

The Maximal Number of Real Roots of a Multihomogeneous System of Polynomial Equations

by

Andrew McLennan

Department of Economics

University of Minnesota

271 19th Avenue South

Minneapolis, MN 55455

`mclennan@atlas.socsci.umn.edu`

TEXed: March 26, 1997

THE MAXIMAL NUMBER OF REAL ROOTS OF A
MULTIHOMOGENEOUS SYSTEM OF POLYNOMIAL EQUATIONS

Abstract

A multihomogeneous system of polynomial equations, with as many equations as degrees of freedom, has instances for which there are as many regular real roots, in the relevant product of projective spaces, as are allowed, for the corresponding dehomogenized system, by Bernshtein's [Ber] theorem. One may, in addition, require that all roots lie in a prescribed open subset of the solution space.

The general form of a sparse system of d polynomial equations in d variables is

$$f(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_d(\mathbf{x})) = 0, \quad (1)$$

where $\mathbf{x} = (x_1, \dots, x_d)$ and, for each $i = 1, \dots, d$, there is a nonempty finite $\mathcal{A}_i \subset \mathbf{N}^d$ such that $f_i(\mathbf{x}) = \sum_{a \in \mathcal{A}_i} c_{ia} \mathbf{x}^a$ for some coefficients c_{ia} . (Here $\mathbf{x}^a = x_1^{a_1} x_2^{a_2} \dots x_d^{a_d}$.) The result of this note concerns the specialization of this framework in which the variables in \mathbf{x} are divided into s groups, so that $\mathbf{x} = (\mathbf{y}^1, \dots, \mathbf{y}^s)$ where $\mathbf{y}^k = (y_1^k, \dots, y_{d^k}^k)$, and there are nonnegative integers m_i^k ($i = 1, \dots, d, k = 1, \dots, s$) such that

$$\mathcal{A}_i = \mathcal{A}_i^1 \times \dots \times \mathcal{A}_i^s, \quad \text{where} \quad \mathcal{A}_i^k = \{a \in \mathbf{N}^{d^k} : a_1 + \dots + a_{d^k} \leq m_i^k\}. \quad (2)$$

To see the significance of this we construct the associated ‘‘multihomogenized’’ system of equations. A new collection of variables is given by $\hat{\mathbf{x}} = (\hat{\mathbf{y}}^1, \dots, \hat{\mathbf{y}}^s)$ where $\hat{\mathbf{y}}^k = (y_0^k, y_1^k, \dots, y_{d^k}^k)$. Let

$$\hat{\mathcal{A}}_i = \hat{\mathcal{A}}_i^1 \times \dots \times \hat{\mathcal{A}}_i^s, \quad \text{where} \quad \hat{\mathcal{A}}_i^k = \{\hat{a} \in \mathbf{N}^{d^k+1} : \hat{a}_0 + \hat{a}_1 + \dots + \hat{a}_{d^k} = m_i^k\}.$$

There are bijections $\beta_i^k : \hat{\mathcal{A}}_i^k \rightarrow \mathcal{A}_i^k$ and $\beta_i : \hat{\mathcal{A}}_i \rightarrow \mathcal{A}_i$ given by $\beta_i^k(\hat{a}_0, \hat{a}_1, \dots, \hat{a}_{d^k}) = (\hat{a}_1, \dots, \hat{a}_{d^k})$ and $\beta_i(\hat{a}^1, \dots, \hat{a}^s) = (\beta_i^1(\hat{a}^1), \dots, \beta_i^s(\hat{a}^s))$. For each instance of (1) there is an associated multihomogeneous system of equations

$$\hat{f}(\hat{\mathbf{x}}) = (\hat{f}_1(\hat{\mathbf{x}}), \dots, \hat{f}_d(\hat{\mathbf{x}})) = 0 \in \mathbb{R}^d \quad \text{where} \quad \hat{f}_i(\hat{\mathbf{x}}) = \sum_{\hat{a} \in \hat{\mathcal{A}}_i} c_{i\beta_i(\hat{a})} \hat{\mathbf{x}}^{\hat{a}}. \quad (3)$$

Each \hat{f}_i is homogeneous of degree m_i^k as a function \mathbf{y}^k , for any fixed values of the other variables. The notion of a *multihomogeneous* polynomial, defined in this way, generalizes the concepts of homogeneous polynomial and multilinear function. The definition of a

totally mixed Nash equilibrium of a normal form game gives rise to a multihomogeneous system in which, for each $k = 1, \dots, s$, there are d^k indices i for which $m_i^k = 0$ and $m_i^{k'} = 1$ for all $k' \neq k$. These systems are studied by McKelvey and McLennan [MM]; the proof below is a straightforward generalization of their argument.

Let $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ and $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$. The products of projective spaces $\mathcal{P}^K = P^{d^1}(K) \times \dots \times P^{d^s}(K)$, for $K = \mathbb{R}^*, \mathbb{R}, \mathbb{C}^*, \mathbb{C}$, are natural spaces in which to look for, and count, solutions of (3). In the obvious way we regard these spaces as nested: $\mathcal{P}^{\mathbb{R}^*} \subset \mathcal{P}^{\mathbb{R}} \subset \mathcal{P}^{\mathbb{C}}$ and $\mathcal{P}^{\mathbb{R}^*} \subset \mathcal{P}^{\mathbb{C}^*} \subset \mathcal{P}^{\mathbb{C}}$. For

$$\hat{\mathbf{x}} = (\hat{\mathbf{y}}^1, \dots, \hat{\mathbf{y}}^s) \in (\mathbb{C}^{d^1+1} \setminus \{0\}) \times \dots \times (\mathbb{C}^{d^s+1} \setminus \{0\})$$

let $[\hat{\mathbf{x}}] = ([\hat{\mathbf{y}}^1], \dots, [\hat{\mathbf{y}}^s])$ denote the corresponding element of $\mathcal{P}^{\mathbb{C}}$. If $\mathbf{x} = (\mathbf{y}^1, \dots, \mathbf{y}^s)$ is a solution of (1), where $\mathbf{y}^k = (y_1^k, \dots, y_{d^k}^k)$, then an associated solution $\hat{\mathbf{x}}$ of (3) is derived by setting $\hat{\mathbf{y}}^k = (1, y_1^k, \dots, y_{d^k}^k)$ and $[\hat{\mathbf{x}}] = ([\hat{\mathbf{y}}^1], \dots, [\hat{\mathbf{y}}^s])$. Although the Theorem below is most naturally understood as pertaining to (3), for several reasons it will be more convenient to argue in terms of system (1).

We will be most interested in real coefficients c_{ia} , but for several points in the argument we consider the space of vectors of complex coefficients. Let $\mathcal{H}^{\mathbb{R}} (\mathcal{H}^{\mathbb{C}})$ be the set of f with real (complex) coefficients in which no f_i vanishes identically. Consider $(f, \mathbf{x}) \in \mathcal{H}^{\mathbb{R}} \times (\mathbb{R}^*)^d (\mathcal{H}^{\mathbb{C}} \times (\mathbb{C}^*)^d)$. Since \mathbf{x} has only nonzero components, no monomial \mathbf{x}^a vanishes. Since the coefficients of each f_i can be varied freely near f_i without leaving $\mathcal{H}^{\mathbb{R}} (\mathcal{H}^{\mathbb{C}})$, it follows that (f, \mathbf{x}) is a regular point of the map $(f, \mathbf{x}) \mapsto f(\mathbf{x})$. The regular value theorem implies that

$$V^{\mathbb{R}} := \{(f, \mathbf{x}) \in \mathcal{H}^{\mathbb{R}} \times (\mathbb{R}^*)^d : f(\mathbf{x}) = 0\} \text{ and } V^{\mathbb{C}} := \{(f, \mathbf{x}) \in \mathcal{H}^{\mathbb{C}} \times (\mathbb{C}^*)^d : f(\mathbf{x}) = 0\}$$

are smooth manifolds of the same dimensions as $\mathcal{H}^{\mathbb{R}}$ and $\mathcal{H}^{\mathbb{C}}$ respectively.

Let $\pi^{\mathbb{R}} : V^{\mathbb{R}} \rightarrow \mathcal{H}^{\mathbb{R}}$ and $\pi^{\mathbb{C}} : V^{\mathbb{C}} \rightarrow \mathcal{H}^{\mathbb{C}}$ be the restrictions of the natural projections. Evidently (f, \mathbf{x}) is a regular point of $\pi^{\mathbb{R}} (\pi^{\mathbb{C}})$ if and only if there are no vectors in $T_{(f, \mathbf{x})} V^{\mathbb{R}} (T_{(f, \mathbf{x})} V^{\mathbb{C}})$ that are parallel to $(\mathbb{R}^*)^d ((\mathbb{C}^*)^d)$, which is precisely the condition that \mathbf{x} is a regular point of f . Let $\mathcal{U}^{\mathbb{R}}$ and $\mathcal{U}^{\mathbb{C}}$ be the sets of f in $\mathcal{H}^{\mathbb{R}}$ and $\mathcal{H}^{\mathbb{C}}$ having only regular roots. Applications of Sard's theorem to $\pi^{\mathbb{R}}$ and $\pi^{\mathbb{C}}$ show that $\mathcal{U}^{\mathbb{R}}$ and $\mathcal{U}^{\mathbb{C}}$ are dense. Note that the discussion in this and the last paragraph is general, in that (2) plays no role.

For systems satisfying (2), a root $[\hat{\mathbf{x}}] \in \mathcal{P}^{\mathbb{C}}$ of \hat{f} is said to be *regular* if $\hat{\mathbf{x}}$ is a regular point of \hat{f} . It is easy to show that if $\hat{\mathbf{x}} = (\hat{\mathbf{y}}^1, \dots, \hat{\mathbf{y}}^s)$ is a regular point of \hat{f} , then so is $(\alpha^1 \hat{\mathbf{y}}^1, \dots, \alpha^s \hat{\mathbf{y}}^s)$ for any $\alpha^1, \dots, \alpha^s \in \mathbb{C}^*$, so this definition is meaningful.

The convex hull of \mathcal{A}_i is denoted by Q_i , and is called the *Newton polytope* of f_i . Generalizing a result of Kushnirenko [Ku], Bernshtein [Ber] shows that, for the general

system (1), every f in a dense subset of $\mathcal{U}^{\mathbb{C}}$ has $\mathcal{MV}(Q_1, \dots, Q_d)$ roots in $(\mathbb{C}^*)^d$, where $\mathcal{MV}(Q_1, \dots, Q_d)$ is the mixed volume of the d -tuple of Newton polytopes. (The concept of mixed volume is due to Minkowski; Betke [Bet] is a recent treatment.)

We show that, if (2) holds, there exists a real coefficient vector for which there are as many real roots as are allowed by Bernshtein's theorem. (The example $d = 1$, $\mathcal{A}_1 = \{0, 3\}$ illustrates how, in general, this may fail to be the case.) More precisely:

Theorem: *When (2) holds, for any open $\hat{U} \subset \mathcal{P}^{\mathbb{R}}$ there is an $f \in \mathcal{U}^{\mathbb{R}}$ for which the corresponding \hat{f} has $\mathcal{MV}(Q_1, \dots, Q_d)$ regular roots in \hat{U} .*

Remark 1: For the general version of system (1) the maximal number of real regular roots of f for $f \in \mathcal{U}^{\mathbb{R}}$ is denoted by $\rho(\mathcal{A}_1, \dots, \mathcal{A}_d)$. In view of the correspondence between roots of (1) and roots of (3), the Theorem implies that $\rho(\mathcal{A}_1, \dots, \mathcal{A}_d) = \mathcal{MV}(Q_1, \dots, Q_d)$ when (2) holds. Other work studying $\rho(\mathcal{A}_1, \dots, \mathcal{A}_d)$ includes the theory of fewnomials of Khovanski [Kh], in which arguments are based on the number $|\mathcal{A}_i|$ of monomials in each equation, and Sturmfels [S] and Itenberg and Roy [IR], who use the methods of the theory of sparse systems of polynomials to derive bounds on $\rho(\mathcal{A}_1, \dots, \mathcal{A}_d)$. These authors also consider the maximal number of roots in the various orthants of \mathbb{R}^d . For the specialized system given by (2) the Theorem shows that these maximums coincide with $\rho(\mathcal{A}_1, \dots, \mathcal{A}_d)$. Supports $\mathcal{A}_1, \dots, \mathcal{A}_d$ with these properties are interesting from the point of view of the conjectures presented by Itenberg and Roy, since the assertions of these conjectures can hold only in quite particular ways. We mention now that the part of the proof of Lemma 3 below that moves real roots into a prescribed orthant depends on a condition ($a \in \mathcal{A}_i$ whenever $a \leq b \in \mathcal{A}_i$) that is more general than (2).

Remark 2: The Theorem suggests that multihomogeneous systems are well behaved in the sense of sharing properties of the general homogeneous system considered in Bezout's theorem. Additional evidence for this point of view is given by Sturmfels and Zelevinsky [SZ], who give Sylvester type (that is, determinantal) formulas for the resultant of the unmixed ($\mathcal{A}_0 = \mathcal{A}_1 = \dots = \mathcal{A}_d$) multihomogeneous system in which the number of equations exceeds the number of degrees of freedom by one.

The Theorem follows from Lemmas 1-3 below.

Lemma 1: *If (2) holds, there is an $f \in \mathcal{U}^{\mathbb{R}}$ for which the corresponding \hat{f} has no roots in $\mathcal{P}^{\mathbb{C}} \setminus \mathcal{P}^{\mathbb{R}^*}$.*

Proof: For each i let

$$f_i(\mathbf{x}) = \prod_{k=1}^s \prod_{\ell=1}^{m_i^k} \mu_i^{k\ell}(\mathbf{y}^k)$$

for some choice of affine functionals $\mu_i^{k\ell} : \mathbb{R}^{d^k} \rightarrow \mathbb{R}$. It is easy to see that

$$\hat{f}_i(\hat{\mathbf{x}}) = \prod_{k=1}^s \prod_{\ell=1}^{m_i^k} \hat{\mu}_i^{k\ell}(\hat{\mathbf{y}}^k),$$

where, for each i, k, ℓ , $\hat{\mu}_i^{k\ell}$ is the linear functional on \mathbb{R}^{d^k+1} that reduces to $\mu_i^{k\ell}$ under the specialization $y_0^k = 1$.

For each k let

$$M^k = \{ \mu_i^{k\ell} : 1 \leq i \leq d, 1 \leq \ell \leq m_i^k \}.$$

We may choose the affine functionals $\mu_i^{k\ell}$ in general position in the sense that, for any subset $M \subset M^k$, the intersection of the kernels of the elements of M is an affine subspace of \mathbb{R}^{d^k} of dimension $d^k - \#M$. We may also insist that the intersection of the kernels has an intersection with each coordinate subspace of the expected dimension, so that when $\#M = d^k$, the unique common root has nonvanishing coordinates.

Let $[\hat{\mathbf{x}}] = ([\hat{\mathbf{y}}^1], \dots, [\hat{\mathbf{y}}^s])$ be a projective root of (3). For each i there must be some (k_i, ℓ_i) such that $\hat{\mu}_i^{k_i \ell_i}(\hat{\mathbf{y}}^{k_i}) = 0$. The general position assumption implies that for each k there are at most d^k indices i with $k_i = k$, and since the total number of indices i is $d = d^1 + \dots + d^s$, for each k there must be exactly d^k such i . Therefore each $\hat{\mathbf{y}}^k$ is a solution of a system of d^k linear equations in $d^k + 1$ variables that has exactly one projective root, which is real, namely $[(1, y_1^k, \dots, y_{d^k}^k)]$ where $\mathbf{y}^k = (y_1^k, \dots, y_{d^k}^k)$ is the unique root of the associated system of d^k affine functionals $\mu_i^{k\ell}$. This shows that all roots of (3) are in $\mathcal{P}^{\mathbb{R}^*}$.

Suppose $k_i = k$. By varying \mathbf{y}^k in the intersection of the kernels of the other $\mu_{i'}^{k_i' \ell_{i'}}$ with $k_{i'} = k$, we may vary f_i without changing the value of any other $f_{i'}$. Thus the image of $Df(\mathbf{x})$ spans the standard basis of \mathbb{R}^d , so \mathbf{x} is a regular root. This shows that all roots are regular, so that $f \in \mathcal{U}^{\mathbb{R}}$. ■

Remark 3: For the system of equations in the proof above, the roots are evidently in 1-1 correspondence with the set of d -tuples $((k_1, \ell_1), \dots, (k_d, \ell_d))$ with $1 \leq \ell_i \leq m_i^{k_i}$ for all i and $\#\{i : k_i = k\} = d^k$ for all k . An inductive formula for the number of roots can be developed from the observation that setting $\mu_i^{k\ell} = 0$ results in a multihomogeneous system with one fewer equation and one less degree of freedom, in that $P^{d^k}(\mathbb{C})$ is replaced by a plane of codimension one. (Cf. [MM, Sect. 5].)

Lemma 2: For $f \in \mathcal{U}^{\mathbb{R}}$, if the corresponding \hat{f} has no roots in $\mathcal{P}^{\mathbb{C}} \setminus \mathcal{P}^{\mathbb{R}^*}$, then

$$\rho(\mathcal{A}_1, \dots, \mathcal{A}_d) = \mathcal{M}\mathcal{V}(Q_1, \dots, Q_d).$$

Proof: Consider a root $[\hat{\mathbf{x}}]$ of \hat{f} . If $\hat{\mathbf{x}} = (\hat{\mathbf{y}}^1, \dots, \hat{\mathbf{y}}^s)$ where $\hat{\mathbf{y}}^k = (y_0^k, y_1^k, \dots, y_{d^k}^k)$, we set $\mathbf{x} = (\mathbf{y}^1, \dots, \mathbf{y}^s)$ where $\mathbf{y}^k = (y_1^k/y_0^k, \dots, y_{d^k}^k/y_0^k)$. Applying the (complex) implicit function theorem at \mathbf{x} , then multihomogenizing, we see that there exist neighborhoods $U_{[\hat{\mathbf{x}}]} \subset \mathcal{H}^{\mathbf{C}}$ of f and $W_{[\hat{\mathbf{x}}]} \subset \mathcal{P}^{\mathbf{C}}$ of $[\hat{\mathbf{x}}]$ for which

$$\{ (f', [\hat{\mathbf{x}}']) \in U_{[\hat{\mathbf{x}}]} \times W_{[\hat{\mathbf{x}}]} : [\hat{\mathbf{x}}'] \text{ is a root of } \hat{f}' \}$$

is the graph of a C^∞ function $r_{[\hat{\mathbf{x}}]} : U_{[\hat{\mathbf{x}}]} \rightarrow W_{[\hat{\mathbf{x}}]}$. At this point we have not used the hypothesis that \hat{f} has no roots outside of $\mathcal{P}^{\mathbf{R}^*}$, and by applying the argument to this point to some \hat{f} with the maximal number of real regular roots, we may conclude that

$$\rho(\mathcal{A}_1, \dots, \mathcal{A}_d) \leq \mathcal{M}\mathcal{V}(Q_1, \dots, Q_d).$$

If every neighborhood of f in $\mathcal{H}^{\mathbf{C}}$ contained an f' for which the associated \hat{f}' had a root in $\mathcal{P}^{\mathbf{C}} \setminus \bigcup_{[\hat{\mathbf{x}}]} W_{[\hat{\mathbf{x}}]}$, then, since this set is compact, it would contain a root of \hat{f} , which is impossible. Consequently it must be possible to choose a neighborhood U of f small enough that, for every $f' \in U$, the only roots of the associated \hat{f}' in $\mathcal{P}^{\mathbf{C}}$ are the various $r_{[\hat{\mathbf{x}}]}(f')$. Since U is open in $\mathcal{H}^{\mathbf{C}}$, generic $f' \in U$ must have $\mathcal{M}\mathcal{V}(Q_1, \dots, Q_d)$ roots, so we see that this number is also a lower bound on $\rho(\mathcal{A}_1, \dots, \mathcal{A}_d)$. ■

We will say that a set $\mathcal{A} \subset \mathbf{N}^d$ is *comprehensive* if $a \in \mathcal{A}$ whenever $a, b \in \mathbf{N}^d$ with $a \leq b \in \mathcal{A}$.

Lemma 3: Suppose that (2) holds, and $f \in \mathcal{U}^{\mathbf{R}}$. If the corresponding \hat{f} has no roots in $\mathcal{P}^{\mathbf{C}} \setminus \mathcal{P}^{\mathbf{R}^*}$, then for any open $\hat{U} \subset \mathcal{P}^{\mathbf{R}}$ there is an $f' \in \mathcal{U}^{\mathbf{R}}$ for which the corresponding \hat{f}' has as many roots as f , with all of them lying in \hat{U} .

Proof: Let $\underline{\mathbf{x}} = (-C, \dots, -C) \in \mathbb{R}^d$ for some $C > 0$. Each \mathcal{A}_i is comprehensive, implying that in the system $f'' : \mathbf{x} \mapsto f(\mathbf{x} + \underline{\mathbf{x}})$ each f_i'' has no monomials with nonzero coefficients that are not already in \mathcal{A}_i . The roots of the system f'' are precisely the points of the form $\mathbf{x} - \underline{\mathbf{x}}$ where \mathbf{x} is a root of f , and by taking C large, we can insure that they have only positive components. This being the case, if the corresponding \hat{f}'' had a root $[\hat{\mathbf{x}}] = ([\hat{\mathbf{y}}^1], \dots, [\hat{\mathbf{y}}^s])$ in $\mathcal{P}^{\mathbf{C}} \setminus \mathcal{P}^{\mathbf{R}^*}$, there would necessarily be some k for which $[\hat{\mathbf{y}}^k]$ was in the hyperplane of $P^{d^k}(\mathbf{C})$ given by $y_0^k = 0$, and there would be a corresponding root of \hat{f} in $\mathcal{P}^{\mathbf{C}} \setminus \mathcal{P}^{\mathbf{R}^*}$, contrary to hypothesis.

By choosing C large, we can force all the projective roots of \hat{f}'' into an arbitrarily small neighborhood of

$$([(0, 1, \dots, 1)], \dots, [(0, 1, \dots, 1)]) \in \mathcal{P}^{\mathbf{R}}.$$

The action of the product group $G = O(d^1 + 1) \times \dots \times O(d^s + 1)$ on $\mathbb{R}^{d^1+1} \times \dots \times \mathbb{R}^{d^s+1}$ induces an action on the space of multilinear systems (3) with real coefficients. Under such

a transformation, the roots of the transformed system are related to the roots of the given system by the corresponding action of G on $\mathcal{P}^R = P^{d^1}(\mathbb{R}) \times \dots \times P^{d^s}(\mathbb{R})$. By choosing an appropriate element of G we can move $([(0, 1, \dots, 1)], \dots, [(0, 1, \dots, 1)])$ to any other point in \mathcal{P}^R , and under the induced action \hat{f}'' is transformed to a system \hat{f}' having all its roots in the corresponding neighborhood of the chosen point. With C sufficiently large, and the point in \mathcal{P}^R chosen suitably, this results in \hat{f}' having all its roots in \hat{U} . Finally, it is easy to check that the transformations considered in this argument preserve the regularity of all roots, so that $f' \in \mathcal{H}^R$. ■

References

- [Ber] D. N. Bernshtein, The number of roots of a system of equations, *Functional Analysis and its Applications* **9** (1975), 183–185.
- [Bet] U. Betke, Mixed volumes of polytopes, *Arch. Math.* **58** (1992), 388–391.
- [IR] I. Itenberg and M.-F. Roy, Multivariate Descartes' rule, *Beitrage zur Algebra und Geometrie* **37** (1996) 337–346.
- [Kh] A. G. Khovanski, *Fewnomials*, American Mathematical Society, Translations of Math. Monographs, Vol. 88, 1991.
- [Ku] A. G. Kushnirenko, The Newton polyhedron and the number of solution of a system of k equations in k unknowns, *Upsekhi Mat. Nauk.* **30** (1975), 266–267.
- [MM] R. D. McKelvey and A. McLennan, The maximal number of regular totally mixed Nash equilibria, *Journal of Economic Theory* **72** (1997), 411–425.
- [SZ] B. Sturmfels and A. Zelevinsky, Multigraded resultants of Sylvester type, *Journal of Algebra* **163** (1994), 115–127.
- [S] B. Sturmfels, On the number of real roots of a sparse polynomial system, *Hamiltonian and gradient flows, algorithms, and control*, A. Block, ed., Am. Math. Soc., Providence, 1994, 137–143.