x^a$, its* **lift with respect to $\boldsymbol{\omega}$** *is the polynomial $\hat{f}(x,t) := \sum_{a\in A} c_a x^a t^{\omega(a)}$. Finally, the* **lift with respect to $(\boldsymbol{\omega_1}, \ldots, \boldsymbol{\omega_n})$** *of a $k \times n$ polynomial system $F := (f_1, \ldots, f_k)$ is simply $\hat{F} := (\hat{f}_1, \ldots, \hat{f}_k)$, where $\hat{f}_i$ is the lift of f_i with respect to ω_i for all i.* $\diamond$

LEMMA 4.1.4. *Following the notation of Definition 4.1.2, we have that for any fixed A, generic lifting functions occur generically. More precisely, there is a finite union, $\mathcal{H}_A$, of proper flats in $\mathbb{R}^{\#A}$ such that $\omega(A) \in \mathbb{R}^{\#A} \setminus \mathcal{H}_A \Longrightarrow \omega$ is a generic lifting for A.* ■

The proof of Lemma 4.1.4 is straightforward once one observes that it suffices to enforce $\omega(S)$ being a $(d+1)$-simplex for all cardinality $d+2$ subsets of A, where $d = \dim \mathrm{Conv}(A)$.

We will now refine and generalize our approach toward Example 4.0.15 as follows: After building $\hat{A}$ and $\hat{F}$ via a generic lifting function, we will build a new point set $\tilde{A}$ and a space $Y_{\tilde{A}}$ with the following properties:

(1) $Y_{\tilde{A}}$ is compact.
(2) There is an h-to-1 map from $(\mathbb{C}^*)^{n+1}$ to a dense open subset of $Y_{\tilde{A}}$, for some positive integer h related to the Hermite factorization of A.
(3) $\hat{F}$ has a well-defined complex zero set $\tilde{Z}$ in $Y_{\tilde{A}}$.
(4) There is a natural map $\pi : Y_{\tilde{A}} \longrightarrow \mathbb{P}^1_{\mathbb{C}}$, where $\mathbb{P}^1_{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ is the usual complex projective line, such that for all $t_0 \in \mathbb{C}^*$, $h\#(\pi^{-1}(t_0) \cap \tilde{Z})$ is exactly the number of roots of $\hat{F}$ in $(\mathbb{C}^*)^n$ with t-coordinate t_0.

Our proof of Kushnirenko's Theorem will then focus instead on (a) showing that $\#\left(\pi^{-1}(1) \cap \tilde{Z}\right) = \#\left(\pi^{-1}(0) \cap \tilde{Z}\right)$ generically, and (b) showing that $h\#(\pi^{-1}(0) \cap \tilde{Z}) = \mathrm{Vol}(A)$ to avoid the use of limits. We've actually already seen an example of (b), from an elementary point of view, in Example 4.0.16 of the last section. So let us now elaborate the framework needed for (a).

DEFINITION 4.1.5. *Given any finite subset $A = \{a_1, \ldots, a_N\} \subset \mathbb{Z}^n$, let $\varphi_A : (\mathbb{C}^*)^n \longrightarrow \mathbb{P}^{N-1}_{\mathbb{C}}$ — the* **generalized Veronese map with respect to $\boldsymbol{A}$** *— be the map defined by $x \mapsto [x^{a_1} : \cdots : x^{a_N}]$. We then let Y_A — the* **toric variety corresponding to the point set $\boldsymbol{A}$** *— denote the closure of $\varphi_A\left((\mathbb{C}^*)^n\right)$ in $\mathbb{P}^{N-1}_{\mathbb{C}}$.* $\diamond$

Being a closed subset of a compact space, we thus see that Y_A is compact as a topological space and this will be important later for guaranteeing that certain limits of curves exist. However, one may wonder if Y_A actually compactifies $(\mathbb{C}^*)^n$ in any reasonable way and what the closure above really means. Here's one way to make this precise.

Lemma 4.1.6. *Following the notation of Definition 4.1.5, let* $a_i := (a_{i1}, \ldots, a_{in})$ *for all* i. *Also let* E *(resp.* $\bar{E}$*) be the* $N \times n$ *(resp.* $N \times (n+1)$*) matrix whose* $i^{\underline{\text{th}}}$ *row is* a_i *(resp.* $(a_{i1}, \ldots, a_{in}, 1)$*). Finally, let* H *be the Hermite normal form of* E, *let* $\bar{U}\bar{E} = \bar{H}$ *be any Hermite factorization of* $\bar{E}$, *and let* $\bar{u}_i$ *(resp.* h*) denote the* $i^{\underline{\text{th}}}$ *row of* $\bar{U}$ *(resp. the product of the diagonal elements of* H*).*

Then $Y_A = \left\{[p_1 : \cdots : p_N] \in \mathbb{P}^{N-1}_{\mathbb{C}} \mid p^{\bar{u}^+_{r+1}} = p^{\bar{u}^-_{r+1}}, \ldots, p^{\bar{u}^+_N} = p^{\bar{u}^-_N}\right\}$, *where* r *is the rank of* $\bar{H}$ *and, for all* i, $\bar{u}^+_i - \bar{u}^-_i = \bar{u}_i$ *and* $\bar{u}^\pm_i$ *has all entries non-negative. Furthermore,* φ_A *is an* h*-to-*1 *map, i.e.,* $\#\varphi^{-1}_A(p) = h$ *for all* $p \in \varphi_A((\mathbb{C}^*)^n)$. ■

The p_i above are sometimes called **toric coordinates**. The proof of Lemma 4.1.6 is a routine application of the Hermite normal form we introduced in the last section, so let us see an example of Y_A now.

Example 4.1.7. *Taking* A *as in our last example, we obtain*

$$\varphi_A(x,y) = \left[1 : x^2 : y : x^7y^5 : x^6y^7\right] \ , \ E = \begin{bmatrix} 0 & 0 \\ 2 & 0 \\ 0 & 1 \\ 7 & 5 \\ 6 & 7 \end{bmatrix} \ , \ \text{and } \bar{E} = \begin{bmatrix} 0 & 0 & 1 \\ 2 & 0 & 1 \\ 0 & 1 & 1 \\ 7 & 5 & 1 \\ 6 & 7 & 1 \end{bmatrix}.$$

Using the `ihermite` *command in* `Maple`, *we then easily obtain that* $H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$ *is the Hermite normal form for* E *and* $\begin{bmatrix} 7 & -3 & -5 & 1 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 15 & -7 & -10 & 2 & 0 \\ 9 & -3 & -7 & 0 & 1 \end{bmatrix} \bar{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ *is a Hermite factorization for* $\bar{E}$. *So Lemma 4.1.6 tells us that our* Y_A *here can also be defined as the zero set in* $\mathbb{P}^4_{\mathbb{C}}$ *of the following collection of binomials:* $\left\langle\ p_1^{15}p_4^2 - p_2^7p_3^{10}\ ,\ p_1^9p_5 - p_2^3p_3^7\ \right\rangle$. *Furthermore, since* $h = 1 \cdot 1 \cdot 1 = 1$, *our map* φ_A *here is thus a bijection between* $(\mathbb{C}^*)^2$ *and an open dense subset of* Y_A. ⋄

The most relevant combinatorial/geometric properties of Y_A can be summarized as follows (see the companion tutorials [**Cox03, Sot03**] in this volume, and [**Stu96**], for other aspects and points of view).

Definition 4.1.8. *Given any finite subset* $A \subset \mathbb{Z}^n$, *for any face* Q *of* $\mathrm{Conv}(A)$, *the* **orbit** O_Q *(or* O_w *when* $Q = \mathrm{Conv}(A)^w$*) of* Y_A *is the subset*

$$\{[p_1 : \cdots : p_N] \in Y_A \mid a_i \notin Q \Longrightarrow p_i = 0\}.$$

Also, for any $p \in O_Q$ *with* Q *a* **proper** *face, we say that* p **lies at toric infinity**. *Finally, given any* $f_1, \ldots, f_k \in \mathbb{C}[x_1, \ldots, x_n]$ *of the form* $f_i(x) = \sum_{a \in A} c_{i,a}x^a$ *for all* i, *the* **zero set of** $\boldsymbol{F = (f_1, \ldots, f_k)}$ **in** $\boldsymbol{Y_A}$ *is simply the set of all* $[p_1 : \cdots : p_N] \in Y_A$ *with* $\sum_{j=1}^N c_{i,a_j}p_j = 0$ *for all* i. ⋄

Example 4.1.9. *Consider the* **new** 2×2 *polynomial system*

$$f_1(x,y) := 1 + x^2 - 3y + 7x^7y^5 - 11x^6y^7$$
$$f_2(x,y) := 1 + x^2 + y + 2x^7y^5 - 5x^6y^7.$$

Then the point $q = [1 : -1 : 0 : 0 : 0] \in Y_A$ — *where* $A := \left\{\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 7 \\ 5 \end{bmatrix}, \begin{bmatrix} 6 \\ 7 \end{bmatrix}\right\}$ — *is a root of* $F := (f_1, f_2)$. *In particular,* $q \notin \varphi_A((\mathbb{C}^*)^2)$, $q \in O_{(0,1)}$, *and is thus a root at* *toric infinity. Furthermore, q lies on the portion of toric infinity corresponding to the sole horizontal edge of* $\mathrm{Conv}(A)$*. Note also that there is a bona-fide root of F in $\mathbb{C}^2$ — $\left(\sqrt{-1},0\right)$ — which maps to q under φ_A, provided we extend the domain of φ_A slightly. (This will* **not** *always be the case with roots at toric infinity.) In general, toric varieties allow us to mold where and what infinity is relative to our applications.* $\diamond$

Just as a polytope can be expressed as a disjoint union of the relative interiors of its faces, Y_A can always be expressed as disjoint union of the O_Q. The lemma below follows routinely from Lemma 4.1.6 and Definition 4.1.8.

Lemma 4.1.10. *Given any finite subset $A \subset \mathbb{Z}^n$, let $N := \#A$ and let Q be any face of* $\mathrm{Conv}(A)$*. Then O_Q is a dense open subset of a d-dimensional algebraic subset of $\mathbb{P}^{N-1}_{\mathbb{C}}$, where $d = \dim Q$. In particular, Y_A is the disjoint union* $\displaystyle\bigsqcup_{Q \text{ a face of } \mathrm{Conv}(A)} O_Q$, *and* $\displaystyle Y_A \setminus \varphi_A\left((\mathbb{C}^*)^n\right) = \bigsqcup_{Q \text{ a } \textbf{proper} \text{ face of } \mathrm{Conv}(A)} O_Q$. *Finally, if $F=(f_1,\ldots,f_k)$ with* $\mathrm{Supp}(f_i)=A$ *for all i, then F has a root in O_w (cf. Definition 4.1.8) iff* $\mathrm{Init}_w(F)$ *has a root in $(\mathbb{C}^*)^n$.* ■

Generalizations of Lemma 4.1.10 can be found in standard references such as [**Stu96**] and [**GKZ94**]. Since all the faces of $\mathrm{Conv}(A)$ have a well-defined inner normal, Lemma 4.1.10 thus gives a complete characterization of when a root of F lies at toric infinity, as well as which piece of toric infinity. This is what will allow us to replace the cumbersome curves $C_{(x_0,y_0,w)}$ mentioned earlier with a single algebraic curve in Y_A.

4.2. The Smooth Case of Kushnirenko's Theorem.

Let us now review some final tools we'll need to start our proof of Kushnirenko's Theorem: The **Cayley Trick**, a simplified characterization of the **discriminant** of a system of equations, and some basic facts on algebraic curves.

Definition 4.2.1. *Given any $k \times n$ polynomial system $F = (f_1,\ldots,f_k)$ with $f_i(x) = \sum_{a \in A_i} c_{i,a} x^a$ for all i, the* **toric** *Jacobian matrix of F is the $k \times n$ matrix* $\mathrm{ToricJac}(F) = \begin{bmatrix} x_1\frac{\partial f_1}{\partial x_1} & \cdots & x_n\frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ x_1\frac{\partial f_k}{\partial x_1} & \cdots & x_n\frac{\partial f_k}{\partial x_n} \end{bmatrix}$*. Assuming F is unmixed and $A_1 = \cdots = A_k = A = \{a_1,\ldots,a_N\}$, we then say that F* **has a degenerate root at $\boldsymbol{p \in Y_A}$** *iff p is a root of F in Y_A (cf. Definition 4.1.8) and* $\mathrm{rank}\, \mathrm{ToricJac}(F)|_p < k$*. We then let the* **discriminant variety***, $\Delta(\underbrace{A,\ldots,A}_{k})$, denote the set of all*

$$(c_{1,a_1},\ldots,c_{1,a_N}) \times \cdots \times (c_{k,a_1},\ldots,c_{k,a_N}) \in (\mathbb{C}^N)^k$$

such that F has a degenerate root in Y_A. Finally, given any k-tuple of point sets from $\mathbb{R}^n$, $(A_1,\ldots,A_k)$, its **Cayley configuration** *is the point set*
$\mathrm{Cay}(A_1,\ldots,A_k) := (A_1 \times (\underbrace{0,\ldots,0}_{k-1})) \cup (A_2 \times (1,\underbrace{0,\ldots,0}_{k-2})) \cup \cdots \cup (A_k \times (\underbrace{0,\ldots,0}_{k-2},1)) \subset \mathbb{R}^{n+k-1}$. $\diamond$

One can of course define discriminant varieties for mixed systems and to do this one instead works with roots in Y_A where $A := A_1 + \cdots + A_k$. The latter **Minkowski sum** will be developed further starting in Section 7.

Remark 4.2.2. *There are deep reasons, coming from complexity theory and algebraic geometry, for why approximating roots is essentially as hard as computing membership in a discriminant variety. In particular, computing discriminants* *efficiently remains a deep and important open problem. For example, in practice one can usually only check membership in a larger hypersurface* **containing** *the discriminant variety, and even doing this is quite expensive.* $\diamond$

Example 4.2.3. *Returning to Example 4.0.16 one last time, consider the root* $p = [1 : -1 : -1 : 0 : 0] \in Y_{\hat{A}}$ *of*

$$\hat{f}_1(x,y,t) := -2t + x^2 - 3y + 5x^7y^5 + 4x^6y^7t$$
$$\hat{f}_2(x,y,t) := 3t + 2x^2 + y + 4x^7y^5 + 2x^6y^7t,$$

Note in particular that $p \in O_{(1,2,2)}$ *and thus lies at the portion of toric infinity corresponding to the smallest triangular cell of* A_ω*. The toric Jacobian matrix, in toric coordinates, is then* $\begin{bmatrix} 2p_2+35p_4+24p_5 & -3p_3+25p_4+28p_5 & -2p_1+4p_5 \\ 4p_2+28p_4+12p_5 & p_3+20p_4+14p_5 & 3p_1+2p_5 \end{bmatrix}$*. Evaluating at p, our matrix then becomes* $\begin{bmatrix} -2 & 3 & -2 \\ -4 & -1 & 3 \end{bmatrix}$*, which clearly has rank* 2*, so p is a non-degenerate root.* $\diamond$

An important and unusual property of discriminants is that the special case of a single multivariate polynomial already contains all the complexities of the general $k \times n$ case. In particular, by a little basic linear algebra, the Cayley configuration enables an explicit reduction. This approach to discriminants is sometimes called the **Cayley trick** [**GKZ94**].

Theorem 4.2.4. *Suppose* $A = \{a_1, \ldots, a_N\}$ *and* $f(x) = \sum_{a \in A} c_a x^a$ *where the* c_a *are indeterminates to be specialized later. Then there is a homogeneous polynomial* $D_A \in \mathbb{C}[c_{a_1}, \ldots, c_{a_N}]$ *such that for all* $(c_{a_1}, \ldots, c_{a_N}) \in \mathbb{C}^N$,

$$D_A(c_{a_1}, \ldots, c_{a_N}) \neq 0 \Longrightarrow (c_{a_1}, \ldots, c_{a_N}) \notin \Delta(A),$$

i.e., $\Delta(A)$ *is always contained in an algebraic hypersurface in* $\mathbb{C}^N$*. In particular, letting* $A_1, \ldots, A_k \subset \mathbb{Z}^n$ *be finite subsets and* $N_i := \#A_i$ *for all i, the expression* $\delta : = D_{\mathrm{Cay}(A_1,\ldots,A_k)}(c_{1,a_1}, \ldots, c_{1,a_{N_1}}, \ldots, c_{k,a_1}, \ldots, c_{k,a_{N_k}})$ *is homogeneous with respect to each* N_i*-tuple* $(c_{1,a_1}, \ldots, c_{1,a_{N_i}})$ *for all i, and*

$$\delta \neq 0 \Longrightarrow (c_{1,a_1}, \ldots, c_{1,a_{N_1}}) \times \cdots \times (c_{k,a_1}, \ldots, c_{k,a_{N_k}}) \notin \Delta(A_1, \ldots, A_k).$$

Furthermore, $\mathrm{Cay}(\underbrace{A, \ldots, A}_{k}) = A \times \{\mathbf{O}, e_1, \ldots, e_{k-1}\}$ *and, in the unmixed case, the degree of* $D_{\mathrm{Cay}(\underbrace{A, \ldots, A}_{k})}$ *is no more than* $\frac{(n+k)!}{n!(k-1)!}\mathrm{Vol}(A)$. ■

The polynomial D_A can in fact be chosen so that (a) it is irreducible over $\mathbb{Z}[c_{a_1}, \ldots, c_{a_N}]$ and (b) both implications above can be strengthened to "$\Longleftrightarrow$" equivalences. These additional conditions then make D_A unique (up to sign) and we then say that D_A is the ***A*-discriminant**. The existence of the weaker version above follows easily from a construction not much more difficult than the **mixed subdivisions** we introduce later: the **toric resultant** [**Emi94, EC00, Roj00a, Stu02, BEM03**]. We omit the proof for the sake of brevity but do provide an explicit example below. The reader interested in a complete proof of Theorem 4.2.4 can see the companion survey [**BEM03**] in this volume or [**Emi94**].

Example 4.2.5. *Suppose* $A := \left\{ \begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}2\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}, \begin{bmatrix}7\\5\end{bmatrix}, \begin{bmatrix}6\\7\end{bmatrix} \right\}$ *and* $\hat{A} := \left\{ \begin{bmatrix}0\\0\\1\end{bmatrix}, \begin{bmatrix}2\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}7\\5\\0\end{bmatrix}, \begin{bmatrix}6\\7\\1\end{bmatrix} \right\}$, *and that we would like to guarantee that a polynomial system of the form*

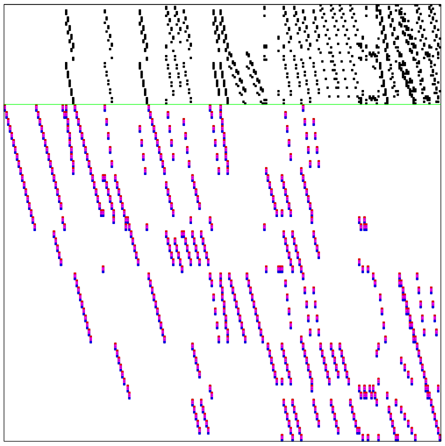
$$f_1(x,y) := c_{1,1} + c_{1,2}x^2 + c_{1,3}y + c_{1,4}x^7y^5 + c_{1,5}x^6y^7$$

$$f_2(x,y) := c_{2,1} + c_{2,2}x^2 + c_{2,3}y + c_{2,4}x^7y^5 + c_{2,5}x^6y^7$$

has **no** *degenerate roots in* Y_A **or** *that a polynomial system of the form*

$$\hat{f}_1(x,y) := c_{1,1}t + c_{1,2}x^2 + c_{1,3}y + c_{1,4}x^7y^5 + c_{1,5}x^6y^7t$$
$$\hat{f}_2(x,y) := c_{2,1}t + c_{2,2}x^2 + c_{2,3}y + c_{2,4}x^7y^5 + c_{2,5}x^6y^7t$$

has **no** *degenerate roots in* $Y_{\hat{A}}$. *Theorem 4.2.4 then tells us that a* **sufficient** *(and most likely not necessary) condition, in both cases, is the non-vanishing of a suitable polynomial in the coefficients* $\{c_{i,a}\}$. *More precisely, one can take* $B = A \times \{0,1\}$ *and try to find a polynomial* $D_B(c_1,\ldots,c_{10})$ *as specified by Theorem 4.2.4. Ordering the points of B into the sequence* $\left(\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}2\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}7\\5\\0\end{bmatrix},\begin{bmatrix}6\\7\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix},\begin{bmatrix}2\\0\\1\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix},\begin{bmatrix}7\\5\\1\end{bmatrix},\begin{bmatrix}6\\7\\1\end{bmatrix}\right)$, *we can then specialize* $(c_1,\ldots,c_{10}) = (c_{1,1},\ldots,c_{1,5},c_{2,1},\ldots,c_{2,5})$ *to check the first non-degeneracy condition or* $(c_1,\ldots,c_{10}) = (tc_{1,1},c_{1,2},c_{1,3},c_{1,4},tc_{1,5},tc_{2,1},c_{2,2},c_{2,3},c_{2,4},tc_{2,5})$ *to check the second non-degeneracy condition.* ◇



To be even more explicit in our last example, one can construct (with the assistance of some combinatorics and `Macaulay 2`) a suitable D_B as the determinant of an explicit 249×249 matrix. (This follows from a beautiful recent result from the Ph.D. thesis of Amit Khetan [**Khe03**].) In particular, D_B turns out to be a polynomial of total degree ≤ 420 in $\{c_1,\ldots,c_{10}\}$, and the corresponding matrix is highly structured and sparse. The nonzero entries of the matrix are either (a) coefficients of $G := \left(f_1, x\left(\frac{\partial f}{\partial x} + s\frac{\partial f_2}{\partial x}\right), y\left(\frac{\partial f}{\partial y} + s\frac{\partial f_2}{\partial y}\right), sf_2\right)$ or (b) determinants of 4×4 matrices whose entries are chosen from the coefficients of G. Where the entries of type (a) (resp. (b)) occur is illustrated in the lower 192 (resp. upper 57) rows of the matrix on the left. (The dark dots indicate entries of type (a) or (b); the absence of dots indicates zeros.)

An important consequence of our observations on discriminants is a concrete approach to the intuitive fact that over-determined polynomial systems usually have no roots.

Corollary 4.2.6. *Suppose F is an $n \times n$ polynomial system with support $(A_1,\ldots,A_n)$ and that there is an $(n-1)$-flat containing translates of $A_1,\ldots,A_n$. Then for fixed $(A_1,\ldots,A_n)$, F generically has no roots in $(\mathbb{C}^*)^n$. In particular, in the unmixed case, F generically has no roots in Y_A.* ■

Corollary 4.2.6 follows immediately from Theorem 4.2.4 upon observing that any root of an $(n+1) \times n$ polynomial system must be degenerate.

The final additional fact we'll need follows easily from the Implicit Function Theorem.

Definition 4.2.7. *If $X \subseteq \mathbb{P}^{N_1}_{\mathbb{C}}$ and $Y \subseteq \mathbb{P}^{N_2}_{\mathbb{C}}$ are algebraic sets, then a* **morphism** *$\psi : X \longrightarrow Y$ is a well-defined map of the form $[p_1 : \cdots : p_{N_1+1}] \mapsto [\phi_1(p_1,\ldots,p_{N_1+1}) : \cdots : \phi_{N_2+1}(p_1,\ldots,p_{N_1+1})]$, where $\phi_1,\ldots,\phi_{N_2+1}$ are homogeneous polynomials of the same degree.* ◇

Lemma 4.2.8. *Suppose $C \subset \mathbb{P}^N_{\mathbb{C}}$ is a smooth algebraic curve (not necessarily connected) and $\psi : C \longrightarrow \mathbb{P}^1_{\mathbb{C}}$ is any morphism. Then either $\#\psi(X) < \infty$ or $\psi(X) = \mathbb{P}^1_{\mathbb{C}}$. In the latter case, there is a positive integer m and a finite set* $\mathrm{Crit}_\psi \subset \mathbb{P}^1_{\mathbb{C}}$*, the* **critical values** *of ψ, such that $\#\psi^{-1}(t) = m \iff t \in \mathbb{P}^1_{\mathbb{C}} \setminus \mathrm{Crit}_\psi$.*

Finally, in the special case where C is the zero set in $Y_{\tilde{A}}$ of an $n \times (n+1)$ polynomial system $\hat{F}(x_1, \ldots, x_n, t)$ with $\mathrm{Supp}(\hat{f}_i) \subseteq \hat{A}$ *for all i, $\tilde{A} := \hat{A} \cup (\hat{A} + e_{n+1})$, and $\psi(\varphi_{\tilde{A}}(x_1, \ldots, x_n, t)) = [1 : t]$ for all $t \in \mathbb{C}^*$, we have that $t_0 \in \mathbb{C}$ lies in* $\mathrm{Crit}_\psi \iff$ *$(\hat{F}, t - t_0)$ has a degenerate root in $Y_{\tilde{A}}$.*

We are now ready to prove Kushnirenko's Theorem in the smooth case.

Theorem 4.2.9. *Fix any finite subset $A \subset \mathbb{Z}^n$ and consider the family of all $n \times n$ polynomial systems $F = (f_1, \ldots, f_n)$ with* $\mathrm{Supp}(f_i) \subseteq A$ *for all i. Then such F generically have exactly* $\mathrm{Vol}(A)$ *roots in $(\mathbb{C}^*)^n$, all of which are non-degenerate.*

Proof of Theorem 4.2.9: Let $N := \#A$ and $\{a_1, \ldots, a_N\} := A$ as before and pick any generic lifting function ω with **integral** range. Following the notation of Definition 4.1.2, let $\hat{A}$ be the lift of A with respect to ω and define $\tilde{A} := \hat{A} \cup \{(a, \omega(a) + 1) \mid a \in A\}$. Letting I denote the set of binomials defining $Y_{\hat{A}}$, observe by Lemma 4.1.6 that the set of binomials defining $Y_{\tilde{A}}$ is simply

$$I \cup \{p_{N+1}p_2 - p_1 p_{N+2}, \ldots, p_{N+1}p_N - p_1 p_{2N}\},$$

where the coordinates of $\mathbb{P}^{N-1}_{\mathbb{C}}$ and $\mathbb{P}^{2N-1}_{\mathbb{C}}$ are respectively ordered $[p_1 : \cdots : p_N]$ and $[p_1 : \cdots : p_{2N-1}]$ so that

$$\varphi_{\tilde{A}}(x, t) = [x^{a_1} t^{\omega(a_1)} : \cdots : x^{a_N} t^{\omega(a_N)} : x^{a_1} t^{\omega(a_1)+1} : \cdots : x^{a_N} t^{\omega(a_N)+1}].$$

This in turn implies that the following convention is well-defined: let us say that $[p_1 : \cdots : p_{2N-1}] \in Y_{\tilde{A}}$ is a root of F iff $\sum_{j=1}^{N} c_{i,a_j} p_j$ for all i.

So $\hat{F}$ now has a well-defined zero set in $Y_{\tilde{A}}$ as well as $Y_{\hat{A}}$, and we can at last define our promised map $\pi : Y_{\tilde{A}} \longrightarrow \mathbb{P}^1_{\mathbb{C}}$ by $p \mapsto [p_1 : p_{N+1}]$. Defining $\tilde{Z}$ (resp. Z) to be the zero set of $\hat{F}$ in $Y_{\tilde{A}}$ (resp. F in Y_A), note that there is an isomorphism between $\pi^{-1}(1) \cap \tilde{Z}$ and Z defined by $[p_1 : \cdots : p_{2N}] \longleftrightarrow [p_1 : \cdots : p_N]$.

Note also that π also induces a natural morphism from $\tilde{Z}$ to $\mathbb{P}^1_{\mathbb{C}}$. Let H be the Hermite normal form of A and h the product of the diagonal elements of H. Since the first n columns of the Hermite normal forms of A and $\tilde{A}$ are the same, Lemma 4.1.6 then tells us that the number of roots of F is exactly $h\#\left(\pi^{-1}(1) \cap \varphi_{\tilde{A}}\left((\mathbb{C}^*)^n\right)\right)$. By applying Theorem 4.2.4 to $(A, \ldots, A)$ it thus suffices to show that $\boldsymbol{h\#\left(\pi^{-1}(1) \cap \varphi_{\tilde{A}}\left((\mathbb{C}^*)^n\right)\right) = \mathrm{Vol}(A)}$ generically.

Next, note all the initial term systems of F are unmixed and have Newton polytopes with volume 0. In particular, by Corollary 4.2.6, any particular initial term system will generically have no roots. Similarly, by Corollary 3.0.13, the initial term systems of $\hat{F}$ will have each have smooth zero set generically. So by Lemma 4.1.10 it will generically be true that F will have **no** roots at toric infinity in Y_A, and all the roots of $\hat{F}$ at toric infinity in $Y_{\tilde{A}}$ will be non-degenerate. Furthermore, by applying Theorem 4.2.4 to $(\hat{A}, \ldots, \hat{A})$, we know that $\tilde{Z}$ is generically smooth. **It thus suffices to show that $\boldsymbol{[\tilde{Z}, Z,}$ and $\boldsymbol{\tilde{Z} \cap (\text{Toric Infinity in } Y_{\tilde{A}})}$ are smooth] $\boldsymbol{\Longrightarrow \#(\pi^{-1}(1) \cap \tilde{Z}) = \mathrm{Vol}(A)}$.**

So let us now assume the hypothesis of the last implication. By Lemma 4.2.8, Z (resp. $\tilde{Z} \cap \pi^{-1}(0)$) smooth $\Longrightarrow$ 1 (resp. 0) is not a critical value of $\pi|_{\tilde{Z}}$. Also, by the Implicit Function Theorem, the smoothness of $\tilde{Z}$ implies that $\pi(\tilde{Z})$ contains a small open ball about 1. So by the first part of Lemma 4.2.8, $\pi(\tilde{Z}) = \mathbb{P}^1_{\mathbb{C}}$.

Clearly, $\mathbb{P}^1_{\mathbb{C}}$ remains path-connected even after a finite set of points is removed, so let L be any continuous path connecting 0 and 1 in $\mathbb{P}^1_{\mathbb{C}} \setminus \mathrm{Crit}(\pi|_{\tilde{Z}})$. By the Implicit Function Theorem once more, and the fact that L is compact (by virtue of the compactness of $\mathbb{P}^1_{\mathbb{C}}$), we must have that $\#(\pi^{-1}(t) \cap \tilde{Z})$ is constant on L. **So we now need only show that $\boldsymbol{h\#(\pi^{-1}(0) \cap \tilde{Z}) = \mathrm{Vol}(A)}$.**

To conclude, note that $\tilde{A}$ and $\hat{A}$ have the same lower hull, so Lemmata 4.1.10 and 4.1.6 then imply that $\pi^{-1}(0) \cap \tilde{Z}$ is nothing more than

$$\{[p_1 : \cdots : p_{2N}] \in Y_{\tilde{A}} \mid \sum_{a_j \in Q} c_{i,a_j} p_j = 0 \text{ for all } i \in \{1, \ldots, n\} \text{ for some cell } Q \text{ of } A_\omega\}.$$

In particular, by our smoothness assumption on $\pi^{-1}(0) \cap \tilde{Z}$, Corollary 3.0.13 tells us that we can restrict to full-dimensional cells. Since A_ω is a triangulation, Corollary 3.0.13 and Lemma 4.1.6 tells us that

$$h\#(\pi^{-1}(0) \cap \tilde{Z}) = \sum_{Q \text{ a full-dimensional cell of } A_\omega} \mathrm{Vol}(Q) = \mathrm{Vol}(A),$$

so we are done. ■

Remark 4.2.10. *Our proof generalizes quite easily to arbitrary algebraically closed fields and positive characteristic, e.g., the algebraic closure of a finite field. One need only use a little algebra to extend Lemma 4.2.8 to algebraically closed fields (e.g.,* [**Sil95**, Ch. II, Sec. 2, Pg. 28, Prop. 2.6]*), and then one can use the same proof above almost verbatim.* ◇

Remark 4.2.11. *David N. Bernstein's seminal paper* [**Ber75**] *contains a proof of Kushnirenko's Theorem similar in spirit to ours. He instead used an elegant recursive construction (based on support functions) that allowed him to reduce the dimension and conclude by induction. His proof occupies less than half a page, so here we have made an effort to keep our proof self-contained and illustrate the underlying toric variety aspects which are useful elsewhere. We also note that his proof makes use of Puiseux series,*[11] *so it does not generalize trivially to positive characteristic.* ◇

We point out in closing that there are at least 3 main approaches to proving Kushnirenko's Theorem: (1) computing the degree of Y_A, (2) computing the multiplicity of a singular point of a related hypersurface, or (3) introducing a clever metric on Y_A and computing the volume of Y_A. Our proof is a combinatorial elaboration of (1), based on an approach pioneered in [**HS95**]. In particular, we have just made constructive all the non-degeneracy assumptions used in [**HS95**], and avoided the use of Puiseux series which wouldn't work in positive characteristic.

5. Path Following, Compactness, and Degenerate Kushnirenko

Let us now allow degeneracies for the zero set of F and prove the following strengthening of Kushnirenko's Theorem. Throughout this paper, we let $\mathbb{N}$ denote the **positive** integers.

Theorem 5.0.12. *Let F be any* **unmixed** *$n \times n$ polynomial system with common support $A = \{a_1, \ldots, a_N\}$, let H be the Hermite normal form of the $N \times n$ matrix whose $i^{\underline{\mathrm{th}}}$ row is a_i, and let h be the product of the diagonal elements of H. Also let Z_A be the zero set of F in Y_A, and let $\{Z_i\}$ be the collection of path-connected components of Z_A. Then there is a natural, well-defined positive*

[11] i.e., power series with fractional powers allowed.

intersection multiplicity $\mu : \{Z_i\} \longrightarrow \mathbb{N}$ *such that* $\sum_i \mu(Z_i) = \mathrm{Vol}(A)/h$ *and* $\mu(Z_i)=1$ *if* Z_i *is a non-degenerate root.*

We actually have all the technical preliminaries we'll need, except for one last simple proposition on path-connectedness. The proof follows easily by restricting to a (complex) line and reducing to the case $N=1$. (The latter special case follows easily by using a path consisting of just two line segments.)

Proposition 5.0.13. *If* $\mathcal{H}$ *is any algebraic hypersurface in* $\mathbb{C}^N$ *then* $B\backslash\mathcal{H}$ *is path-connected for any open ball* B *in* $\mathbb{C}^N$. ■

Proof of Theorem 5.0.12: Let $N := \#A$ as usual and note that the space of all polynomials in $\mathbb{C}[x_1, \ldots, x_n]$ with support contained in A can be identified with $\mathbb{C}^N$. Since zero sets of polynomials are unchanged under scaling of the coefficients, we will then let $(\mathbb{P}^{N-1}_{\mathbb{C}})^n$ be the space we'll use to consider our possible F. Let us also use Δ' to denote the image of $\Delta(\underbrace{A, \ldots, A}_{n})$ in $(\mathbb{P}^{N-1}_{\mathbb{C}})^n$.

Note now that if $F \in (\mathbb{P}^{N-1}_{\mathbb{C}})^n \backslash \Delta'$ then we are done by Theorem 4.2.9 (simply setting $\mu(Z_i)=1$ for every root Z_i). Indeed, since $(\mathbb{P}^{N-1}_{\mathbb{C}})^n \backslash \Delta'$ is path-connected by Proposition 5.0.13, the Implicit Function Theorem tells us that F must have the same number of roots in Y_A as any F with smooth zero set and no roots at toric infinity.

Essentially the same idea can be used for $F \in \Delta'$. In particular, for any such F, let $(F^{(i)}) \subset (\mathbb{P}^{N-1}_{\mathbb{C}})^n \backslash \Delta'$ be any sequence such that $F^{(i)} \longrightarrow F$. Then, letting $Z^{(i)}$ be the zero set of $F^{(i)}$ in Y_A, let ζ be any limit point of $\{Z^{(i)}\}$. By the continuity of $F(p)$ as a function of F and p, we must then have $F(\zeta)=0$ and thus Z_A must be non-empty.

Now let $\{U_i\}$ be disjoint open sets with $Z_i \subset U_i$ for all i. (Such a collection of open sets must exist since Y_A is compact and the Z_i must be of positive distance from each other, using the usual **Fubini-Study distance** in $\mathbb{P}^{N-1}_{\mathbb{C}}$.)[12] Note then that $Y_A \backslash \bigcup_i U_i$ must be compact. By the continuity of F as a function of its variables **and** coefficients, there must then be a ball B about F in $(\mathbb{P}^{N-1}_{\mathbb{C}})^n$ such that the roots of any $G \in B$ are contained in $\bigcup_i U_i$.

We may now define $\mu(Z_i)$ as follows: Take any $G \in B \backslash \Delta'$ and define $\mu(Z_i)$ to be the number of roots of G in U_i. Since $B\backslash\Delta'$ is path-connected by Proposition 5.0.13, the Implicit Function Theorem tells us that the number of roots is independent of whatever $G \in B \setminus \Delta'$ we picked.

To see that $\mu(Z_i)$ is always positive, let $V := \{(F,p) \in (\mathbb{P}^{N-1}_{\mathbb{C}})^n \times Y_A \mid F(p)=0\}$ and note that V is connected. (This follows easily by fibering over Y_A and using monomial curves, mimicking [**BCSS98**, Pg. 194]). Now let $\Sigma := \{(F,p) \in V \mid F \in \Delta'\}$ and $V_i := (B \times U_i) \cap V$ for all i. Then V_i is an open subset of V containing $B \times Z_i$. Since Σ is a proper subset of the connected set V, V_i cannot be contained in Δ' for any i. Hence the projection of V_i into B contains a non-empty open set. So there is indeed a $G \in B \setminus \Delta(\underbrace{A, \ldots, A}_{n})$ with at least one root in U_i. ■

[12] This is just the metric which assigns a distance of $\mathrm{ArcCos}\left(\frac{\langle x,y \rangle}{\|x\| \|y\|}\right)$ between any two points $x := [x_1 : \cdots : x_N]$ and $y := [y_1 : \cdots : y_N]$ in $\mathbb{P}^{N1}_{\mathbb{C}}$, where $\langle \cdot, \cdot \rangle$ (resp. $\|\cdot\|$) denotes the usual Hermitian inner product (resp. Hermitian norm) in $\mathbb{C}^N$.

Remark 5.0.14. *Our approach above is inspired by an elegant proof of an extended version of Bézout's Theorem by Mike Shub (see* [**Shu93**] *and* [**BCSS98**, Pg. 199]*). Note that neither theorem generalizes the other, but the theorems overlap in the special case where A is the set of integral points in a scaled standard simplex. On the other hand, Theorem 8.0.32 below will generalize both Kushnirenko's Theorem and Bézout's Theorem simultaneously. The elegance of Shub's approach is that it gives a rigourous and simple approach to intersection theory for a broad class of polynomial systems.* $\diamond$

Remark 5.0.15. *Combining Theorems 5.0.12 and 4.2.9 we immediately obtain our earlier, coarser statement of Kushnirenko's Theorem (Theorem 4.0.14).* $\diamond$

6. Multilinearity and Reducing Bernstein to Kushnirenko

The big question now is how to count the roots of a **mixed** polynomial system, since being unmixed is such a strong restriction. Toward this end, let us consider another consequence of the basic properties of discriminant varieties.

Lemma 6.0.16. *Let F and G be any $n\times n$ polynomial systems with support contained in $(A_1,\ldots,A_n)$ component-wise. Then, generically, F and G share no roots in $(\mathbb{C}^*)^n$. Furthermore, the number of roots of F is generically a fixed constant.* $\blacksquare$

As an immediate consequence, we can obtain a preliminary answer to our big question.

Definition 6.0.17. *Let $S_1,\ldots,S_k$ be any subsets of $\mathbb{R}^n$. Then their* **Minkowski sum** *is simply $S_1+\cdots+S_k:=\{y_1+\cdots+y_k \mid y_i\in S_i \text{ for all } i\}$.* $\diamond$

It is easily proved that $\mathrm{Newt}(fg)=\mathrm{Newt}(f)+\mathrm{Newt}(g)$, once one observes that the vertices of $\mathrm{Newt}(fg)$ are themselves Minkowski sums of vertices of $\mathrm{Newt}(f)$ and $\mathrm{Newt}(g)$. So it should come as no surprise that Minkowski sums will figure importantly in our discussion relating polyhedra and polynomials.

Lemma 6.0.18. *Let $\mathcal{N}(A_1,\ldots,A_n)$ denote the generic number of roots in $(\mathbb{C}^*)^n$ of an $n\times n$ polynomial system F with support $(A_1,\ldots,A_n)$. Then $\mathcal{N}(A_1,\ldots,A_n)$ is a non-negative symmetric function of $\mathrm{Conv}(A_1),\ldots,\mathrm{Conv}(A_n)$ which is multilinear with respect to Minkowski sum.*

Proof: That $\mathcal{N}(A_1,\ldots,A_n)$ is a well-defined non-negative symmetric function of $A_1,\ldots,A_n$ is clear, thanks to the last part of Lemma 6.0.16. The formula for $\mathcal{N}(A_1,\ldots,A_n)$ in the unmixed case then follows immediately from Theorem 4.2.9. Translation invariance follows easily since the roots of F in $(\mathbb{C}^*)^n$ are the same as the roots of $(x^{a_1}f_1,\ldots,x^{a_n}f_n)$ in $(\mathbb{C}^*)^n$. Defining $x^{[u_{ij}]}:=(x_1^{u_{11}}\cdots x^{u_{n1}},\ldots,x_1^{u_{1n}}\cdots x^{u_{nn}})$ for any $n\times n$ matrix $[u_{ij}]$, it is then easily checked that $\mathrm{Supp}(F(x^U))=(UA_1,\ldots,UA_n)$, and that U **unimodular** (cf. Definition 3.0.6) implies that the map $x\mapsto x^U$ is an analytic bijection of $(\mathbb{C}^*)^n$ (with analytic inverse) into itself. So it is clear that $\mathcal{N}(UA_1,\ldots,UA_n)=\mathcal{N}(A_1,\ldots,A_n)$.

We thus need only show that $\mathcal{N}$ is a multilinear function of the volumes of the convex hulls of $A_1,\ldots,A_n$. To see the multilinearity, note that the zero set of $(f_1\bar{f}_1,f_2,\ldots,f_n)$ in $(\mathbb{C}^*)^n$ is exactly the union of the zero sets of $(f_1,f_2,\ldots,f_n)$ and $(\bar{f}_1,f_2,\ldots,f_n)$. So by the first part of Lemma 6.0.16, and the symmetry of $\mathcal{N}$, multlinearity follows. However, the dependence on $\mathrm{Conv}(A_1),\ldots,\mathrm{Conv}(A_n)$ alone is not yet clear.

So recall now the **polarization identity**:

$$n!m(x_1,\ldots,x_n) = \sum_{\emptyset \neq I \subseteq \{1,\ldots,n\}} (-1)^{n-\#I} m\left(\sum_{i\in I} x_i, \ldots, \sum_{i\in I} x_i\right),$$

valid for any symmetric multilinear function. (The identity is not hard to prove via inclusion-exclusion [**GKP94**]. See also [**Gol03**] in this volume for another point of view.) Therefore, we must have

$$n!\mathcal{N}(A_1,\ldots,A_n) = \sum_{\emptyset \neq I \subseteq \{1,\ldots,n\}} (-1)^{n-\#I} \mathcal{N}\left(\sum_{i\in I} A_i, \ldots, \sum_{i\in I} A_i\right),$$

and thus $\mathcal{N}(A_1,\ldots,A_n)$ depends only the convex hulls of $A_1,\ldots,A_n$, thanks to Kushnirenko's Theorem. ■

So we have answered our big question, assuming we know a function $\mathcal{M}(P_1,\ldots,P_n)$, defined on n-tuples $(P_1,\ldots,P_n)$ of polytopes in $\mathbb{R}^n$, that satisfies the obvious analogues of the properties of $\mathcal{N}(A_1,\ldots,A_n)$ specified in Lemma 6.0.18. However, such a function indeed exists: it is called the **mixed volume** and we denote it by $\mathcal{M}(\cdot)$. Abusing notation slightly by setting $\mathcal{M}(A_1,\ldots,A_n) := \mathcal{M}(\mathrm{Conv}(A_1),\ldots,\mathrm{Conv}(A_n))$, we immediately obtain the following result.

THEOREM 6.0.19 (Bernstein's Theorem). *Suppose F is any $n \times n$ polynomial system with fixed support $A_1,\ldots,A_n$. Then F generically has exactly $\mathcal{M}(A_1,\ldots,A_n)$ roots in $(\mathbb{C}^*)^n$.* ■

Of course, we now appear to have an even bigger question: what is mixed volume? This we now answer.

REMARK 6.0.20

Ferdinand Minding
1806–1885

Hermann
Minkowski[13]
1864–1909

Bernstein's original proof in [**Ber75**] *makes a similar reduction to the unmixed case. There he also derived an algebraic criterion for when the number of roots is exactly the mixed volume. (Here, such a criterion is implicit in our definitions of initial term systems and Lemma 4.1.10.) It is worth noting that his paper is just 3 pages long. Interestingly, Ferdinand Minding appears to have been the first to prove the special case $n=2$ in 1841, and mixed volume wasn't even defined until near the end of the 19th century by Hermann Minkowski.* ◇

7. Mixed Subdivisions and Mixed Volumes from Scratch

There are many different definitions of mixed volume but the two most important use Minkowski sums in an essential way. More to the point, if one can subdivide $P_1 + \cdots + P_n$ in a special way, then one is well on the way to computing mixed volume. This is where **mixed subdivisions** enter.

DEFINITION 7.0.21. [**HS95**] *Given polytopes $P_1,\ldots,P_k \subset \mathbb{R}^n$, a* **subdivision of $(P_1,\ldots,P_k)$** *is a finite collection of k-tuples $\{(C_1^\alpha,\ldots,C_k^\alpha)\}_{\alpha\in S}$ satisfying the following axioms:*

[13] We also point out that Minkowski was born on 22 June, 1864, in a town named Alexotas (Aleksotas in Lithuanian), on the left bank of the river Nemunas. This town, founded around the 15th century, belonged to Prussia from 1795 and from 1814 to 1918 belonged to what was the Russian empire at the time. In 1931 Alexotas became a district in the Lithuanian city of Kaunas, temporary capital of Lithuania.

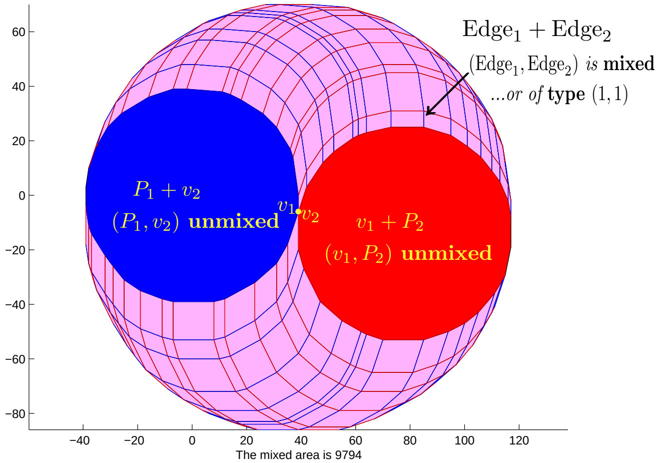
(1) $\bigcup_{\alpha\in S} C_i^\alpha = P_i$ *for all* i

(2) $C_i^\alpha \cap C_i^\beta$ *is a face of both* C_i^α *and* C_i^β *for all* α, β, i.

(3) C_i^β *a face of* C_i^α *for all* $i \Longrightarrow$ *there is a* $w\in\mathbb{R}^n$ *such that* $C_i^\beta = (C_i^\alpha)^w$ *for all* i.

Furthemore, if we have in addition that $\sum_i \dim C_i^\alpha = \dim \sum_i C_i^\alpha$ *for all* α, *then we call* $\{(C_1^\alpha, \ldots, C_k^\alpha)\}_{\alpha\in S}$ *a* **mixed subdivision**. $\diamond$

Example 7.0.22.



Here we see a very special kind of subdivision $\{Q_i\}$ *of the Minkowski sum of two polygons* P_1 *and* P_2, *each with many vertices. In particular, the subdivision of* $P_1 + P_2$ *above is built in such a way as to encode a* **mixed** *subdivision* $\{(C_1^\alpha, C_2^\alpha)\}$ *of the pair* (P_1, P_2). *We also see that each* P_i *has a distinguished vertex* v_i, *and that we can read off a mixed subdivision of* (P_1, P_2) *as follows: there are two cells* (P_1, v_2) *and* (v_1, P_2), *corresponding to the two cells* $P_1 + v_2$ *and* $v_1 + P_2$ *of* $\{Q_i\}$. *The remaining cells of* $\{(C_1^\alpha, C_2^\alpha)\}$ *are of the form* (E_1, E_2) *where* E_i *is an edge of* P_i *for all* i. *In particular, all but two of the* 2*-dimensional cells of* $\{Q_i\}$ *are parallelograms.* $\diamond$

Definition 7.0.23. *Following the notation above, the* **type** *of a cell* $(C_1^\alpha, \ldots, C_k^\alpha)$ *of a subdivision of* $(P_1, \ldots, P_n)$ *is simply the vector* $(\dim C_1^\alpha, \ldots, \dim C_k^\alpha)$. *In particular, the cells of type* $(1, \ldots, 1)$ *are called* **mixed cells**. $\diamond$

It is easily verified that any subdivision of $(P_1, \ldots, P_k)$ immediately induces a subdivision of $(\lambda P_1, \ldots, \lambda P_k)$, and vice-versa, for any $\lambda_1, \ldots, \lambda_k \geq 0$. In particular, note that the volume of a cell from a subdivision of $\lambda P_1 + \cdots + \lambda P_k$ induced by a mixed subdivision of $(P_1, \ldots, P_k)$ scales — as a function of $\lambda_1, \ldots, \lambda_k$ — according to its type.

The first lemma below follows from a slight modification of the proof of Lemma 4.1.4, while the second lemma follows directly from the first, thanks to the existence of mixed subdivisions.

Lemma 7.0.24. *Following the notation of Definition 4.1.2, recall that* $\pi : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}^n$ *is the natural projection which forgets the last coordinate. Then, given finite point sets* $A_1, \ldots, A_n \subset \mathbb{Z}^n$ *and lifting functions* ω_i *for* A_i *for all* i, *the collection* $(A_1, \ldots, A_n)_\omega :=$ $\{(\pi(\text{Conv}(\hat{A}_1)^{(v,1)}), \ldots, \pi(\text{Conv}(\hat{A}_n)^{(v,1)})) \mid v \in \mathbb{R}^n\}$ *always forms a subdivision of* $(\text{Conv}(A_1), \ldots, \text{Conv}(A_n))$ — *the* **subdivision of** $\boldsymbol{(\text{Conv}(A_1), \ldots, \text{Conv}(A_n))}$ **induced by** $\boldsymbol{(\omega_1, \ldots, \omega_n)}$. *In particular, for fixed* $(A_1, \ldots, A_n)$, $(A_1, \ldots, A_n)_\omega$ *will* *generically*[14] *be a* **mixed** *subdivision. In this case, we say that* $(\omega_1, \dots, \omega_n)$ *is an* n*-tuple of* **generic lifting functions** *for* $(A_1, \dots, A_n)$. ■

Lemma 7.0.25. *For* $\lambda_1 \dots, \lambda_n \geq 0$*, and any polytopes* $P_1, \dots, P_n \subset \mathbb{R}^n$*, the quantity* $\mathrm{Vol}\left(\sum_{i=1}^n \lambda_i P_i\right)$ *is a homogeneous polynomial of degree* n *with non-negative coefficients.* ■

We then at last arrive at the following definition of the mixed volume.

Definition 7.0.26. *Given any polytopes* $P_1, \dots, P_n \subset \mathbb{R}^n$*, their mixed volume is the coefficient of* $\lambda_1 \lambda_2 \cdots \lambda_n$ *in the polynomial* $\mathrm{Vol}'\left(\sum_{i=1}^n \lambda_i P_i\right)$*, where* Vol' *denotes volume normalized so that the standard unit* n*-cube has volume* 1. ⋄

Example 7.0.27 (The Unmixed Case). *It is easily checked that* $\mathcal{M}(P, \dots, P) = \mathrm{Vol}(P)$*. Note also that the multilinearity of* $\mathcal{M}(\cdot)$ *with respect to Minkowski sum also follows immediately from the preceding definition.* ⋄

Example 7.0.28 (Line Segments). *It is also easily checked that* $\mathcal{M}(\{0, a_1\}, \dots, \{0, a_n\}) = |\det[a_1, \dots, a_n]|$*, where* $a_1, \dots, a_n$ *are any points in* $\mathbb{R}^n$ *and* $[a_1, \dots, a_n]$ *is the matrix whose columns are* $a_1, \dots, a_n$. ⋄

Example 7.0.29 (Bézout's Theorem). *Taking* $d_1, \dots, d_n \in \mathbb{N}$ *and* $P_i = d_i \mathrm{Conv}(\{\mathbf{0}, e_1, \dots, e_n\})$ *for all* i*, it is easily checked by multilinearity that* $\mathcal{M}(P_1, \dots, P_n) = \prod_{i=1}^n d_i$*. So, modulo roots on the coordinate hyperplanes or at the hyperplane at projective infinity, Bernstein's Theorem includes Bézout's Theorem as a special case. Alternatively, if we use toric varieties, Bernstein's Theorem contains Bézout's Theorem without qualification.* ⋄

The next two characterizations follow easily from the last lemma, and inclusion-exclusion [**GKP94**].

Lemma 7.0.30. *For any mixed subdivision* $\{(C_1^\alpha, \dots, C_n^\alpha)\}$ *of* $(P_1, \dots, P_n)$*, we have* $\mathcal{M}(P_1, \dots, P_n) := \sum_{\substack{(C_1, \dots, C_n) \\ \text{a cell of type } (1, \dots, 1)}} \mathrm{Vol}'\left(\sum_i C_i\right)$*. Furthermore, we have*

$$\mathcal{M}(P_1, \dots, P_n) := \sum_{\emptyset \neq I \subseteq \{1, \dots, n\}} (-1)^{n - \#I} \mathrm{Vol}'\left(\sum_{i \in I} P_i\right). \quad \diamond$$

Example 7.0.31 (Bricks, a.k.a. the fine multigraded case). *Via multilinearity, it easily follows that* $\mathcal{M}([0, d_{11}] \times \cdots \times [0, d_{1n}], \dots, [0, d_{n1}] \times \cdots \times [0, d_{nn}]) = \mathrm{Perm}[d_{ij}]$*, where* Perm *denotes the* **permanent**.[15] *In particular, this immediately shows that computing mixed volume is* $\#\mathbf{P}$*-hard* [**Pap95, DGH98**]*. In can also be shown that mixed volume computation is in the complexity class* $\#\mathbf{P}$ [**DGH98**]. ⋄

Let us now prove Main Theorem 2.

Proof of Main Theorem 2: Note that by Bernstein's Theorem, it suffices to find an algorithm for computing $\mathcal{M}(A_1, A_2)$ with arithmetic complexity $O(\bar{N} \log \bar{N})$. The main idea of the proof can then already be visualized in the first mixed subdivision we illustrated: one computes the mixed area of (A_1, A_2) by first efficiently

[14] The genericity of ω is of course in the sense of Lemma 4.1.4: there is a finite collection of $(N-1)$-flats, where $N := \#A_1 + \cdots + \#A_n$, such that $\omega \in \mathbb{R}^N \setminus \mathcal{H} \Longrightarrow (A_1, \dots, A_n)_\omega$ is a mixed subdivision.

[15] Recall that this function can be defined as the variant of the determinant where all alternating signs in the full determinant expansion are replaced by +1's.

computing the convex hulls of A_1 and A_2, and then computing the sum of the areas of the mixed cells. However, one must represent this sum of areas **compactly and without building the entire mixed subdivision**. This is quite possible, provided one views the mixed cells in the right way.

More precisely, first recall that the convex hulls of A_1 and A_2 can be computed within $O(\bar{N}\log\bar{N})$ arithmetic operations, via the usual well-known 2-dimensional convex hull algorithms [**PS85**]. In particular, with this much work, we can already assume we know the inner edge normals of $P_1 := \mathrm{Conv}(A_1)$ and $P_2 := \mathrm{Conv}(A_2)$, and the vertices of P_1 and P_2 in counter-clockwise order.

Recall that an **angle cone** of a vertex v is the cone generated by the edge vectors emanating from v. Let us then pick vertices $v_1 \in P_1$ and $v_2 \in P_2$ such that their angle cones intersect only at the origin. Then there is a mixed subdivision (which we will never calculate explicitly!) with exactly 2 non-mixed cells — (P_1, v_2) and (v_1, P_2) — and several other mixed cells.[16]

Note then that the union of the mixed cells can be partition into a union of strips. In particular, by construction, there are disjoint contiguous sequences of edges $(E^{(i)}_1, \ldots, E^{(i)}_{a_i})$ and $(E^{(i')}_1, \ldots, E^{(i')}_{a_{i'}})$, with $E^{(i)}_1$ and $E^{(i')}_1$ incident to v_i, for all i and i'. Furthermore, every mixed cell of $(A_1, A_2)_\omega$ is of the form $(E^{(1)}_i, E^{(2)}_j)$ or $(E^{(1')}_i, E^{(2')}_j)$.

The partition into strips then arises as follows: the mixed cells of $(A_1, A_2)_\omega$ can be partitioned into lists of one of the following two forms:

$$(E^{(1)}_j, E^{(2)}_{m_j}), \ldots, (E^{(1)}_1, E^{(2)}_{n_j}) \qquad \text{or} \qquad (E^{(1')}_j, E^{(2')}_{m_j}), \ldots, (E^{(1')}_1, E^{(2')}_{n_j}),$$

where $j \in \{1, \ldots, a_1\}$ (resp. $j \in \{1, \ldots, a_{1'}\}$), $m_j \leq n_j$, and $n_j \leq a_2$ (resp. $n_j \leq a_{2'}$). In particular, the union of the mixed cells in any such list is simply the Minkowski sum of a contiguous portion of the boundary of P_2 and an edge of P_1, and its area can thus be expressed as the absolute value of a determinant of differences of vertices of the P_i. Furthermore, these formulae can easily be found by a binary search on the sorted edge normals using a total of $O(\bar{N}\log\bar{N})$ comparisons.

Since there are no more than $\bar{N}$ such strips, the total work we do is bounded above by the specified complexity bound, so our upper bound is proved.

To obtain our lower bound, note that the mixed area of (A_1, A_2) is zero iff [[P_1 or P_2 is a point] or [P_1 and P_2 are parallel line segments]]. So just knowing whether the mixed area is positive or not amounts to a rank computation on a matrix of size $O(\bar{N})$ and thus requires no less than $\Omega(N)$ arithmetic operations in the worst case [**BCS97**]. ■

8. A Stronger Bernstein Theorem Via Mixed Subdivisions

Here we prove two generalization of Theorem 5.0.12: Theorem 8.0.32 and the full version of Main Theorem 1. Having introduced all the necessary background, our proofs will be simply be minor modifications of the earlier proofs of our earlier extensions of Kushnirenko's Theorem.

Theorem 8.0.32. *Following the notation of Theorem 6.0.19, let $A := A_1 + \cdots + A_n$, let Z_A be the zero set of F in Y_A, and let $\{Z_i\}$ be the collection of path-connected components of Y_A. Then there is a natural, well-defined positive*

[16] This is easily seen by picking a lifting function ω_1 for P_1 that is identically zero, and a non-constant linear lifting function ω_2 for P_2 that is 0 at v_2, constant on a line that intersects the angle cones of v_1 and v_2 only at the origin, and non-negative on P_2.

intersection multiplicity[17] $\mu : \{Z_i\} \longrightarrow \mathbb{N}$ *such that* $\sum_i \mu(Z_i) = \mathcal{M}(A_1, \ldots, A_n)$ *and* $\mu(Z_i) = 1$ *if* Z_i *is a non-degenerate root.*

The proof will be almost exactly the same as that of our extended version of Kushnirenko's Theorem, so let us first see an illustration of a toric deformation for a **mixed** system.

EXAMPLE 8.0.33. *Take* $n = 2$ *and*

$$f_1(x,y) := c_{1,\mathbf{O}} + c_{1,(\alpha,0)}x^\alpha + c_{1,(0,\beta)}y^\beta + c_{1,(\alpha,\beta)}x^\alpha y^\beta$$
$$f_2(x,y) := c_{2,\mathbf{O}} + c_{2,(\gamma,0)}x^\gamma + c_{2,(0,\delta)}y^\delta + c_{2,(\gamma,\delta)}x^\gamma y^\delta.$$

By Bernstein's Theorem, the number of roots should be $\alpha\delta + \beta\gamma$, *so let us try to prove this.*

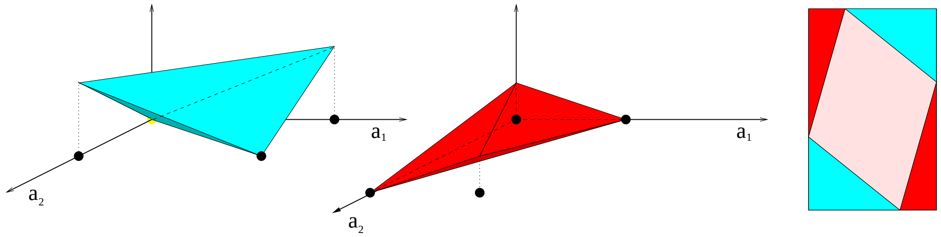
Let us take the following lifting of F*:*

$$\hat{f}_1(x,y,t) := c_{1,\mathbf{O}} + c_{1,(\alpha,0)}x^\alpha t + c_{1,(0,\beta)}y^\beta t + c_{1,(\alpha,\beta)}x^\alpha y^\beta$$
$$\hat{f}_2(x,y) := c_{2,\mathbf{O}}t + c_{2,(\gamma,0)}x^\gamma + c_{2,(0,\delta)}y^\delta + c_{2,(\gamma,\delta)}x^\gamma y^\delta t$$

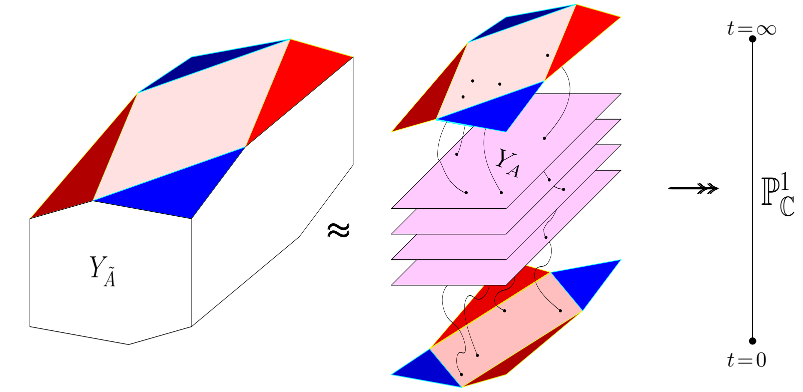
In particular, we see that there will be exactly one mixed cell for $(A_1, A_2)_\omega$ *and its corresponding initial term system will be*

$$\mathrm{Init}_{(0,0,1)}(\hat{F})(x,y,t) = (c_{1,\mathbf{O}} + c_{1,(\alpha,\beta)}x^\alpha y^\beta, c_{2,(\gamma,0)}x^\gamma + c_{2,(0,\delta)}y^\delta)$$

The lifted Newton polytopes and induced subdivisions appear below.



The idea of our proof of Bernstein's Theorem then mimics our earlier proof of Kushnirenko's Theorem: our lifting induces a lifted version $\hat{A} := \hat{A}_1 + \hat{A}_2$ *of* $A := A_1 + A_2$ *and we'll then try to build a map from our lifted zero set to the projective line. To do so, we'll define* $\tilde{A} := \hat{A} \times \{0,1\}$ *and this is illustrated below.*



In particular, the only portion of the lower hull of $\tilde{A}$ *(i.e., the "lower portion" of toric infinity on* $Y_{\tilde{A}}$*) which is touched by the zero set of* $\hat{F}$ *in* $Y_{\tilde{A}}$ *is the parallelogram facet, and the projection of this facet has area exactly* $\alpha\delta + \beta\gamma$. ⋄

[17] Another version of Bernstein's Theorem which took intersection multiplicity into account appeared earlier in [**Dan78**]. However, the proof there requires considerably more machinery than our approach here.

Proof of Theorem 8.0.32: We will first prove the generic case, and then derive the degenerate case just as we did for the unmixed case.

At this point, we could just use Theorem 6.0.19 to get the generic case and proceed with our proof of the degenerate case. However, let us observe that we could instead use mixed subdivisions to directly obtain Theorem 6.0.19 without reducing to the unmixed case. The proof proceeds exactly like the proof of Theorem 4.0.14, except for the following differences:

(1) The map π is instead defined by a ratio of coordinates depending on the lifting ω of $A_1, \ldots, A_n$.
(2) The only portions of toric infinity in $Y_{\tilde{A}}$ that intersect $\pi^{-1}(0) \cap \tilde{Z}$ are those corresponding to facets on the lower hull of $\tilde{P}$ that project to **mixed cells** of $(A_1, \ldots, A_n)_\omega$.
(3) The final count of roots becomes a sum of the number of roots of a collection of **binomial** systems.

To prove the degenerate case, we then proceed exactly as in the proof of Theorem 5.0.12, except with the following minor modifications: (1) We use the notational changes above, and (2) the space of F we work with is instead $\mathbb{P}^{N_1-1}_{\mathbb{C}} \times \cdots \times \mathbb{P}^{N_n-1}_{\mathbb{C}}$, where $N_i = \#A_i$ for all i. ■

We are now finally ready to state and prove the full version of Main Theorem 1:

Main Theorem 1 (Full Version) *Suppose* $F = (f_1, \ldots, f_k)$ *is any* $k \times n$ *polynomial system with* $\mathrm{Supp}(f_i) \subset (\mathbb{N} \cup \{0\})^n$ *for all* i *and let* $Z_{\mathbb{C}}(F)$ *denote the zero set of* F *in* $\mathbb{C}^n$*. For all* $(i,j) \in \{1, \ldots, k\} \times \{1, \ldots, n\}$*, let* $s_{ij} := \min\limits_{(a_1,\ldots,a_n) \in \mathrm{Supp}(f_i)} a_j$ *and let* t_{ij} *be* $s_{ij} - 1$ *or* 0 *according as* s_{ij} *is positive or not. Finally, let* $A'_i := \mathrm{Supp}(f_i) - (t_{i1}, \ldots, t_{in})$ *for all* i*.*

Then the number of connected components of $Z_{\mathbb{C}}(F)$*, counting multiplicities,*[18] *is no more than* $\mathrm{Vol}\left(\{\mathbf{O}, e_1, \ldots, e_n\} \cup \bigcup_{i=1}^k A'_i\right)$*, and an improved bound of* $\mathcal{M}\left(\{\mathbf{O}, e_1\} \cup A'_1, \ldots, \{\mathbf{O}, e_n\} \cup A'_n\right)$ *holds when* $k = n$*. In particular, when* $k = n$ *and* $\{\mathbf{O}, e_i\} \subseteq \mathrm{Supp}(f)_i$ *for all* i*, the latter bound is tight.*

Proof: Let $f'_i := x_1^{-t_{i1}} \cdots x_n^{-t_{in}} f_i$ for all i and $F' := (f'_1, \ldots, f'_k)$. Then, by the definition of the t_{ij}, F and F' clearly have the **same** zero set in $\mathbb{C}^n$. So it suffices to work with F' instead of F.

Now let $B' := \{\mathbf{O}, e_1, \ldots, e_n) \cup \bigcup_{i=1}^k A'_i$. If $k < n$ then simply set $f'_{k+1} = \cdots = f'_n = f'_1$. We can then apply Theorem 5.0.12, noting that the $\mathrm{Supp}(f'_i) \subseteq B'$ for all i. In particular, it is easily checked that $Y_{B'}$ actually contains an embedded copy of $\mathbb{C}^n$, so the first bound is now proved for $k \leq n$.

To prove the case $k > n$, let $g_1, \ldots, g_n$ be n generic linear combinations of the f'_i. Clearly, $Z_{\mathbb{C}}(f'_1, \ldots, f'_k) \subseteq Z_{\mathbb{C}}(g_1, \ldots, g_n)$, and by Theorem 4.2.4 it is not difficult to show that $Z_{\mathbb{C}}(g_1, \ldots, g_n) \backslash \mathbb{Z}_{\mathbb{C}}(f'_1, \ldots, f'_k)$ is generically a finite set of points (see, e.g., [**GH93**, Sec. 3.4.1] for a complete proof). So we can assume $k = n$ and proceed

[18] It can be shown that — off the coordinate hyperplanes — our homotopically defined intersection multiplicity agrees with the more high-powered definition from intersection theory when $k = n$, provided one sums over the **distinguished** components lying in a given connected component [**Ful98**, Ch. 7]. (For connected components lying in the coordinate hyperplanes, our intersection can be less than the algebraic geometry definition (but still of the same sign), depending on the distances of the supports to the coordinate hyperplanes.) However, for $k > n$, the definition from intersection theory no longer applies, while our multiplicity remains positive.

as in the last paragraph (using Theorem 8.0.32 instead of Theorem 5.0.12) to obtain our second bound.

That the second bound is bounded above by the first follows immediately from the monotonicity of the mixed volume. The last statement of the theorem then follows easily from the fact that $Y_{\bar{A}}$ contains an embedded copy of $\mathbb{C}^n$, and the sharpness of Bernstein's Theorem.

The final technicality to take care of is the disruption in connectivity in passing from $Y_{\tilde{A}}$ to $\mathbb{C}^n$. However, this can be handled easily by altering our earlier proof of Theorem 5.0.12 to work in a large compact subset S of $\mathbb{C}^n$. Letting S tend to $Y_{\tilde{A}}$, we recover all possible connected components in $\mathbb{C}^n$ and preserve the bound we had for connected components in $Y_{\tilde{A}}$. So we are done. ■

REMARK 8.0.34. *A beautiful exposition by Askold Khovanski on Bernstein's Theorem can be found in* [**BZ88**, Ch. 4, Sec. 27, Addendum 3]*. The approach there is philosophically quite similar to ours but has some differences. For instance, while Khovanski avoids fractional power series as we do, he uses a special lemma on the intersection of space curves with hypersurfaces to reduce the dimension and conclude by induction. Also, while he mentions intersection multiplicity briefly, his theorems do not address degenerate polynomial systems. He also avoids the construction of toric varieties by resorting to Riemann surfaces to compactify his curves.* ◇

REMARK 8.0.35. *The problem of tightly estimating the number of roots in $\mathbb{C}^n$ (as opposed to $(\mathbb{C}^*)^n$) was never quite directly addressed until the 1990's. It was at least observed in the late 1970's by Khovanski that adding the origin to the supports and using the Newton polytopes so modified instead yields a formula for the generic number of affine roots. Tight general upper bounds for the affine case, along with explicit algebraic conditions for exactness, finally appeared in* [**RW96, HS97, Roj99a**]. ◇

Acknowledgements

I was truly overwhelmed by Ron Goldman and Rimvydas Krasauskas' kind and generous invitation to speak in Vilnius. I also thank David Cox, Ron Goldman, Miriam Lucian, Jorg Peters, Frank Sottile, and Joe Warren for some wonderful discussions. Finally, special thanks go to David Cox, Rimvydas Krasauskas, and an anonymous referee for detailed comments on an earlier version of this paper, and to Amit Khetan for generously providing me the 249 × 249 resultant matrix stated in Section 4.2.

In early 1995, Ilya Itenberg asked me for an explanation of Bernstein's Theorem. Regretably, I mistook explanation for proof. So I hope the present paper provides explanation as well as proof. I thus thank Ilya Itenberg for inspiring me to think more about why certain facts are true.

References

[Ber75] Bernstein, David Naumovich, *"The Number of Roots of a System of Equations,"* Functional Analysis and its Applications (translated from Russian), Vol. 9, No. 2, (1975), pp. 183–185.

[BCSS98] Blum, Lenore; Cucker, Felipe; Shub, Mike; and Smale, Steve, *Complexity and Real Computation,* Springer-Verlag, 1998.

[BCS97] Bürgisser, Peter; Clausen, Michael; and Shokrollahi, M. Amin, *Algebraic complexity theory,* with the collaboration of Thomas Lickteig, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 315, Springer-Verlag, Berlin, 1997.

[BZ88] Burago, Yu. D. and Zalgaller, V. A., *Geometric Inequalities,* Grundlehren der mathematischen Wissenschaften 285, Springer-Verlag (1988).

[BEM03] Buse, Laurent; Elkadi, Mohammed; and Mourrain, Bernard, *"Using Projection Operators in Computer Aided Geometric Design,"* this volume, pp. 321–342.

[Cox03] Cox, David A., *"What is a Toric Variety,"* this volume, pp. 203–223.

[Dan78] Danilov, V. I., *"The Geometry of Toric Varieties,"* Russian Mathematical Surveys, 33 (2), pp. 97–154, 1978.

[EGA1] Dieudonné, Jean and Grothendieck, Alexander, *Éléments de géométrie algébrique I: Le langage des schémas,* Inst. Hautes Études Sci. Publ. Math. No. 4, 1960.

[DGH98] Dyer, Martin; Gritzmann, Peter; and Hufnagel, Alexander, *"On the Complexity of Computing Mixed Volumes,"* SIAM J. Comput. **27** (1998), no. 2, pp. 356–400.

[Emi94] Emiris, Ioannis Z., *"Sparse Elimination and Applications in Kinematics,"* Ph.D. thesis, Computer Science Division, U. C. Berkeley (December, 1994), available on-line at http://cgi.di.uoa.gr/~emiris/publis.html .

[EC00] Emiris, Ioannis Z. and Canny, John, *"A Subdivision-Based Algorithm for the Sparse Resultant,"* J. ACM **47** (2000), no. 3, pp. 417–451.

[EP02] Emiris, Ioannis Z. and Pan, Victor Y., *"Symbolic and Numeric Methods for Exploiting Structure in Constructing Resultant Matrices,"* Journal of Symbolic Computation, Vol. 33, No. 4, April 1, 2002.

[FH95] Forsythe, Keith and Hatke, Gary, *"A Polynomial Rooting Algorithm for Direction Finding,"* preprint, MIT Lincoln Laboratories, 1995.

[Ful98] Fulton, William, *Intersection Theory*, 2nd ed., Ergebnisse der Mathematik und ihrer Grenzgebiete 3, **2**, Springer-Verlag, 1998.

[Ful93] Fulton, William, *Introduction to Toric Varieties*, Annals of Mathematics Studies, no. 131, Princeton University Press, Princeton, New Jersey, 1993.

[Gat01] Gatermann, Karin, *"Counting Stable Solutions of Sparse Polynomial Systems in Chemistry,"* Contemporary Mathematics, vol. 286, AMS-IMS-SIAM Joint Summer Research Conference Proceedings of "Symbolic Computation: Solving Equations in Algebra, Geometry, and Engineering (June 11–15, 2000, Mount Holyoke College)," edited by R. Laubenbacher and V. Powers, pp. 53–69, AMS Press, 2001.

[GKZ94] Gel'fand, I. M., Kapranov, M. M., and Zelevinsky, A. V., *Discriminants, Resultants and Multidimensional Determinants,* Birkhäuser, Boston, 1994.

[GH93] Giusti, Marc and Heintz, Joos, *"La détermination des points isolés et la dimension d'une variété algébrique peut se faire en temps polynomial,"* Computational Algebraic Geometry and Commutative Algebra (Cortona, 1991), Sympos. Math. XXXIV, pp. 216–256, Cambridge University

[Gol03] Goldman, Ron, *"Polar Forms in Geometric Modelling and Algebraic Geometry,"* this volume, pp. 3–24.

[GKP94] Graham, R. L., Knuth, D. E., and Patashnik, O., *Concrete Mathematics: A Foundation for Computer Science*, 2nd edition, Addison-Wesley, 1994.

[GH94] Griffiths, Phillip and Harris, Joseph, *Principles of Algebraic Geometry,* Reprint of the 1978 original, Wiley Classics Library, John Wiley & Sons, Inc., New York, 1994.

[Har77] Hartshorne, Robin, *Algebraic Geometry,"* Graduate Texts in Mathematics, No. 52, Springer-Verlag.

[HS95] Huber, Birk and Sturmfels, Bernd, *"A Polyhedral Method for Solving Sparse Polynomial Systems,"* Math. Comp. **64** (1995), no. 212, pp. 1541–1555.

[HS97] ________________________________, *"Bernstein's Theorem in Affine Space,"* Discrete and Computational Geometry, **17** (1997), no. 2, pp. 137–141.

[Ili89] Iliopoulos, Costas S., *"Worst Case Complexity Bounds on Algorithms for Computing the Canonical Structure of Finite Abelian Groups and the Hermite and Smith Normal Forms of an Integer Matrix,"* SIAM Journal on Computing, 18 (1989), no. 4, pp. 658–669.

[Jac85] Jacobson, Nathan, *Basic Algebra I*, 2nd edition, W. H. Freeman and Company, 1985.

[JKSS03] Jeronimo, Gabriela; Krick, Teresa; Sabia, Juan; and Sombra, Martin, *"The Computational Complexity of the Chow Form,"* Math ArXiV preprint math.AG/0210009.

[Kap00] Kapranov, Mikhail M., *"Amoebas Over Non-Archimedean Fields,"* manuscript, University of Toronto, 2000.

[KM97] Karpinski, Marek and Macintyre, Angus J., *"Polynomial bounds for VC dimension of sigmoidal and general Pfaffian neural networks," J. Comp. Sys. Sci.*, 54, pp. 169–176, 1997.

[Khe03] Khetan, Amit, *"Formulas for Resultants",* Ph.D. thesis, U. C. Berkeley, 2003.

[Kho91] Khovanski, Askold Georgievich, *Fewnomials,* AMS Press, Providence, Rhode Island, 1991.

[Kus77] Kushnirenko, Anatoly Georgievich, *"Newton Polytopes and the Bézout Theorem,"* Functional Analysis and its Applications (translated from Russian), vol. 10, no. 3, July–September (1977), pp. 233–235.

[Li97] Li, Tien Yien, *"Numerical solution of multivariate polynomial systems by homotopy continuation methods,"* Acta numerica, 1997, pp. 399–436, Acta Numer., 6, Cambridge Univ. Press, Cambridge, 1997.

[LRW03] Li, Tien-Yien; Rojas, J. Maurice; and Wang, Xiaoshen, *"Counting Real Connected Components of Trinomial Curves Intersections and m-nomial Hypersurfaces,"* Discrete and Computational Geometry, to appear.

[MR03] Malajovich, Gregorio and Rojas, J. Maurice, *"High Probability Analysis of the Condition Number of Sparse Polynomial Systems,"* Theoretical Computer Science, special issue on algebraic and numerical algorithms, to appear.

[Man98] Manocha, Dinesh, *"Numerical Methods for Solving Polynomial Equations,"* Applications of Computational Algebraic Geometry (San Diego, CA, 1997), pp. 41–66, Proc. Sympos. Appl. Math., 53, Amer. Math. Soc., Providence, RI, 1998.

[McD02] McDonald, John, *"Fractional power series solutions for systems of equations,"* Discrete Comput. Geom. 27 (2002), no. 4, pp. 501–529.

[McL97] McLennan, Andrew, *"The maximal number of regular totally mixed Nash equilibria,"* J. Econom. Theory 72 (1997), no. 2, pp. 411–425.

[Mik03] Mikhalkin, Grigory, *"Amoebas of Algebraic Varieties,"* preprint, downloadable from `http://www.math.utah.edu/~gmikhalk`.

[Min41] Minding, Ferdinand, *"Über die Bestimmung des Grades einer durch Elimination hervorgehenden Gleichung,"* J. Reine. Angew. Math. 22, pp. 178–183 (1841); J. Math. Pures Appl., Ser. 1, 6, pp. 412–418 (1841).

[MS87] Morgan, Alexander and Sommese, Andrew, *"A homotopy for solving general polynomial systems that respects m-homogeneous structures,"* Appl. Math. Comput. 24 (1987), no. 2, pp. 101–113.

[MP98] Mourrain, Bernard and Pan, Victor, *"Asymptotic Acceleration of Solving Multivariate Polynomial Systems of Equations,"* Proc. STOC '98, pp. 488–496, ACM Press, 1998.

[Mum95] Mumford, David, *Algebraic Geometry I: Complex Projective Varieties,* reprint of the 1976 edition, Classics in Mathematics, Springer-Verlag, Berlin, 1995.

[NM99] Nešić, D. and Mareels, Ivan M. Y., *"Controllability of structured polynomial systems,"* IEEE Trans. Automat. Control 44 (1999), no. 4, pp. 761–764.

[Pap95] Papadimitriou, Christos H., *Computational Complexity,* Addison-Wesley, 1995.

[PS85] Preparata, Franco P. and Shamos, Michael Ian, *Computational Geometry: An Introduction,* Texts and Monographs in Computer Science, Springer-Verlag, New York-Berlin, 1985.

[Roj94] Rojas, J. Maurice, *"A Convex Geometric Approach to Counting the Roots of a Polynomial System,"* Theoretical Computer Science (1994), vol. 133 (1), pp. 105–140. (Additional notes and corrections available on-line at `http://www.math.tamu.edu/~rojas/list2.html` .)

[Roj97] ________________, *"A New Approach to Counting Nash Equilibria,"* Proceedings of the IEEE/IAFE Conference on Computational Intelligence for Financial Engineering, Manhattan, New York, March 23–25, 1997, pp. 130–136.

[Roj99a] ________________, *"Toric Intersection Theory for Affine Root Counting,"* Journal of Pure and Applied Algebra, vol. 136, no. 1, March, 1999, pp. 67–100.

[Roj99b] ________________, *"Solving Degenerate Sparse Polynomial Systems Faster,"* Journal of Symbolic Computation, vol. 28 (special issue on elimination theory), no. 1/2, July and August 1999, pp. 155–186.

[Roj00a] ________________, *"Algebraic Geometry Over Four Rings and the Frontier to Tractability,"* Contemporary Mathematics, vol. 270, Proceedings of a Conference on Hilbert's Tenth Problem and Related Subjects (University of Gent, November 1-5, 1999), edited by Jan Denef, Leonard Lipschitz, Thanases Pheidas, and Jan Van Geel, pp. 275–321, AMS Press (2000).

[Roj02] ————————, *"Additive Complexity and the Roots of Polynomials Over Number Fields and $\mathtt{p}$-adic Fields,"* Proceedings of the 5th Annual Algorithmic Number Theory Symposium (ANTS V), Lecture Notes in Computer Science #2369, pp. 506–515, Springer-Verlag (2002).

[Roj03] ————————, *"Arithmetic Multivariate Descartes' Rule,"* American Journal of Mathematics, to appear.

[RW96] Rojas, J. M., and Wang, Xiaoshen, *"Counting Affine Roots of Polynomial Systems Via Pointed Newton Polytopes,"* Journal of Complexity, vol. 12, June (1996), pp. 116–133.

[RY02] Rojas, J. M. and Ye, Yinyu, *"On Solving Fewnomials Over an Interval in Fewnomial Time,"* submitted for publication, also available as Math ArXiV preprint `math.NA/0106225`.

[Sha94] Shafarevich, Igor R., *Basic Algebraic Geometry I,* second edition, Springer-Verlag (1994).

[Shu93] Shub, Mike, *"Some Remarks on Bézout's Theorem and Complexity Theory,"* From Topology to Computation: Proceedings of the Smalefest (Berkeley, 1990), pp. 443–455, Springer-Verlag, 1993.

[Sil95] Silverman, Joseph H., *The Arithmetic of Elliptic Curves,* corrected reprint of the 1986 original, Graduate Texts in Mathematics 106, Springer-Verlag (1995).

[Smi61] Smith, H. J. S., *"On Systems of Integer Equations and Congruences,"* Philos. Trans. 151, pp. 293–326 (1861).

[Sot03] Sottile, Frank, *"Toric Ideals, Real Toric Varieties, and the Moment Map,"* this volume, pp. 225–240.

[Stu96] Sturmfels, Bernd, *Gröbner Bases and Convex Polytopes,* University Lecture Series, 8, American Mathematical Society, Providence, RI, 1996.

[Stu02] ————————, *Solving systems of polynomial equations,"* CBMS Regional Conference Series in Mathematics, 97. Published for the Conference Board of the Mathematical Sciences, Washington, DC, by the American Mathematical Society, Providence, RI, 2002.

[Sus98] Sussmann, Héctor J., *"Some Optimal Control Applications of Real-Analytic Stratifications and Desingularization,"* Singularities Symposium — Łojasiewicz 70 (Kraków, 1996; Warsaw, 1996), 211–232, Banach Center Publ., 44,

[vdK00] Van Der Kallen, Wilberd, *"Complexity of the Havas, Majewski, Matthews LLL Hermite normal form algorithm,"* J. Symbolic Comput. 30 (2000), no. 3, pp. 329–337.

[Ver00] Verschelde, Jan, *"Toric Newton method for polynomial homotopies,"* Symbolic computation in algebra, analysis, and geometry (Berkeley, CA, 1998), J. Symbolic Comput. 29 (2000), no. 4–5, pp. 777–793.

[VR02] Vidyasagar, M. and Rojas, J. Maurice, *"An Improved Bound on the VC-Dimension of Neural Networks with Polynomial Activation Functions,"* submitted for publication.

[Vir01] Viro, Oleg, *"Dequantization of Real Algebraic Geometry on a Logarithmic Paper,"* Proceedings of the 3rd European Congress of Mathematicians, Birkhuser, Progress in Math, 201, (2001), pp. 135–146.

[Zie95] Ziegler, Gunter M., *Lectures on Polytopes*, Graduate Texts in Mathematics, Springer Verlag, 1995.

Department of Mathematics, Texas A&M University, TAMU 3368, College Station, Texas 77843-3368, USA.

E-mail address: `rojas@math.tamu.edu` `http://www.math.tamu.edu/~rojas`