The Expected Number of 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

June 1998

---

* Software used to perform calculations reported herein used source code developed by the Gambit Project. I have benefited from numerous discussions with Maurice Rojas.

Running head: **The Expected Number of Real Roots**

Correspondent:
Andrew McLennan
Department of Economics
University of Minnesota
271 19th Avenue South
Minneapolis, MN 55455

Abstract

The methods of Shub and Smale [SS93] are extended to the class of multihomogeneous systems of polynomial equations. Theorem 1 is a formula for the mean (with respect to a particular distribution on the space of coefficient vectors) number of real roots that has as corollaries: (a) Shub and Smale's result that the expected number of real roots of the general homogeneous system is the square root of the generic number of complex roots given by Bezout's theorem; (b) Rojas' [Roj96] characterization of the mean number of real roots of an "unmixed" multihomogeneous system. Theorem 2 asserts that, for unmixed systems, the square of the mean number of roots always exceeds the generic number of complex roots, as determined by Bernstein's [Ber75] theorem, and there is extensive computational evidence in support of the conjecture that this inequality also holds for mixed multihomogeneous systems.

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

1. Introduction

Shub and Smale [SS93] study the mean number of real roots (in real projective space) of the general homogeneous system of n polynomial equations of fixed degrees $d_1, \ldots, d_n$, with respect to a particular distribution on the space of real coefficient vectors, arriving at the lovely result that the mean number of real roots is $\sqrt{\prod_i d_i}$, the square root of the generic number of complex projective roots, as given by Bezout's theorem. Their approach is descended from [Kac43] which spawned an extensive literature that is ably surveyed, and extended, by Edelman and Kostlan [EC95].

This note generalizes the methods of [SS93] to the class of systems of multihomogeneous polynomial equations: there are n equations in $n + k$ variables that are broken into k groups such that each equation is homogeneous separately in the variables of each group. Fixing n, the decomposition into groups, and the degree of each equation in each group, we study the average number of real roots (in the natural k-fold product of projective spaces) of this system for random vectors of coefficients. The distribution on the space of coefficient vectors is equivalent to the one proposed by Rojas [Roj96]. Below we will show that for multihomogeneous systems this distribution is canonical in the sense of uniquely satisfying certain natural conditions.

Our main result, Theorem 1, gives a formula that expresses the average number of real roots as a product of the mean absolute value of the determinant of a random matrix and a ratio of products of Euler's function Γ, evaluated at multiples of $1/2$. The Shub-Smale theorem is a corollary, as is the formula given [Roj96] for unmixed systems. (Section 2 gives definitions for the terms 'unmixed' and 'mixed,' as they are used in the theory of sparse polynomial systems.)

The celebrated theorem of Bernshtein [Ber75] (see also [Kus75]) is the generalization of Bezout's theorem to sparse systems of polynomial equations, and it is natural to wonder whether the theorem of [SS93] extends, in some sense, to a relationship, for sparse systems, between the mean number of real roots and the generic (in the space of complex coefficient vectors) number of complex roots, as given by Bernshtein's theorem. Indeed, Rojas' *Square Root Volume Conjecture* [Roj96] would be such an extension, since it proposes that within certain similarity classes of problems, the mean number of roots should be proportional to the square root of the generic number of complex roots. Here we propose:

Conjecture 1: The mean (with respect to the probability measure on coefficient vectors described herein) number of real roots of a multihomogeneous system is at least as large as the square root of the generic number of complex roots.

Theorem 2 uses Rojas' formula to establish this claim for unmixed multihomogeneous systems.

The formula of Theorem 1 is easy to evaluate on the computer. (The main obstacle to accuracy is an integral that we estimate by Monte Carlo methods that have standard error proportional to the inverse of the square of the sample size.) In addition, for multilinear systems it is possible to compute the maximal number of roots recursively, as we describe in Section 5. The author has implemented these algorithms, and compared the numbers they yield for thousands of multilinear systems, without finding any counterexamples to Conjecture 1.

The author's own interest in this topic is motivated by concepts of noncooperative game theory. [(1)]

[(1)] This is not the place to give a general introduction to noncooperative game theory; Fudenberg and Tirole (1991) is a standard text. For the internal logic of this paper the description of *quasiequilibrium* (Section 2) is sufficient. For the connection between this notion and the standard concepts of *Nash equilibrium* and *totally mixed Nash equilibrium* see [MM97, McL97].

McLennan and McKelvey [MM97] give a method for constructing games that give rise to systems of polynomials having as many regular (real) quasiequilibria as are permitted by Bernshtein's theorem. Roughly, the message of this result is that the maximal number of Nash equilibria is large, at least compared to most game theorists' prior intuition.

Games that have the maximal number of equilibria are thought to be very atypical, and there arises the question of whether the set of equilibria is not only potentially large, but also large on average. (Part of the problem is to find an appealing definition of "on average.") McLennan [McL97] investigates the application, to this problem, of the methods of [SS93]. From the point of view of determining the mean number of quasiequilibria, the results there are less general versions of Theorem 1 of this note, but additional problems arise in characterizing the relations between these results and the mean numbers of totally mixed equilibria and Nash equilibria of all sorts. From the point of view of the material here, the most important point is that Conjecture 1 above would imply that, in a variety of senses, the mean number of Nash equilibria grows exponentially with various measures of the size of the game.

In connection with speculation concerning whether Conjecture 1 might hold for more general classes of sparse systems than the multihomogeneous ones, it is interesting to note that multihomogeneous systems are potentially special insofar as they can have as many real regular roots as are permitted by Bernshtein's theorem. McLennan [McL98] proves this by pointing out that the argument in [MM97], which establishes this claim for the systems arising in game theory, is actually valid for any multihomogeneous system. In particular, this is true of the general homogeneous system of Bezout's theorem.

The remainder has the following organization. Section 2 describes multihomogeneous systems as a certain type of sparse system. In Section 3 we introduce a certain group action, and use it to motivate the definition of an inner product on the space of coefficient vectors of a multihomogeneous system. This allows us to state the problem precisely, insofar as the notion of average is defined with respect to the uniform distribution on a cartesian product of unit spheres, relative to this inner product, in the space of coefficient vectors. Section 4 states Theorem 1 and presents a number of consequences. Section 5 defines mixed volume, states Bernshtein's theorem precisely, shows how the generic number of complex roots of a multihomogeneous system may be computed recursively, and applies this recursion to Rojas' formula to prove Theorem 2. Sections 6–8 present the proof of Theorem 1.

2. Multihomogeneous Systems

We adopt some of the standard notation of the theory of sparse systems of polynomial equations. The general form of a *sparse system* of n polynomial equations in ℓ variables is

$$f(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_n(\mathbf{x})) = 0, \tag{1}$$

where $\mathbf{x} = (x_1, \ldots, x_\ell)$ and, for each $i = 1, \ldots, n$, there is a nonempty finite $\mathcal{A}_i \subset \mathbf{N}^\ell$ such that $f_i(\mathbf{x}) = \sum_{a \in \mathcal{A}_i} f_{ia}\mathbf{x}^a$ for some system of coefficients f_{ia}. (Here $\mathbf{x}^a$ denotes the monomial $x_1^{a_1} x_2^{a_2} \cdots x_\ell^{a_\ell}$.) The general approach of the theory of sparse systems is to hold the n-tuple of *supports* $(\mathcal{A}_1, \ldots, \mathcal{A}_n)$ fixed while treating the coefficients f_{ia} as variables, for instance in the sense of studying properties that are generic in the space of vectors of coefficients. Such a system is said to be *unmixed* if $\mathcal{A}_1 = \ldots = \mathcal{A}_n$; otherwise it is *mixed*. Identifying a polynomial with its vector of coefficients, we regard $\mathcal{H}_i := I\!R^{\mathcal{A}_i}$ as the vector space of polynomials with real coefficients whose supports are contained in $\mathcal{A}_i$. Let $\mathcal{H} := \mathcal{H}_1 \times \ldots \times \mathcal{H}_n$.

The framework studied in this paper is a specialization of this general framework in which the variables in $\mathbf{x}$ are divided into k groups, so that $\mathbf{x} = (\mathbf{y}_1, \ldots, \mathbf{y}_k)$ where $\mathbf{y}_j = (y_{j0}, y_{j1}, \ldots, y_{jn_j})$, and each equation is homogeneous of degree δ_{ij} as a function of $\mathbf{y}_j$, for any given values of the other variables

$(\mathbf{y}_1, \ldots, \mathbf{y}_{j-1}, \mathbf{y}_{j+1}, \ldots, \mathbf{y}_k)$. More precisely, we require that there are nonnegative integers δ_{ij} $(i = 1, \ldots, n$, $j = 1, \ldots, k)$ such that

$$\mathcal{A}_i = \mathcal{A}_{i1} \times \ldots \times \mathcal{A}_{ik}, \quad \text{where} \quad \mathcal{A}_{ij} = \{\, \alpha \in \mathbf{N}^{n_j+1} : \alpha_0 + \alpha_1 + \ldots + \alpha_{n_j} = \delta_{ij} \,\}.$$

Henceforth we will always assume that $\ell = n + k$, which means that there are effectively the same number of equations and unknowns:

$$n_1 + \ldots + n_k = n.$$

Insofar as n is now determined by the vector $\mathbf{n} := (n_1, \ldots, n_k)$, an instance of the type of system studied here is specified by the vector $\mathbf{n}$ and the $n \times k$ matrix $\delta := (\delta_{ij})$.

Three particular types of multihomogeneous system figure in our discussion:

(a) When $k = 1$ we have the general homogeneous system, for which the problem studied here was analyzed in [SS93].

(b) The *unmixed* multihomogeneous system studied in [Roj96] is described by the condition that all equations have the same support:

$$\delta_{1j} = \ldots = \delta_{nj} \quad (j = 1, \ldots, k).$$

(c) The systems arising, in game theory, from the concept of quasiequilibrium of a finite normal form game, are characterized by:

$$\delta_{ij} = \begin{cases} 0 & \text{if } q(i) = j, \\ 1 & \text{otherwise,} \end{cases} \tag{2}$$

where $q : \{1, \ldots, n\} \to \{1, \ldots, k\}$ is the function defined implicitly by the inequality

$$n_1 + \ldots + n_{q(i)-1} < i \leq n_1 + \ldots + n_{q(i)}.$$

3. An Invariant Inner Product

Since $n_1 + \ldots + n_k = n$, we may index the components of an exponent vector $a \in \mathbf{N}^{n+k}$ by the pairs (j, h) for $j = 1, \ldots, k$ and $h = 0, \ldots, n_j$. For such an a let

$$\eta(a) := \frac{a_{10}! \cdot \ldots \cdot a_{1n_1}!}{(a_{10} + \ldots + a_{1n_1})!} \cdot \ldots \cdot \frac{a_{k0}! \cdot \ldots \cdot a_{kn_k}!}{(a_{k0} + \ldots + a_{kn_k})!} = \binom{a_{10} + \ldots + a_{1n_1}}{a_{10}, \ldots, a_{1n_1}}^{-1} \cdot \ldots \cdot \binom{a_{k0} + \ldots + a_{kn_k}}{a_{k0}, \ldots, a_{kn_k}}^{-1}.$$

We endow each $\mathcal{H}_i$ with the inner product

$$\langle f_i, f_i' \rangle_i := \sum_{a \in \mathcal{A}_i} \eta(a) f_{ia} f_{ia}'. \tag{3}$$

Let $\|\cdot\|$ be the norm derived from $\langle \cdot, \cdot \rangle_i$, let M_i be the unit sphere in $\mathcal{H}_i$, and let $M := M_1 \times \ldots \times M_n$.

As a submanifold of $\mathcal{H}$, M inherits a measure corresponding to the notion of volume, and the *uniform distribution* on M is the probability distribution derived by normalizing so that the measure of M is unity. When we speak of the mean or average number of real roots, it will always be with respect to this underlying probability distribution. An equivalent (in the sense of having the same implied distribution of roots) formulation emphasized by [EK95] and [Roj96] takes the coefficients f_{ia} to be independent Gaussian random variables with mean 0 and variance $\eta(a)^{-1}$. [Roj96] presents a definition of these variances that is geometric and general, in the sense of pertaining to any sparse system. Here, following [SS93], we will motivate the inner product by appealing to invariance.

Consider the product group

$$G := O(n_1 + 1) \times \ldots \times O(n_k + 1).$$

There is the obvious component-wise action of G on $\mathbb{R}^{n_1+1} \times \ldots \times \mathbb{R}^{n_k+1}$, and for $f_i \in \mathcal{A}_i$ and $O \in G$, $f_i \circ O^{-1}$ is easily seen to be a polynomial function that is multihomogeneous for the same numbers δ_{ij}, so $f_i \circ O^{-1}$ is an element of $\mathcal{H}_i$. Thus the formula $Of_i := f_i \circ O^{-1}$ defines an action from the left of G on $\mathcal{A}_i$.

Lemma 3.1: The inner product (4) is the unique (up to multiplication by a scalar) inner product on $\mathcal{H}_i$ that is invariant under the action of G and with respect to which the monomials are pairwise orthogonal.

Proof: We only sketch the method, leaving calculations to motivated readers. The claim may be proved in the case $k = 1$, $n_1 = 1$, by observing that, for $f_i = \sum_{h=0}^{\delta_{i1}} a_h x^h y^{\delta_{i1}-h}$ and $g_i = \sum_{h=0}^{\delta_{i1}} b_h x^h y^{\delta_{i1}-h}$, we must have $\langle f_i, g_i \rangle = \sum_h \gamma_h a_h b_h$ for some numbers $\gamma_0, \ldots, \gamma_{\delta_{i1}} > 0$. The necessary information is obtained by equating coefficients in the identity

$$\gamma_{\delta_{i1}} = \|x^{\delta_{i1}}\|_i^2 = \|(\alpha x + (1-\alpha^2)^{1/2} y)^{\delta_{i1}}\|_i^2 = \sum_{h=0}^{\delta_{i1}} \gamma_h \binom{\delta_{i1}}{h}^2 \alpha^{2h}(1-\alpha^2)^{\delta_{i1}-h}.$$

The generalization to higher values of k and n_j may be obtained by applying this calculation repeatedly in an inductive manner. For example, we may equate coefficients in the identity

$$\|x^h y^{m-h}\|_i^2 = \|x^h(\alpha y + (1-\alpha^2)^{1/2} z)^{m-h}\|_i^2. \blacksquare$$

Combining the actions of G on the various $\mathcal{H}_i$, we obtain an action of G on $\mathcal{H}$ given by

$$Of := (f_1 \circ O^{-1}, \ldots, f_n \circ O^{-1}).$$

Each M_i is invariant under the action of G on $\mathcal{H}_i$, of course, so M is an invariant of the action of G on $\mathcal{H}$, and the restriction of this action to M is an action of G on M.

In general there will be other inner products on the coefficient vectors that are invariant under the action of G, but in which the monomials are not pairwise orthogonal.(2) It is interesting to point out that, in the systems of equations characterizing totally mixed equilibrium of normal form games, the requirement of invariance automatically entails orthogonality of the monomials. For example, consider i such that $q(i) = 1$. Then the monomials $y_{20}y_{30}\ldots y_{k0}$ and $y_{21}y_{30}\ldots y_{k0}$ are in the support of f_i. Let γ_1 be the element of $O(n_1+1)$ that maps $(1,0,\ldots,0)$ to $(0,1,\ldots,0)$ and $(0,1,\ldots,0)$ to $(-1,0,\ldots,0)$, while leaving all other standard unit basis vectors fixed, and let $\gamma = (\gamma_1, \ldots, \gamma_k) \in G$ where $\gamma_2, \ldots, \gamma_k$ are the respective identity transformations. Then invariance implies that

$$\langle y_{20}y_{30}\ldots y_{k0}, y_{21}y_{30}\ldots y_{k0}\rangle = \langle \gamma(y_{20}y_{30}\ldots y_{k0}), \gamma(y_{21}y_{30}\ldots y_{k0})\rangle = \langle y_{21}y_{30}\ldots y_{k0}, -y_{20}y_{30}\ldots y_{k0}\rangle = 0.$$

---

(2) A linear subspace of $\mathcal{H}_i$ is *invariant* under the action of G if it is mapped into itself by each element of G. The action of G on $\mathcal{H}_i$ is *irreducible* if it is not possible to decompose $\mathcal{H}_i$ as a direct sum of two invariant subspaces of positive dimension. Given such a decomposition, any pair of invariant inner products for the summands can be combined to create an invariant inner product for $\mathcal{H}_i$, so if the invariant inner product is unique up to multiplication by a scalar, then the action of G must be irreducible. The converse is a basic result of the theory of group representations. To see that $\mathcal{H}_i$ may be reducible, suppose that $\delta_{i1}, \ldots, \delta_{ik}$ are all even, and consider that the one dimensional subspace spanned by $\prod_j (y_{j0}^2 + y_{j1}^2 + \ldots + y_{jn_j}^2)^{\delta_{ij}/2} \in \mathcal{H}_i$ is invariant. I am indebted to Jonathan Robbins for a helpful consultation on this point.

4. Statement and Consequences of the Main Result

For $j = 1, \ldots, k$ let $N_j := \mathbf{P}^{n_j}(\mathbb{R})$ be n_j-dimensional projective space. In concrete calculations we regard N_j as the space of pairs $\zeta_j = [\tilde{\zeta}_j, -\tilde{\zeta}_j]$ of antipodal points in $\tilde{N}_j$, where $\tilde{N}_j$ is the unit sphere in $\mathbb{R}^{n_j+1}$. Let $N := N_1 \times \ldots \times N_k$ and $\tilde{N} := \tilde{N}_1 \times \ldots \times \tilde{N}_k$. In the usual way, the equation $f_i(\zeta) = 0$ is meaningful for $f_i \in M_i$ and $\zeta \in N$ even though f_i is not a function defined on N. The *incidence variety* is

$$V := \{ (f, \zeta) \in M \times N : f(\zeta) = 0 \}.$$

In general, the mean of an integrable real valued function g on a compact Riemannian manifold P, with respect to the uniform distribution, will be denoted by

$$\mathbf{E}_P(g) := \frac{1}{\mathrm{vol}(P)} \int_P g \, dP.$$

(Here we are letting P also denote the measure on P corresponding to the notion of volume induced by the Riemannian metric.) Let $\pi_1 : V \to M$ and $\pi_2 : V \to N$ be the restrictions of the projections $M \times N \to M$ and $M \times N \to N$. Our goal is the computation of

$$E(\mathbf{n}, \delta) := \mathbf{E}_M\big(\#(\pi_1^{-1}(\cdot))\big). \tag{4}$$

Let $\Gamma(s) := \int_0^\infty \exp(-z) z^{s-1} \, dz$ be Euler's function[3].

We are now able to state the central result.

Theorem 1: For $i = 1, \ldots, n$, let X_i be $\mathbb{R}^n$ with components indexed by the pairs (j, h) $(j = 1, \ldots, k, h = 1, \ldots, n_j)$. Let $\Xi_i : X_i \to \mathbb{R}^n$ be the linear transformation mapping x_i to the vector with entries $\sqrt{\delta_{ij}} \cdot x_i^{jh}$. Let $X := X_1 \times \ldots \times X_n$, and for $x \in X$ let $\Xi(x)$ be the $n \times n$ matrix whose i^{th} row is $\Xi_i(x_i)$. Let $L := L_1 \times \ldots \times L_n$, where each L_i is the unit sphere in X_i.

(a)

matrix is important.

$$E(\mathbf{n}, \delta) = \left(\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \right)^n \cdot \left(\prod_{j=1}^{k} \frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{n_j+1}{2})} \right) \cdot \mathbf{E}_L\big(|\det \Xi(\cdot)|\big). \tag{5}$$

(b) The induced distribution of roots is uniform: for any open $W \subset N$,

$$\mathbf{E}_M\big(\#(\pi_1^{-1}(\cdot) \cap \pi_2^{-1}(W))\big) = \frac{\mathrm{vol}(W)}{\mathrm{vol}(N)} E(\mathbf{n}, \delta).$$

The proof occupies Sections 6–9.

The remainder of this section develops some of the consequences of (5). Insofar as the function Γ is familiar, this is primarily a matter of analyzing the term $\mathbf{E}_L\big(|\det \Xi(\cdot)|\big)$. Although means of random determinants have been studied extensively [Gir90] there seems to be little prior work on mean absolute values of random determinants.

We begin by considering systems in which there is a subset of the variables that are determined by equations involving only those variables. Specifically, suppose there is some integer k' between 1 and k such that $\delta_{ij} = 0$ whenever $q(i) \le k' < j$ where q is the function defined at the end of Section 2. Set $n' := n_1 + \ldots + n_{k'}$. Then

$$\delta = \begin{bmatrix} \delta^{11} & 0 \\ \delta^{21} & \delta^{22} \end{bmatrix}.$$

(3) In the statement of the theorem, and subsequently, we write $\Gamma(\frac{1}{2})$ when it might seem simpler and more informative to substitute $\sqrt{\pi}$, since the author has found it helpful to adhere to a style in which the sum of the arguments to Γ in the numerator is equal to the sum in the denominator.

where δ^{11}, δ^{21}, and δ^{22} have dimensions $n' \times k'$, $(n-n') \times k'$, and $(n-n') \times (k-k')$ respectively. For any $\ell \in L$, $\Xi(\ell)$ has an $n' \times (n-n')$ block of zeros in its upper right corner, so its determinant is the product of the determinants of the $n' \times n'$ submatrix in the upper left and the $(n-n') \times (n-n')$ submatrix in the lower right. In particular, the determinant of $\Xi(\ell)$ does not depend on δ^{21}. Consequently (5) implies that $E(\mathbf{n}, \delta)$ is also independent of δ^{21}. When we set $\delta^{21} = 0$ we have a cartesian product of two independent systems. The distribution of coefficients for the combined system that we are studying is easily seen to be the product measure of the assumed distributions for the subsystems. For any particular coefficient vector for the combined system, the number of roots is the product of the numbers of roots of the subsystems, so the following is a consequence of the fact that the mean of a product of independent random variables is the product of their means.

Corollary 1: Suppose there is some $1 \leq k' < k$ such that $\delta_{ij} = 0$ whenever $q(i) \leq k' < j$, and let δ^{11} and δ^{22} be as above. Then

$$E(\mathbf{n}\delta) = E((n_1, \ldots, n_{k'}), \delta^{11}) \cdot E((n_{k'+1}, \ldots, n_k), \delta^{22}).$$

A second general principle results from the effect on the determinant of multiplying a row or a column by a scalar.

Corollary 2: If there are nonnegative integers $d_1, \ldots, d_n$ and $e_1, \ldots, e_k$ such that $\delta'_{ij} = d_i \cdot e_j \cdot \delta_{ij}$, then

$$E(\mathbf{n}, \delta') = \sqrt{\prod_{i=1}^{n} d_i} \cdot \sqrt{\prod_{j=1}^{k} e_j^{n_j}} \cdot E(\mathbf{n}, \delta).$$

Consider now the particular case of $k = 1$ and $\delta_{11} = \ldots = \delta_{n1} = 1$. This is a system of n linear functionals in $n+1$ variables, and there is be exactly one projective roots for almost all coefficient vectors. In view of (5) we must have:

Proposition 4.1: The mean absolute value of the determinant of a random $n \times n$ matrix whose rows are i.i.d. uniformly distributed points in S^{n-1} is $\Gamma(\frac{n}{2})^n / \Gamma(\frac{1}{2}) \Gamma(\frac{n+1}{2})^{n-1}$.

The results of [SS93] and Rojas follow directly from these last two results.

Corollary 3: ([SS93]) If $k = 1$, then $E(\mathbf{n}, \delta) = \sqrt{\prod_{i=1}^{n} \delta_{i1}}$.

Corollary 4: (Rojas) If there are nonnegative integers $e_1, \ldots, e_k$ such that $\delta_{ij} = e_i$, then

$$E(\mathbf{n}, \delta) = \sqrt{\prod_{j=1}^{k} e_j^{n_j}} \cdot \frac{\Gamma(\frac{1}{2})^{k-1} \Gamma(\frac{n+1}{2})}{\prod_{j=1}^{k} \Gamma(\frac{n_j+1}{2})}.$$

There is a class of systems for which $E(\mathbf{n}, \delta)$ can be computed exactly by combining Corollaries 1 and 2 with Proposition 4.1. I know of no case outside this class in which the integral $\int_L |\det \Xi(\ell)| \, d\ell$ evaluates to a closed form expression. For the systems arising from normal form games we are able to evaluate in closed form only when $k = 2$, which corresponds to a game with two players. The computation can be executed using the results above, but we leave details to the reader, in part because the result can also be derived that the observation that the system consists of two systems of linear equations.

Corollary 5: In the case of the game equilibrium system given by (2), if $k = 2$ then

$$E(\mathbf{n}, \delta) = \begin{cases} 1 & \text{if } n_1 = n_2, \\ 0 & \text{otherwise.} \end{cases}$$

5. Comparison with the Generic Number of Complex Roots

We now state Bernshtein's theorem, and analyze its consequences for multihomogeneous systems. Let $f(z) = (f_1(z), \dots, f_n(z))$ be a general sparse system of n equations in the n variables $z_1, \dots, z_n$, where f_i has support $\mathcal{A}_i \subset \mathbf{N}^n$. The *Newton polytope* of f_i is the convex polytope $Q_i = \text{con}(\mathcal{A}_i)$. The *mixed volume* of $Q_1, \dots, Q_n$, which was first defined and studied by Minkowski, and which we denote by $\mathcal{MV}(Q_1, \dots, Q_n)$, may be defined to be the coefficient of the monomial $\lambda_1 \cdot \ldots \cdot \lambda_n$ in the polynomial $\text{vol}(Q_\lambda)$ where

$$Q_\lambda = \lambda_1 Q_1 + \ldots + \lambda_n Q_n.$$

Theorem: [Ber75] Let $\mathbf{C}^* := \mathbf{C} \setminus \{0\}$. Let $\mathcal{H}^{\mathbf{C}} = \mathcal{H}_1^{\mathbf{C}} \times \ldots \times \mathcal{H}_n^{\mathbf{C}}$ where $\mathcal{H}_i^{\mathbf{C}} = \mathbf{C}^{\mathcal{A}_i}$ is the space of complex polynomials with support $\mathcal{A}_i$. For systems f in the complement, in $\mathcal{H}^{\mathbf{C}}$, of an algebraic set of positive (complex) codimension, there are $\mathcal{MV}(Q_1, \dots, Q_n)$ roots in $(\mathbf{C}^*)^n$.

We will apply this result to the "demultihomogenized" system obtained, from the given multihomogeneous system, by setting $y_{10} = \ldots = y_{k0} = 1$. In comparing the roots of the latter system, in $(\mathbf{C}^*)^n$, with the roots, in N, of the given multihomogeneous system, there is the possibility of roots in one of the coordinate subspaces (in the projective sense) along which one of the variables vanishes, but invariance under the action of G quickly implies that generic systems do not have such roots. Similarly, for generic systems there are no roots at projective infinity. Thus, generically, there is a one-to-one correspondence between the roots of the given multihomogeneous system and the demultihomogenized system.

The Newton polytope of the i^{th} demultihomogenized equation is $Q_i = \prod_{j:n_j>0} \delta_{ij} \Delta(n_j)$ where

$$\Delta(n_j) := \{ (z_{j1}, \dots, z_{jn_j}) \in \mathbb{R}_+^{n_j} : z_{j1} + \ldots + z_{jn_j} \leq 1 \}.$$

Let

$$M(\mathbf{n}, \delta) := \mathcal{MV}\Big(\prod_{j:n_j>0} \delta_{1j} \Delta(n_j), \dots, \prod_{j:n_j>0} \delta_{nj} \Delta(n_j) \Big)$$

be the generic number of complex roots. McLennan [McL98] shows that there are open subsets of the space of real coefficient vectors for which all $M(\mathbf{n}, \delta)$ roots are real.

The next result is a straightforward (both in statement and proof) generalization of one given by [MM97] for the systems arising from games. We adopt the convention that

$$M((0, \dots, 0), \delta) = E((0, \dots, 0), \delta) = 1, \tag{6}$$

which means that the "null system" with no variables and no equations has one root. Given this, all other values of $M(\mathbf{n}, \delta)$ are determined by the recursive relationship specified below. This relationship gives an obvious algorithm for computing $M(\mathbf{n}, \delta)$ that is, in the author's experience, much faster than the Canny-Emeris algorithm [CE95]. (The Canny-Emiris algorithm solves a harder problem, in that it computes the mixed volume of general n-tuples of polytopes.)

Proposition 5.1: For $i = 1, \dots, n$ let δ^{-i} be the $(n-1) \times k$ matrix obtained from δ by deleting the i^{th} row. Then, for each i,

$$M(\mathbf{n}, \delta) = \sum_{j:n_j>0} \delta_{ij} \cdot M((n_1, \dots, n_j - 1, \dots, n_k), \delta^{-i}). \tag{7}$$

Proof: Of course there is no loss of generality in assuming that $i = n$. We compute that

$$\text{vol}(Q_\lambda) = \text{vol}\left(\sum_{i=1}^n \lambda_i Q_i \right) = \text{vol}\left(\sum_{i=1}^n \lambda_i \left(\prod_{j:n_j>0} \delta_{ij} \Delta(n_j) \right) \right) = \text{vol}\left(\prod_{j:n_j>0} (\sum_{i=1}^n \lambda_i \delta_{ij}) \Delta(n_j) \right)$$

$$= \prod_{j:n_j>0} \mathrm{vol}\left(\left(\sum_{i=1}^{n} \lambda_i \delta_{ij}\right)\Delta(n_j)\right) = \prod_{j:n_j>0} \left(\mathrm{vol}(\Delta(n_j)) \cdot \left(\lambda_n \delta_{nj} + \sum_{i=1}^{n-1} \lambda_i \delta_{ij}\right)^{n_j}\right).$$

We can now see that

$$\mathrm{coeff}_{\lambda_1 \cdot \ldots \cdot \lambda_n}(\mathrm{vol}(Q_\lambda)) =$$

$$\sum_{j:n_j>0} \delta_{nj} \cdot \mathrm{coeff}_{\lambda_1 \cdot \ldots \cdot \lambda_{n-1}} \left[n_j \cdot \mathrm{vol}(\Delta(n_j)) \cdot \left(\sum_{i=1}^{n-1} \lambda_i \delta_{ij}\right)^{n_j - 1} \cdot \prod_{\substack{h \neq j \\ n_h > 0}} \left(\mathrm{vol}(\Delta(n_h)) \cdot \left(\sum_{i=1}^{n-1} \lambda_i \delta_{ih}\right)^{n_h}\right)\right].$$

Since $n_j \cdot \mathrm{vol}(\Delta(n_j)) = \mathrm{vol}(\Delta(n_j - 1))$, this is the desired result except in the case of $n = 1$, which means that there is some j such that $n_j = 1$ and $n_h = 0$ for all $h \neq j$. The claim holds in this case by virtue of the convention announced above, and the fact that a nonzero linear functional in two variables has a single root in projective space. ∎

The next two results are analogues of Corollaries 1 and 2.

Proposition 5.2: Suppose there is some $1 \leq k' < k$ such that $\delta_{ij} = 0$ whenever $q(i) \leq k' < j$, and let δ^{11}, δ^{21}, and δ^{22} be as in Section 4. Then

$$M(\mathbf{n}, \delta) = M((n_1, \ldots, n_{k'}), \delta^{11}) \cdot M((n_{k'+1}, \ldots, n_k), \delta^{22}).$$

Proof: Arguing by induction on $n' := n_1 + \ldots + n_{k'}$, (7) implies that $M(\mathbf{n}, \delta)$ does not depend on δ^{21}. When $\delta^{21} = 0$ we have two independent systems, and for generic coefficients the set of solutions of the combined system is the cartesian product of the sets of solutions of the subsystems. ∎

Proposition 5.3: If there are nonnegative integers $d_1, \ldots, d_n$ and $e_1, \ldots, e_k$ such that $\delta'_{ij} = d_i \cdot e_j \cdot \delta_{ij}$, then

$$M(\mathbf{n}, \delta') = \left(\prod_{i=1}^{n} d_i\right) \cdot \left(\prod_{j:n_j>0} e_j^{n_j}\right) \cdot M(\mathbf{n}, \delta).$$

Proof: Obviously it suffices to establish this when all but one of the numbers $d_1, \ldots, d_n, e_1, \ldots, e_k$ are one. If $d_i \neq 1$, the claim follows directly from (7). If $e_j \neq 1$, then

$$M(n_1, \ldots, n_k; \delta') = e_j^{n_j} M(n_1, \ldots, n_k; \delta)$$

holds, obviously, when $n_1 = \ldots = n_k = 0$, and that it holds generally follows from (7) by induction on $n_1 + \ldots + n_k$. ∎

We now consider unmixed systems, finding a case in which the inequality of Conjecture 1 holds strictly.

Theorem 2: Suppose that there is $e = (e_1, \ldots, e_k)$ such that $\delta_{ij} = e_j$ for all i and j. Then

$$E(\mathbf{n}, \delta)^2 \geq M(\mathbf{n}, \delta),$$

with strict inequality unless there is some j with $n_j = n$ (so that $n_h = 0$ for all $h \neq j$).

Proof: (In view of the last result and Corollary 4, it would suffice to prove the claim in the case $e_1 = \ldots = e_k = 1$, but this does not simplify the argument.) The argument is by induction on n, and begins with the observation that, by (6), the claim holds when $\mathbf{n} = (0, \ldots, 0)$. Substituting e_j for δ_{ij}, (7) specializes to

$$M(\mathbf{n}, \delta) = \sum_{n_j > 0} e_j \cdot M((n_1, \ldots, n_j - 1, \ldots, n_k), \delta),$$

so the claim will follow from induction if we can show that

$$E(\mathbf{n}, \delta)^2 \geq \sum_{n_j > 0} e_j \cdot E((n_1, \ldots, n_j - 1, \ldots, n_k), \delta)^2,$$

with strict inequality unless there is some j with $n_j = n$. Substitution of the formula for $E(\mathbf{n}; \delta)^2$ given by Corollary 4, followed by simplification, shows that this last inequality is equivalent to

$$\left(\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})}\right)^2 \geq \sum_{n_j > 0} \left(\frac{\Gamma(\frac{n_j+1}{2})}{\Gamma(\frac{n_j}{2})}\right)^2 = n \cdot \sum_{n_j > 0} \frac{n_j}{n} \cdot \frac{1}{n_j} \left(\frac{\Gamma(\frac{n_j+1}{2})}{\Gamma(\frac{n_j}{2})}\right)^2.$$

Since $n = n_1 + \ldots + n_k$, this inequality (and the fact that it holds strictly unless there is some j with $n_j = n$) is a consequence of the following lemma. ∎

Lemma 5.4: $\frac{1}{n}\left(\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})}\right)^2$ is a strictly increasing function of the integers $n \geq 1$.

Proof: Since $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, $\Gamma(1) = 1$, and $s\Gamma(s) = \Gamma(s+1)$ for all $s > 0$, for odd n the inequality

$$\frac{1}{n}\left(\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})}\right)^2 < \frac{1}{n+1}\left(\frac{\Gamma(\frac{n+2}{2})}{\Gamma(\frac{n+1}{2})}\right)^2 < \frac{1}{n+2}\left(\frac{\Gamma(\frac{n+3}{2})}{\Gamma(\frac{n+2}{2})}\right)^2$$

is equivalent to

$$\frac{1}{n}\left(\frac{\frac{2}{2} \cdot \frac{4}{2} \cdot \ldots \cdot \frac{n-1}{2}}{\sqrt{\pi} \cdot \frac{1}{2} \cdot \frac{3}{2} \cdot \ldots \cdot \frac{n-2}{2}}\right)^2 < \frac{1}{n+1}\left(\frac{\sqrt{\pi} \cdot \frac{1}{2} \cdot \frac{3}{2} \cdot \ldots \cdot \frac{n}{2}}{\frac{2}{2} \cdot \frac{4}{2} \cdot \ldots \cdot \frac{n-1}{2}}\right)^2 < \frac{1}{n+2}\left(\frac{\frac{2}{2} \cdot \frac{4}{2} \cdot \ldots \cdot \frac{n+1}{2}}{\sqrt{\pi} \cdot \frac{1}{2} \cdot \frac{3}{2} \cdot \ldots \cdot \frac{n}{2}}\right)^2.$$

Taking square roots, multiplying by a suitable factor, and simplifying, shows that this in turn is equivalent to

$$\sqrt{\frac{n}{n+1}}\left(\frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \ldots \cdot \frac{n-1}{n} \cdot \frac{n+1}{n}\right) < \frac{\pi}{2} < \sqrt{\frac{n+1}{n+2}}\left(\frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \ldots \cdot \frac{n-1}{n} \cdot \frac{n+1}{n}\right).$$

Straightforward, but tedious, algebraic calculations show that the left hand expression is a strictly increasing function of odd n, while the right hand expression is strictly decreasing. That this inequality holds for all odd $n \geq 1$ follows from the well known [CS61, p. 268] fact that

$$\frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \ldots \cdot \frac{n-1}{n} \cdot \frac{n+1}{n} \to \frac{\pi}{2}. \ ∎$$

In view of the argument proof of Theorem 2, and the fact that Bernshtein's proof involves a similar recursion, it is tempting to speculate as follows:

Conjecture 2:

$$E(\mathbf{n}, \delta)^2 \geq \sum_{j: n_j > 0} \delta_{nj} \cdot E((n_1, \ldots, n_j - 1, \ldots, n_k), \delta^{-n})^2,$$

with strict inequality unless there is some j with $n_j = n$.

By the same argument used to prove Theorem 2, this would imply Conjecture 1. As of this date (June 3, 1998) Conjecture 2 has been verified computationally for one thousand randomly generated systems.

6. π_2 is a Fibration

The next three sections constitute the proof of Theorem 1, which, in both its overall outline and many of its details, follows the analysis of [SS93]. This section develops manifold-theoretic properties of the incidence variety.

Let $\hat{\varepsilon} : \mathcal{H} \times \prod_{j=1}^{k} I\!R^{n_j+1} \to I\!R^n$ be the evaluation map: $\hat{\varepsilon}(f,\zeta) = f(\zeta)$. It is straightforward to argue that 0 is a regular value of the restriction of $\hat{\varepsilon}$ to $M \times \tilde{N}$, so the regular value theorem [GP65] implies that $\{ (f,\tilde{\zeta}) \in M \times \tilde{N} : f(\tilde{\zeta}) = 0 \}$ is a C^∞ $(\dim M)$-dimensional submanifold of $M \times \tilde{N}$. Consequently:

Lemma 6.1: V is a C^∞ submanifold of $M \times N$ with $\dim V = \dim M$.

It will be helpful to distinguish carefully between the two "fibres"

$$V_\zeta := \pi_2^{-1}(\zeta) \subset V \subset M \times N \quad \text{and} \quad \hat{V}_\zeta := \{ f \in M : (f,\zeta) \in V_\zeta \}$$

over a point $\zeta \in N$. For each i let $\hat{V}_{\zeta,i}$ be the set of $f_i \in M_i$ with $f_i(\zeta) = 0$. As the intersection of M_i with a hyperplane, this set is a subsphere of M_i of codimension one. Thus $\hat{V}_\zeta = \hat{V}_{\zeta,1} \times \ldots \times \hat{V}_{\zeta,n}$ has a simple topology that is independent of ζ, and, as one might expect, π_2 is a C^∞ fibration. As usual, to argue this point in detail would be a longwinded and mundane affair, and we shall not do so. It is, perhaps, worth mentioning that the "group" of the fibration may be taken to be G, and that a suitable atlas of coordinate functions(4) is given by the following maps: given $\zeta_0 \in N$, a neighborhood $W \subset N$ of ζ_0, and a C^∞ map $h : W \to G$ satisfying $h(\zeta)\zeta_0 = \zeta$ for all $\zeta \in W$, let $\phi : \hat{V}_{\zeta_0} \times W \to \pi_2^{-1}(W)$ be given by $\phi(f,\zeta) := (h(\zeta)f, \zeta)$.

7. An Integral Formula

Sard's theorem implies that almost all points of M are regular values of π_1. Consider a regular point $(f,\zeta) \in V$ of π_1. Since M and V have the same dimension, the inverse function theorem implies that π_1 maps a neighborhood of (f,ζ) diffeomorphically onto a neighborhood of f. There is a measure μ on V, assigning measure 0 to the set of singular points of π_1, such that if $U \subset V$ is open and the restriction of π_1 to U is injective, then $\mu(U)$ is the measure assigned to $\pi_1(U)$ by the uniform distribution on M. In turn, there is an induced measure $\nu = \mu \circ \pi_2^{-1}$ on N. We have $E(\mathbf{n},\delta) = \mu(V) = \nu(N)$. A detailed consideration of the consequences of these ideas leads to the following integral formula from [SS93, p. 273] which we quote, together with preparatory discussion, verbatim(5):

Let M, N be (real) compact Riemannian manifolds and V a compact submanifold of the product $M \times N$ with $\dim V = \dim M$. Suppose that the restriction $\pi_2 : V \to N$ of the projection $M \times N \to N$ is a locally trivial fibration. Let $V_y := {\pi_2}^{-1}(y)$. Let x be a regular value of $\pi_1 : V \to M$, the restriction of the projection $M \times N \to M$. Define $A(x,y) : T_y(N) \to T_x(M)$ to be the linear map whose graph is the orthogonal complement to $TV_y(x,y)$ in $TV(x,y)$. Let U be an open subset of V and $\#(x)$ be the number of points in $\pi_1^{-1}(x) \cap U$.

Theorem 3.

$$\int_{x \in \pi_1 U} \#(x)\, dM = \int_N \int_{V_y \cap U} \det(A^*(x,y)A(x,y))^{1/2} dV_y dN.$$

(Here $TV(x,y)$ is the tangent space of V at (x,y), and $A^*(x,y) : T_x(M) \to T_y(N)$ is the adjoint of $A(x,y)$.)

(4) This terminology, and the definition of "fibration" we are appealing to, are from [Ste51, §2].

(5) To avoid confusion we have changed the designator from 'Theorem 1' to 'Theorem 3.'

We now develop some of the consequences of this result for the problem at hand. Clearly $\tilde{N}$ is invariant under the action of G on $\prod_{j=1}^k \mathbb{R}^{n_j+1}$, and there is an induced action on N given by

$$O\zeta = ([O_1\tilde{\zeta}_1, -O_1\tilde{\zeta}_1], \ldots, [O_k\tilde{\zeta}_k, -O_1\tilde{\zeta}_k]).$$

Thus we have actions of G on $M \times \tilde{N}$ and $M \times N$ given by $O(f,\tilde{\zeta}) := (Of, O\tilde{\zeta})$ and $O(f,\zeta) := (Of, O\zeta)$. For any $O \in G$, $f \in M$, and $\tilde{\zeta} \in \tilde{N}$ we have $Of(O\tilde{\zeta}) = f \circ O^{-1}(O\tilde{\zeta}) = f(\tilde{\zeta})$, so:

Lemma 7.1: V is an invariant of the action of G on $M \times N$: $OV = V$ for all $O \in G$. Consequently $O(V_\zeta) = V_{O\zeta}$ and $O(\hat{V}_\zeta) = \hat{V}_{O\zeta}$.

We endow each $\tilde{N}_j$ with the Riemannian metric inherited from its inclusion in $\mathbb{R}^{n_j+1}$, and N_j is endowed with the Riemannian metric derived from this in the obvious way. The Riemannian metrics on $\tilde{N}$, N, $M \times \tilde{N}$, and $M \times N$ are the natural product metrics. As in Theorem 3, for $(f,\zeta) \in V$ the function $A(f,\zeta) : T_\zeta(N) \to T_f(M)$ is defined to be the linear map whose graph is the orthogonal complement of $TV_\zeta(f,\zeta)$ in $TV(f,\zeta)$. As is customary, we will not present a proof of the following; although a certain amount of detail is involved, the argument is obvious and straightforward.

Lemma 7.2: For any $\zeta \in N$, $f \in \hat{V}_\zeta$, and $O \in G$,

$$\det(A^*(Of,O\zeta)A(Of,O\zeta)) = \det(A^*(f,\zeta)A(f,\zeta)).$$

The integral $\int_{\hat{V}_\zeta} \det(A^*(f,\zeta)A(f,\zeta))^{1/2}\, df$ does not depend on ζ. In particular, for any open $W \subset N$ and any $\zeta \in N$,

$$\mathbf{E}_M\left(\#(\pi_1^{-1}(\cdot) \cap \pi_2^{-1}(W))\right) = \frac{\mathrm{vol}(W)}{\mathrm{vol}(M)} \cdot \int_{\hat{V}_\zeta} \det(A^*(f,\zeta)A(f,\zeta))^{1/2}\, df.$$

Note that (b) of Theorem 1 is a consequence of the last assertion.

In evaluating the integral over $\hat{V}_\zeta$ in the last result, we are free to let ζ be any convenient point in N. For $j = 1, \ldots, k$ let $\mathbf{e}_{j0}, \mathbf{e}_{j1}, \ldots, \mathbf{e}_{jn_j}$ be the standard unit basis vectors of $\mathbb{R}^{n_j+1}$. We will compute at $\zeta_0 := ([\mathbf{e}_{10}, -\mathbf{e}_{10}], \ldots, [\mathbf{e}_{k0}, -\mathbf{e}_{k0}]) \in N$. Abusing notation, we let ζ_0 also denote the point $(\mathbf{e}_{10}, , \ldots, \mathbf{e}_{k0}) \in \tilde{N}$, and in general our notation will not distinguish between points in $\tilde{N}$ and the corresponding point in N. Let $a_{ij}^0 = (\delta_{ij}, 0, \ldots, 0) \in \mathcal{A}_{ij}$, and let $a_i^0 = (a_{i1}^0, \ldots, a_{ik}^0) \in \mathcal{A}_i$. Since $\zeta_0^a = 0$ for all $a \in \mathcal{A}_i$ other than a_i^0, and $\zeta_0^{a_i^0} = 1$,

$$\hat{V}_{\zeta_0} = \{\, f \in M : f_{ia_i^0} = 0 \ (i = 1, \ldots, n) \,\}.$$

Suppose that $\delta_{ij} > 0$. For $h = 1, \ldots, n_j$, let a_i^{jh} be a_i^0 with a_{ij}^0 replaced by $(\delta_{ij} - 1, 0, \ldots, 0, 1, 0, \ldots, 0)$ (the '1' is component h). The tangent space $T_{\zeta_0}\tilde{N}$ is spanned by the n vectors

$$\mathbf{b}_{jh} := (0, \ldots, \mathbf{e}_{jh}, \ldots, 0) \qquad (1 \le j \le k, 1 \le h \le n_j).$$

Then

$$Df_i(\zeta_0)\mathbf{b}_{jh} = \begin{cases} f_{ia_i^{jh}} & \text{if } \delta_{ij} > 0, \\ 0 & \text{if } \delta_{ij} = 0. \end{cases}$$

In this way we obtain a description of $Df(\zeta_0)$ as an $n \times n$ matrix with rows indexed by $f_1, \ldots, f_n$ and columns indexed by the pairs (j,h), with this (i,jh)–entry.

We now compute $A(f,\zeta_0)$ concretely. We have $T_f(M) \approx T_{f_1}(M_1) \times \ldots \times T_{f_n}(M_n)$, and in our calculations below we will treat each $T_{f_i}(M_i)$ as a hyperplane in $\mathbb{R}^{\mathcal{A}_i}$. We identify $T_{\zeta_0}(N)$ with $T_{\zeta_0}(\tilde{N})$, so that a vector $w \in T_{\zeta_0}N$ may be written as $\sum_{j=1}^{k}\sum_{h=1}^{n_j} w_{jh}\mathbf{b}_{jh}$. Differentiation of the formula $f(\zeta) = 0$ shows that a pair

$$(v, w) \in T_{(f,\zeta_0)}(M \times N) \approx T_f(M) \times T_{\zeta_0}(N)$$

is in $TV(f, \zeta_0)$ if and only if, for each i,

$$v_{ia_i^0} = -\sum_{j:\delta_{ij}>0}\sum_{h=1}^{n_j} f_{ia_i^{jh}} w_{jh}.$$

A pair (v, w) is in the orthogonal complement of $TV_{\zeta_0}(f, \zeta_0)$ in $T_f(M) \times T_{\zeta_0}(N)$ if and only if, for each i, $v_{ia} = 0$ for all $a \in \mathcal{A}_i$ other than a_i^0. To see this, begin by observing that the orthogonal complement of $TV_{\zeta_0}(f, \zeta_0)$ in $T_f(M) \times T_{\zeta_0}(N)$ is the cartesian product of $T_{\zeta_0}(N)$ with the orthogonal complement of $T_f(\hat{V}_{\zeta_0})$ in $T_f(M)$. Since $T_f(M) \approx T_{f_1}(M_1) \times \ldots \times T_{f_n}(M_n)$, the orthogonal complement of $T_f(\hat{V}_{\zeta_0})$ in $T_f(M)$ will be the cartesian product of the orthogonal complements of the $T_{f_i}(\hat{V}_{\zeta_0 i})$ in $T_{f_i}(M_i)$. A vector $v_i \in \mathbb{R}^{\mathcal{A}_i}$ is in $T_{f_i}(M_i)$ if it is orthogonal to f_i, and if it is also orthogonal to $T_{f_i}(\hat{V}_{\zeta_0 i})$ then it is orthogonal to the hyperplane in $\mathbb{R}^{\mathcal{A}_i}$ spanned by $\hat{V}_{\zeta_0 i}$, meaning that $v_{ia} = 0$ for all $a \notin \mathcal{A}_i \setminus \{a_i^0\}$. Conversely, if $v_{ia} = 0$ for all $a \notin \mathcal{A}_i \setminus \{a_i^0\}$, then v_i is orthogonal to both f_i and $T_{f_i}(\hat{V}_{\zeta_0 i})$. Thus the orthogonal complement of $TV_{\zeta_0}(f, \zeta_0)$ in $T_{(f,\zeta_0)}(M \times N)$ consists of those (v, w) with $v_{ia} = 0$ for all i and $a \notin \mathcal{A}_i \setminus \{a_i^0\}$.

Combining the conclusions above, we see that the (i, a_i^0)-component of $A(f, \zeta_0)w$ is given by the formula above, while all other components vanish. Thus $A(f, \zeta_0)$ is described by a matrix with rows indexed by the various pairs (i, a) with $a \in \mathcal{A}_i$ and columns indexed by the n pairs (j, h), with the row for (i, a) vanishing unless $a = a_i^0$, and the (j, h) entry of row (i, a_i^0) being $-f_{ia_i^{jh}}$ if $\delta_{ij} > 0$ and 0 otherwise. This matrix consists of the rows of $-Df(\zeta_0)$ together with a number of rows of zeros. Since the matrix of the adjoint of a linear transformation is the transpose of the matrix of the transformation, and the domain and range of $Df(\zeta_0)$ have the same dimension, elementary facts concerning matrices and determinants yield:

Proposition 7.3: For all $f \in \hat{V}_{\zeta_0}$,

$$A^*(f, \zeta_0)A(f, \zeta_0) = Df(\zeta_0)^* Df(\zeta_0)$$

and $\det(Df(\zeta_0)^* Df(\zeta_0))^{1/2} = |\det Df(\zeta_0)|$.

For the calculations that occupy the remainder it is notationally convenient to define $\mathcal{D}(f, \zeta_0) := |\det Df(\zeta_0)|$. With this notation we may combine the last two results, and set the stage for the computations in the next section, as follows:

Lemma 7.4:

$$E(\mathbf{n}, \delta) = \frac{\mathrm{vol}(N) \cdot \mathrm{vol}(V_{\zeta_0})}{\mathrm{vol}(M)} \cdot \mathbf{E}_{V_{\zeta_0}}(\mathcal{D}(\cdot, \zeta_0)). \tag{8}$$

8. A Change of Variables

As we pointed out above, (b) of Theorem 1 has already been established, and the remaining task is to prove (a) of that result. This is largely a matter of exploiting the fact that only certain of the coefficients f_{ia} enter $Df(\zeta_0)$, so that by changing variables we obtain an integral on a domain of low dimension.

Let Λ_i be the unit sphere in the subspace of $\mathcal{H}_i$ consisting of those f_i with $f_{ia} = 0$ for all a that are a_i^{jh} for some indices j and h such that $\delta_{ij} > 0$ and $1 \leq h \leq n_j$. Note that $\Lambda_i \subset \hat{V}_{\zeta_0,i}$. Let J_i be the unit sphere in the subspace of $\mathcal{H}_i$ consisting of those f_i with $f_{ia_i^0} = 0$ and $f_{ia_i^{jh}} = 0$ for all j and h such that $\delta_{ij} > 0$ and $1 \le h \le n_j$. Note that

$$\dim M_i = \dim \Lambda_i + \dim J_i + 2.$$

Let $\Lambda := \Lambda_1 \times \ldots \times \Lambda_n$ and $J = J_1 \times \ldots \times J_n$.

The following two formulas will be proved subsequently:

$$\mathbf{E}_{\hat{V}_{\zeta_0}}(\mathcal{D}(\cdot,\zeta_0)) = \frac{\mathrm{vol}(J)\cdot\mathrm{vol}(\Lambda)}{\mathrm{vol}(\hat{V}_{\zeta_0})}\cdot 2^{-n}\cdot\Big(\prod_{i=1}^{n}\frac{\Gamma(\frac{\dim J_i+1}{2})\Gamma(\frac{\dim \Lambda_i+2}{2})}{\Gamma(\frac{\dim\Lambda_i+\dim J_i+3}{2})}\Big)\cdot\mathbf{E}_{\Lambda}(\mathcal{D}(\cdot,\zeta_0)); \tag{9}$$

$$\mathbf{E}_{\Lambda}(\mathcal{D}(\cdot,\zeta_0)) = \prod_{i=1}^{n}\frac{\Gamma(\frac{\dim\Lambda_i+1}{2})\Gamma(\frac{n+1}{2})}{\Gamma(\frac{\dim\Lambda_i+2}{2})\Gamma(\frac{n}{2})}\cdot\mathbf{E}_L(|\det\Xi(\cdot)|). \tag{10}$$

We now explain how these formulas lead to (a) of Theorem 1. Combining (8), (9), and (10) yields:

$$E(\mathbf{n},\delta) =$$

$$\frac{\mathrm{vol}(N)\cdot\mathrm{vol}(J)\cdot\mathrm{vol}(\Lambda)}{\mathrm{vol}(M)}\cdot 2^{-n}\cdot\Big(\prod_{i=1}^{n}\frac{\Gamma(\frac{\dim J_i+1}{2})\Gamma(\frac{\dim\Lambda_i+1}{2})}{\Gamma(\frac{\dim\Lambda_i+\dim J_i+3}{2})}\Big)\cdot\Big(\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})}\Big)^n\cdot\mathbf{E}_L(|\det\Xi(\cdot)|). \tag{11}$$

Recall the well known [Fed69 p. 251] formula

$$\mathrm{vol}(S^{m-1}) = 2\frac{\Gamma\left(\frac{1}{2}\right)^m}{\Gamma\left(\frac{m}{2}\right)}, \tag{12}$$

which is valid for any integer $m \ge 1$. Of course the volume of $(m-1)$-dimensional real projective space is half this. Since $\dim N_j = n_j$, we have

$$\begin{aligned}\frac{\mathrm{vol}(N)\cdot\mathrm{vol}(J)\cdot\mathrm{vol}(\Lambda)}{\mathrm{vol}(M)} &= \Big(\prod_{j=1}^{k}\mathrm{vol}(N_j)\Big)\Big(\prod_{i=1}^{n}\frac{\mathrm{vol}(J_i)\cdot\mathrm{vol}(\Lambda_i)}{\mathrm{vol}(M_i)}\Big)\\ &= \Big(\prod_{j=1}^{k}\frac{\Gamma(\frac{1}{2})^{n_j+1}}{\Gamma(\frac{n_j+1}{2})}\Big)\cdot 2^n\cdot\Big(\prod_{i=1}^{n}\frac{\Gamma(\frac{\dim\Lambda_i+\dim J_i+3}{2})}{\Gamma(\frac{1}{2})\cdot\Gamma(\frac{\dim J_i+1}{2})\cdot\Gamma(\frac{\dim\Lambda_i+1}{2})}\Big)\end{aligned} \tag{13}$$

It is now straightforward to combine (11) with this to obtain (a) of Theorem 1.

It now remains to establish (9) and (10). We begin by describing, in general terms, the change of variables at the heart of the calculations. For integers m, h with $1 \le h \le m-1$, let S^{m-1} be the unit sphere in $\mathbb{R}^m \approx \mathbb{R}^h \times \mathbb{R}^{m-h}$, and let S^{h-1} and S^{m-h-1} be the unit spheres in the cartesian factors. Define

$$g_{m;h} : S^{h-1}\times S^{m-h-1}\times(0,1)\to S^{m-1} \quad\text{by}\quad g_{m;h}(y,z,t) := ty + (1-t^2)^{1/2}z.$$

Note that $g_{m;h}$ is a diffeomorphism onto its image, which is an open subset of S^{m-1} whose complement has measure zero. Preparatory to applying the change of variables formula, observe that

$$\|\frac{\partial g_{m;h}}{\partial t}(y,z,t)\| = \|y - \frac{t}{\sqrt{1-t^2}}z\| = (1+\frac{t^2}{1-t^2})^{1/2} = (1-t^2)^{-1/2}.$$

With respect to orthonormal bases for T_yS^{h-1} and T_zS^{m-h-1}, the matrix of $Dg_{m;h}(y,z,t)$ is a diagonal matrix, and computing the diagonal elements, then multiplying them together, gives

$$\begin{aligned}\mathcal{D}(g_{m;h},(y,z,t)) &= t^{h-1}\cdot\left((1-t^2)^{1/2}\right)^{m-h-1}\cdot(1-t^2)^{-1/2}\\ &= t^{h-1}(1-t^2)^{\frac{m-h}{2}-1}.\end{aligned} \tag{14}$$

The change of variables given by $g_{m;h}$ is already useful in the following calculation.

Lemma 8.1: Let $\pi_h : S^{m-1} \to \mathbb{R}^h$ be the projection $\pi_h(x) := (x_1, \ldots, x_h)$. Then

$$\mathbf{E}_{S^{m-1}}(\|\pi_h(\cdot)\|) = \frac{\Gamma(\frac{h+1}{2})\Gamma(\frac{m}{2})}{\Gamma(\frac{h}{2})\Gamma(\frac{m+1}{2})}.$$

Proof: The following calculation applies the change of variables formula, then (14):

$$\begin{aligned}\mathbf{E}_{S^{m-1}}(\|\pi_h(\cdot)\|) &= \frac{1}{\mathrm{vol}(S^{m-1})}\int_{S^{m-1}\times S^{m-h-1}\times(0,1)} \|\pi_h(g_{m;h}(y,z,t))\| \cdot \mathcal{D}(g_{m;h},(y,z,t))\, d(y,z,t) \\ &= \frac{1}{\mathrm{vol}(S^{m-1})}\int_{S^{m-1}\times S^{m-h-1}\times(0,1)} t\cdot t^{h-1}(1-t^2)^{\frac{m-h}{2}-1}\, d(y,z,t) \\ &= \frac{\mathrm{vol}(S^{h-1})\cdot \mathrm{vol}(S^{m-h-1})}{\mathrm{vol}(S^{m-1})}\int_0^1 t^h(1-t^2)^{\frac{m-h}{2}-1}\, dt.\end{aligned}$$

The desired assertion is now obtained by using (12) to evaluate the sphere volumes and applying the formula [SS93, p. 273]

$$\int_0^1 t^h(1-t^2)^{\frac{m-h}{2}-1}\, dt = \frac{1}{2}\frac{\Gamma(\frac{m-h}{2})\Gamma(\frac{h+1}{2})}{\Gamma(\frac{m+1}{2})}. \blacksquare \tag{15}$$

The specific application of $g_{m;h}$ that we need is given by defining $g : \Lambda \times J \times (0,1)^n \to \hat{V}_{\zeta_0}$ by

$$g(\lambda, \psi, t) := (g_1(\lambda_1, \psi_1, t_1), \ldots, g_n(\lambda_n, \psi_n, t_n))$$

where

$$g_i := g_{(\dim\Lambda_i + \dim J_i + 2);(\dim\Lambda_i + 1)} : \Lambda_i \times J_i \times (0,1) \to \hat{V}_{\zeta_0,i}.$$

Applying (14) yields

$$\mathcal{D}(g,(\lambda,\psi,t)) = \prod_{i=1}^n (1-t_i{}^2)^{\frac{\dim\Lambda_i - 1}{2}} \cdot t_i^{\dim J_i}. \tag{16}$$

Lemma 8.2: For all $(\lambda, \psi, t) \in \Lambda \times J \times (0,1)^n$,

$$\mathcal{D}(g(\lambda,\psi,t),\zeta_0) = \Big(\prod_{i=1}^n (1-t_i{}^2)^{\frac{1}{2}}\Big)\mathcal{D}(\lambda,\zeta_0).$$

Proof: As we saw in the proof of Proposition 7.3, the i^{th} row of $Df(\zeta_0)$ corresponds to $Df_i(\zeta_0)$, with entries corresponding to the monomials that are the coordinates of λ_i. In comparing $D\lambda(\zeta_0)$ with $D[g(\lambda,\psi,t)](\zeta_0)$, we see that the entries of the ith row of the latter matrix are obtained from the corresponding entries of the former matrix by multiplying by $(1-t_i{}^2)^{\frac{1}{2}}$, so the claim follows from elementary properties of the determinant. ∎

The next two results complete the proof of (a) of Theorem 1 by establishing (9) and (10).

Lemma 8.3:

$$\mathbf{E}_{\hat{V}_{\zeta_0}}(\mathcal{D}(\cdot,\zeta_0)) = \frac{\mathrm{vol}(J)\cdot\mathrm{vol}(\Lambda)}{\mathrm{vol}(\hat{V}_{\zeta_0})}\cdot 2^{-n}\cdot\Big(\prod_{i=1}^n \frac{\Gamma(\frac{\dim J_i + 1}{2})\Gamma(\frac{\dim\Lambda_i + 2}{2})}{\Gamma(\frac{\dim\Lambda_i + \dim J_i + 3}{2})}\Big)\cdot \mathbf{E}_\Lambda(\mathcal{D}(\cdot,\zeta_0)).$$

Proof: Applying the change of variables formula, the computations above, and (15), yields:

$$\begin{aligned}
\mathbf{E}_{\hat{V}_{\zeta_0}}(\mathcal{D}(\cdot,\zeta_0)) &= \frac{1}{\mathrm{vol}(\hat{V}_{\zeta_0})}\cdot\int_{\hat{V}_{\zeta_0}} \mathcal{D}(f,\zeta_0)\,df \\
&= \frac{1}{\mathrm{vol}(\hat{V}_{\zeta_0})}\cdot\int_{\Lambda\times J\times(0,1)^n} \mathcal{D}(g(\lambda,\psi,t),\zeta_0)\cdot\mathcal{D}(g,(\lambda,\psi,t))\,d(\lambda,\psi,t) \\
&= \frac{1}{\mathrm{vol}(\hat{V}_{\zeta_0})}\cdot\int_{\Lambda\times J\times(0,1)^n} \left(\prod_{i=1}^n (1-t_i{}^2)^{\frac{\dim\Lambda_i}{2}}\cdot t_i^{\dim J_i}\right)\cdot\mathcal{D}(\lambda,\zeta_0)\,d(\lambda,\psi,t) \\
&= \frac{\mathrm{vol}(J)\cdot\mathrm{vol}(\Lambda)}{\mathrm{vol}(\hat{V}_{\zeta_0})}\cdot\left(\prod_{i=1}^n\int_0^1 (1-t_i{}^2)^{\frac{\dim\Lambda_i}{2}}\cdot t_i^{\dim J_i}\,dt_i\right)\cdot\mathbf{E}_\Lambda(\mathcal{D}(\cdot,\zeta_0)) \\
&= \frac{\mathrm{vol}(J)\cdot\mathrm{vol}(\Lambda)}{\mathrm{vol}(\hat{V}_{\zeta_0})}\cdot\left(\prod_{i=1}^n\frac{\Gamma(\frac{\dim J_i+1}{2})\Gamma(\frac{\dim\Lambda_i+2}{2})}{2\Gamma(\frac{\dim\Lambda_i+\dim J_i+3}{2})}\right)\cdot\mathbf{E}_\Lambda(\mathcal{D}(\cdot,\zeta_0)).\ \blacksquare
\end{aligned}$$

Lemma 8.4:

$$\mathbf{E}_\Lambda(\mathcal{D}(\cdot,\zeta_0)) = \prod_{i=1}^n\frac{\Gamma(\frac{\dim\Lambda_i+1}{2})\Gamma(\frac{n+1}{2})}{\Gamma(\frac{\dim\Lambda_i+2}{2})\Gamma(\frac{n}{2})}\cdot\mathbf{E}_L(|\det\Xi(\cdot)|).$$

Proof: The argument refers to the maps

$$X_i\xrightarrow{p_i}Y_i\xrightarrow{c_i}Z_i \qquad\text{and}\qquad Z_i\setminus\{0\}\xrightarrow{\rho_i}\Lambda_i\quad(i=1,\dots,n)$$

which we now explain. As in the statement of Theorem 1, X_i is $\mathbb{R}^n$ with coordinates indexed by the pairs (j,h) $(1\le j\le k, 1\le h\le n_j)$. Let $Y_i := \mathbb{R}^{\{(j,h):\delta_{ij}>0\}}$, and let $p_i : X_i\to Y_i$ be the obvious projection. Let $Z_i\subset\mathcal{H}_i$ be the set of polynomials with support $\{a_i^{jh}:\delta_{ij}>0\}$, and let $c_i : Y_i\to Z_i$ be the map taking $y_i\in Y_i$ to the polynomial

$$\sum_{j:\delta_{ij}>0}\sum_{h=1}^{n_j}\sqrt{\delta_{ij}}\,y_i^{jh}\zeta^{a_i^{jh}}.$$

Observing that Λ_i is the unit sphere in Z_i, let $\rho_i : Z_i\setminus\{0\}\to\Lambda_i$ be the map $\rho_i(z_i) := z_i/||z_i||_i$. (Recall that $||\cdot||_i$ is the norm derived from $\langle\cdot,\cdot\rangle_i$.)

By evaluating the defining formula, we obtain $\eta(a_i^{jh}) = 1/\delta_{ij}$ when $\delta_{ij}>0$, so that c_i is a linear isometry. The map Ξ_i described in the statement of Theorem 1 is (the numerical representation of) of $c_i\circ p_i$. Since p_i is an orthogonal projection, and c_i is an isometry, if $\tilde{\ell}_i$ is a uniformly distributed point in $L_i\setminus\Xi_i^{-1}(0)$, then $||\Xi(\tilde{\ell}_i)||_i$ and $\rho_i(\Xi(\tilde{\ell}_i))$ are statistically independent, and $\rho_i(\Xi(\tilde{\ell}_i))$ is uniformly distributed in Λ_i. These facts allow us to compute that

$$\begin{aligned}
\mathbf{E}_L\Big(|\det(\Xi_1(\ell_1),\dots,\Xi_n(\ell_n))|\Big) &= \mathbf{E}_L\Big(|\det(\rho_i(\Xi_1(\ell_1)),\dots,\rho_i(\Xi_n(\ell_n)))|\cdot\prod_i||\Xi(\ell_i)||_i\Big) \\
&= \mathbf{E}_\Lambda\Big(\mathcal{D}(\cdot,\zeta_0)\Big)\cdot\prod_i\mathbf{E}_{L_i}(||\Xi_i(\ell_i)||_i).
\end{aligned}$$

(Here the second equality follows from independence and the description of $\det D\lambda(\zeta_0)$ developed in the discussion leading up to Proposition 7.3.) The final step in the calculation is the evaluation of $\mathbf{E}_{L_i}(||\Xi_i(\ell_i)||_i)$ by means of Lemma 8.1. ■

References

[Ber75] D. N. Bernshtein, The number of roots of a system of equations, *Functional Analysis and its Applications* **9** (1975), 183–185.

[CS61] R. D. Carmichael and E. R. Smith, *Mathematical Tables and Formulas*, Dover, New York, (1961).

[CE95] J. F. Canny and I. Emiris, Efficient Incremental Algorithms for the Sparse Resultant and the Mixed Volume, *Journal of Symbolic Computation* **20** (1995), 117–149.

[EK95] A. Edelman and E. Kostlan, How many zeros of a random polynomial are real?, *Bulletin of the American Mathematical Society* **32** (1995), 1–37.

[Fed69] H. Federer, *Geometric Measure Theory*, Springer, New York, (1969).

[FT91] D. Fudenberg and J. Tirole, *Game Theory*, MIT Press, Cambridge, (1991).

[Gir90] V.L. Girko, *The Theory of Random Determinants*, Kluwer, Boston, (1990).

[GP65] V. Guillemin and A. Pollack, *Differential Topology*, Prentice-Hall, Englewood Cliffs, (1965).

[Kac43] M. Kac, On the average number of real roots of a random algebraic equation, *Bulletin of the American Mathematical Society* **49** (1943), 314–320 and 938.

[Kus75] 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.

[MM97] R. D. McKelvey and A. McLennan, The maximal number of regular totally mixed Nash equilibria, *Journal of Economic Theory* **72** (1997), 411–425.

[McL97] A. McLennan, On the Expected Number of Nash Equilibria of a Normal Form Game, mimeo, University of Minnesota, (1997).

[McL98] A. McLennan, The maximal number of real roots of a multihomogeneous system of polynomial equations, forthcoming in *Beiträge zur Algebra und Geometrie*, (1998).

[Roj96] J. M. Rojas, On the average number of real roots of certain random sparse polynomial systems, *Lectures on Applied Mathematics Series*, ed. by J. Renegar, M. Shub, and S. Smale, American Mathematical Society, (1996).

[SS93] M. Shub and S. Smale, Complexity of Bezout's theorem II: volumes and probabilities, *Computational Algebraic Geometry* (F. Eyssette and A. Galligo, eds.), Progr. Math., vol. 109 (1993), Birkhauser, Boston, 267–285.

[Ste51] N. Steenrod, *The Topology of Fibre Bundles*, Princeton University Press, Princeton (1951).