The Expected Number of Nash Equilibria of a Normal Form Game

by

Andrew McLennan

Discussion Paper #315, May 2002

Center for Economic Research

Department of Economics

University of Minnesota

Minneapolis, Minnesota 55455

The Expected Number of Nash Equilibria of a Normal Form Game*

by

Andrew McLennan

Department of Economics
University of Minnesota
271 19th Avenue South
Minneapolis, MN 55455
mclennan@atlas.socsci.umn.edu
http://www.econ.umn.edu/~mclennan

April, 2002

* This research was supported in part by National Science Foundation grant SBR-9308862 to the University of Minnesota. During various phases of the manuscript's preparation I have enjoyed the hospitality and support of Bristol University, Caltech, the University of Chicago, and the Institute for Social and Economic Research of Osaka University. I am grateful for the comments of participants of presentations at the University of Minnesota, the Conference on Mathematical Economics at Keio University, Caltech, Pennsylvania State University, the Midwest Mathematical Economics Conference at the University of Illinois, and the Southeast Economic Theory Conference at Georgetown University. This work has benefited from numerous discussions with Richard McKelvey and Maurice Rojas. Anonymous referees and the editor made very helpful comments on an earlier draft.

Running head: **The Expected Number of Nash Equilibria**

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

Abstract

Fix finite pure strategy sets $S_1, \ldots, S_n$, and let $S = S_1 \times \cdots \times S_n$. In our model of a random game the agents' payoffs are statistically independent, with each agent's payoff uniformly distributed on the unit sphere in $\mathbb{R}^S$. For given nonempty $T_1 \subset S_1, \ldots, T_n \subset S_n$ we give a computationally implementable formula for the mean number of Nash equilibria in which each agent i's mixed strategy has support T_i. The formula is the product of two expressions. The first is the expected number of totally mixed equilibria for the truncated game obtained by eliminating pure strategies outside the sets T_i. The second may be construed as the "probability" that such an equilibrium remains an equilibrium when the strategies in the sets $S_i \backslash T_i$ become available. *Journal of Economic Literature* Classification Number C72.

The Expected Number of Nash Equilibria of a Normal Form Game

1. Introduction

This paper states and proves a formula characterizing the mean number of equilibria of a normal form game, relative to a particular distribution on the space of possible payoffs. In the interest of brevity, and to maintain a single focus, the body of the paper will be narrowly concerned with the mathematical underpinnings of the formula. Implications of the formula that are known at present will be sketched in broad strokes in this section and studied in more detail in other papers. But before entering into such a conceptual discussion we first give a precise formulation of the model, and a brief description of the result.

Fix nonempty finite pure strategy sets $S_1, \ldots, S_n$, and let $S = S_1 \times \cdots \times S_n$. In our model of a random game the agents' payoffs are statistically independent, with each agent's payoff uniformly distributed on the unit sphere in $\mathbb{R}^S$. Since Nash equilibrium is invariant under positive affine transformations of the vNM payoffs, from the point of view of the induced distribution of equilibria it would be equivalent to assume that the payoffs of each agent at each strategy profile are i.i.d. normally distributed random variables.

A *support* is a n-tuple $T = (T_1, \ldots, T_n)$ where each T_i is a nonempty subset of S_i. The support of a Nash equilibrium is the support T in which T_i is the set of pure strategies that receive positive probability under agent i's mixed strategy. Recall (Harsanyi (1973)) that for generic payoffs, hence with probability one under our model of a random game, all equilibria are regular. One consequence is that, with probability one, all equilibria are strict, meaning that no agent has a pure strategy that is assigned no probability by her mixed strategy, but which would yield the equilibrium expected utility. Also, regular equilibria satisfy all major refinements, including Kohlberg-Mertens (1986) stability.

Our main result is a formula giving the mean number of Nash equilibria with support T. Without saying anything about the particulars of the formula, the rest of this section will describe some of the implications of this result in connection with larger conceptual issues and related directions of research. At this point we are unable to say much about the purely mathematical interest of the analysis, which is considerable.

Many doubts have been expressed about the scientific status of the Nash equilibrium concept, but certainly some of the most important derive from the concept's computational complexity. In introspective theories[1] of how equilibrium is (perhaps only sometimes) achieved, each agent must be able, somehow, to compute her component of the equilibrium that is played, using only the data defining the game. Recently, partly in response to the work of Bernheim (1984) and Pearce (1984), which showed that introspection based on common knowledge of rationality implies the notion of rationalizability, which is weaker than Nash equilibrium, it has become more popular to motivate the salience of equilibrium in terms of scenarios in which it results from evolution or learning. (Cf., Fudenberg and Levine (1998), Samuelson (1997), Weibull (1995).) Many variations are

[1] See Kuhn (1994) for a discussion of the history of interpretations of the Nash equilibrium concept.

possible, but in general such stories do not require the individual agents to be computationally sophisticated. Nonetheless, computational complexity remains important, since in these interpretations the social adjustment process acts as a computer which solves for (by converging to) an equilibrium. The mean number of equilibria is less directly relevant to this computation than to the problem of finding all equilibria; nonetheless, the number of equilibria provides some information concerning the ruggedness of the terrain that such an adjustment process must traverse.

The formula given by the Main Theorem has more tangible implications for the complexity of the problem of computing the set of all Nash equilibria, as measured by the concepts of theoretical computer science. Standard notions of complexity depend on the rate at which the resources (time and/or memory) required by an algorithm grow as the size of the input increases. The most fundamental division is between algorithms whose running times grow at rates that are bounded by polynomial functions of characteristics of the input, and those for which the rate of growth is exponential, or perhaps even faster. Generally, algorithms with polynomial time and space requirements are described as "practical," while those with exponential rates of growth are regarded as "impractical." But many interesting games are small, and still hard for people to solve by hand, so in game theory even exponential algorithms have considerable practical utility.

The most concrete and immediate applications of the formula given here show that the mean numbers of equilibria of various sorts grow at rates that are bounded below by exponential functions. These results feed into a standard tactic for establishing an exponential lower bound on the complexity class of a computation: show that the size of the output grows at an exponential rate. The results derived from the formula are somewhat unusual insofar as most analyses in computer science relate to "worst case complexity," which measures the rate of growth, as a function of input size, of the resource requirements of the most burdensome input of each size. The prevalence of worst case complexity in computer science is not due to a belief that it is more relevant to practical issues of computation than mean complexity, but rather to the fact that worst case complexity is systematically more tractable: mean complexity requires an analysis of all possible inputs, whereas consideration of a single sequence of inputs of increasing size can establish a lower bound on worst case complexity. An economically important example is the simplex algorithm of linear programming, which has been shown to have an exponential worst case complexity, but has a mean running time (under a natural model of a random problem instance) that grows quadratically with the size of the problem[(2)].

One result of this sort has to do with two player games. In Berg and McLennan (2002) techniques from statistical mechanics are applied to the formula given by the Main Theorem to show that the mean number of equilibria of a two person game in which both agents have N pure strategies is $\exp(N[B + O(\log N/N)])$, where $B \approx 0.281644$ is a constant. This implies an exponential lower bound on the mean complexity of the computation, but obvious procedures have exponential worst case running times. Thus both the mean and worst case complexity are seen to be exponential.

Berg and McLennan (2002) establish other interesting results. There is a constant β, which is approximately 0.3195, such that, for large N, most equilibria of the game in

(2) See Gritzman and Klee (1993, §7) for a survey of related literature.

which both players have N pure strategies assign positive probability to approximately βN pure strategies(3). Also, if the number M of pure strategies of one agent is held fixed, when the number N of pure strategies of the second agent is large the mean number of equilibria is approximately $\left(\sqrt{\pi \log N}/2\right)^{M-1}/\sqrt{M}$.

Similar calculations for games with more than two agents seem likely to be feasible, but may not yield such exact results. Our formula is a product of two factors: (a) the mean number of totally mixed equilibria of a random game which has the support $(T_1, \ldots, T_n)$ as its n-tuple of sets of pure strategies; (b) the probability, for a random game on the strategy sets $(S_1, \ldots, S_n)$, that a totally mixed equilibrium of the truncated game obtained by eliminating the pure strategies in $S_1 \setminus T_1, \ldots, S_n \setminus T_n$ is an equilibrium of the given game. For two player games the first factor is a negative power of two, and second factor has a characterization that allows accurate approximation, both computationally and in the relevant theoretical calculations. But when $n \geq 3$ the first factor has a more complex expression.

There are already some results pertaining to the case $n \geq 3$. As will be explained in more detail as we go along, results from McLennan (2002) give a lower bound in terms of the maximal number of totally mixed equilibria of a game with strategy sets $T_1, \ldots, T_n$. McKelvey and McLennan (1997) give a recursive combinatoric characterization of this number, and also provide upper and lower bounds in closed form. Combining these results with the methods of statistical mechanics should lead at least to lower bounds on asymptotic rates of growth of the mean number of equilibria of all sorts. One result described in Section 4 is that when $n \geq 12$, the mean number of totally mixed equilibria of a random game in which each agent has k pure strategies is bounded below by an exponential function of k.

We characterize not only the mean number of equilibria with support $(T_1, \ldots, T_n)$, but also their distribution in the space of mixed strategy profiles. That is, there is a measure that assigns the mean number of equilibria lying in E to each measurable set E in the space of mixed strategy profiles. Representing this measure as the mean number of equilibria with support $(T_1, \ldots, T_n)$ times a probability measure on the space of mixed strategy profiles, the probability measure turns out to be a product of probability measures on the agents' spaces of mixed strategies. As is explained in detail in McLennan (1999), to a surprising extent the mass of these measures is concentrated near the barycenters of the agents' simplices of mixed strategies.

Numerical computation of the mean number of equilibria is simple to program, and feasible for fairly large games. (The time consuming step is Monte Carlo approximation of the mean absolute value of a random square matrix with as many rows and columns as the dimension of the space of mixed strategies, which is the total number of pure strategies less the number of agents.) McLennan (1999) and Berg and McLennan (2002) present various computational results which, in addition to their intrinsic interest, suggest various monotonicity conjectures concerning the relationship between the numbers of pure strategies for the agents and the mean number of equilibria.

(3) Expressed precisely, the result is that for any $\epsilon > 0$ it is the case, for sufficiently large N, that the mean number of equilibria that assign positive probability to between $(1-\epsilon)\beta N$ and $(1+\epsilon)\beta N$ of each player's pure strategies is at least $1-\epsilon$ times the mean number of equilibria of all sorts.

Some additional related literature is as follows. McKelvey and McLennan (1997) characterize the maximal number of totally mixed equilibria. The maximal number of equilibria (on all supports) of a two player normal form game has been studied by Quint and Shubik (1997), Keiding (1995), McLennan and Park (1999), and von Stengel (1999), but is not completely understood. In contrast, as a result of work of Dresher (1970), Powers (1990), and Stanford (1995), the asymptotic distribution of pure equilibria for random normal form games has been characterized with considerable precision. The maximal number of pure strategy Nash equilibria for generic payoffs is determined in McLennan (1997).

The next section states our main result. The contents of the remainder are described at the end of that section.

2. Statement of the Result

This section presents little more than the minimal amount of information required to state the main result. This consists of a description of the model of a random game for the given normal form, the definition of (strict) Nash equilibrium as a system of equations and inequalities, the definitions of objects entering the formula for the mean number of equilibria, and finally the statement itself.

2.1. *Conventions Concerning Manifolds*

In general, whenever X is a d-dimensional C^1 submanifold of a Euclidean space, $\mathrm{vol}_X(\cdot)$ (or simply $\mathrm{vol}(\cdot)$ if there is no ambiguity) denotes the measure on X corresponding to the notion of d-dimensional volume derived from the inner product of the ambient space. In integrals we will sometimes write X in place of vol_X, e.g., $\int f\, dX$. When $\mathrm{vol}(X)$ is finite, the *uniform distribution* on X is the probability measure $\mathbf{U}_X := \mathrm{vol}_X(\cdot)/\mathrm{vol}_X(X)$.

We will also consider integrals over d-dimensional real projective space, denoted by $\mathbf{P}(\mathbb{R}^{d+1})$, which is the set of one dimensional linear subspaces of $\mathbb{R}^{d+1}$ endowed with the C^∞ differential structure that makes the map $x \mapsto \mathrm{span}\{x\}$ from S^d to $\mathbf{P}(\mathbb{R}^{d+1})$ a local diffeomorphism. The local inverses of this map give embeddings of "small" subsets of $\mathbf{P}(\mathbb{R}^{d+1})$ in $\mathbb{R}^{d+1}$, and the notions of volume and uniform distribution for submanifolds of Euclidean space are extended to $\mathbf{P}(\mathbb{R}^{d+1})$ by endowing it with the notion of volume derived, in the obvious way, from these embeddings.

2.2. *The Model of a Random Game*

Let $S = S_1 \times \cdots \times S_n$. Endow $\mathbb{R}^S$ with the standard inner product and associated norm. For $i = 1, \ldots, n$ let M_i be a copy of the unit sphere in $\mathbb{R}^S$, and let

$$M := M_1 \times \cdots \times M_n.$$

The model of a random game studied here, which we describe as the *spherical model*, is given by $\mathbf{U}_M$. For any vector of nonzero utilities $u = (u_1, \ldots, u_n)$ the set of Nash equilibria is the same as the set of Nash equilibria for the vector of normalized utilities

$(u_1/\|u_1\|, \ldots, u_n/\|u_n\|) \in M$, so, from the point of view of the induced distribution of Nash equilibria, any distribution on the space of games $(\mathbb{R}^S \setminus \{0\})^n$ is equivalent to the distribution on M induced by this normalization. As we mentioned above, a natural model that is equivalent to the spherical model is that the payoffs of the various agents at the various pure strategy profiles are independently and identically distributed normal random variables.

2.3. *Nash Equilibrium as a System of Polynomial Equations and Inequalities*

For each $i = 1, \ldots, n$ we index the elements of S_i by setting

$$T_i = \{t_i^0, t_i^1, \ldots, t_i^{p_i}\} \quad \text{and} \quad S_i \setminus T_i = \{s_i^1, \ldots, s_i^{q_i}\}.$$

Let $S_{-i} := \prod_{h \neq i} S_h$ and $T_{-i} := \prod_{h \neq i} T_h$. Typical elements of T_{-i} are written $t_{-i} = (t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_n)$.

We wish to express the equations and inequalities that make up the definition of Nash equilibrium in the manner that is standard in the theory of polynomial systems. For this purpose we introduce notation that is uncommon in game theory. Let

$$\mathcal{B}_i := \{\, b \in \{0,1\}^{T_1 \cup \ldots \cup T_n} : \text{for each } h \neq i \text{ there is exactly one } t_h \in T_h \text{ such that}$$
$$b(t_h) = 1, \text{ and all other components are zero} \,\}.$$

The elements of $\mathcal{B}_i$ are thought of as exponent vectors. (The general idea is described at the beginning of Section 4.) For $\tau \in \mathbb{R}^{T_1} \times \cdots \times \mathbb{R}^{T_n}$ and $b \in \mathcal{B}_i$ let

$$\tau^b := \prod_{i=1}^{n} \left(\prod_{j=0}^{p_i} \tau_i^{b(t_i^j)} \right).$$

Let $\tau^{\mathcal{B}_i}$ be the element of $\mathbb{R}^{\mathcal{B}_i}$ whose component for each $b \in \mathcal{B}_i$ is τ^b.

The following notation will be useful in combining the components of $\tau^{\mathcal{B}_i}$ with the components of $u_i \in \mathbb{R}^S$ to form expected utilities, and differences between the expected utilities associated with pairs of pure strategies in S_i. Let $\beta_i : T_{-i} \to \mathcal{B}_i$ be the obvious bijection: for each $h \neq i$ the h-component t_h of t_{-i} is the element of T_h whose component in $\beta_i(t_{-i})$ is one. Let $\iota_i : \mathbb{R}^{T_{-i}} \to \mathbb{R}^{\mathcal{B}_i}$ be the associated linear isomorphism given by this relabelling of coordinates: for each $t_{-i} \in T_{-i}$ the $\beta_i(t_{-i})$-component of $\iota_i(z)$ is $z_{t_{-i}}$.

We think of $u_i \in \mathbb{R}^S$ as an $\#S_i$-tuple of elements of $\mathbb{R}^{S_{-i}}$ indexed by the elements of S_i. In turn, the components associated with t_i^j ($j = 0, 1, \ldots, p_i$) and s_i^k ($k = 1, \ldots, q_i$) are decomposed as $(v_i^{t_i^j}, w_i^{t_i^j})$ and $(v_i^{s_i^k}, w_i^{s_i^k})$ where $v_i^{t_i^j}, v_i^{s_i^k} \in \mathbb{R}^{T_{-i}}$ and $w_i^{t_i^j}, w_i^{s_i^k} \in \mathbb{R}^{S_{-i} \setminus T_{-i}}$. For $j = 1, \ldots, p_i$ define

$$\kappa_i^j : \mathbb{R}^S \to \mathbb{R}^{\mathcal{B}_i} \text{ by } \kappa_i^j(u_i) = \iota_i(v_i^{t_i^j} - v_i^{t_i^0}),$$

and for $k = 1, \ldots, q_i$ define

$$\lambda_i^k : \mathbb{R}^S \to \mathbb{R}^{\mathcal{B}_i} \text{ by } \lambda_i^k(u_i) = \iota_i(v_i^{s_i^k} - v_i^{t_i^0}).$$

In general we will write $\langle \chi_i, \psi_i \rangle$ to denote the inner product of $\chi_i, \psi_i \in I\!R^{\mathcal{B}_i}$.

In this system of notation a point $\tau \in I\!R^{T_1} \times \cdots \times I\!R^{T_n}$ is a *strict*[4] *Nash equilibrium for* $u = (u_1, \ldots, u_n) \in (I\!R^S)^n$ *with support* $(T_1, \ldots, T_n)$ if, for all $i = 1, \ldots, n$:

(N1) $\langle \kappa_i^j(u_i), \tau^{\mathcal{B}_i} \rangle = 0$ for all $j = 1, \ldots, p_i$;
(N2) $\langle \lambda_i^k(u_i), \tau^{\mathcal{B}_i} \rangle < 0$ for all $k = 1, \ldots, q_i$;
(N3) τ_i is in the interior of the positive orthant of $I\!R^{T_i}$;
(N4) the sum of the components of τ_i is one.

Each component of $\tau^{\mathcal{B}_i}$ is multilinear, in the sense of being linear separately in each τ_h. Therefore the truth values of (N1)-(N3) are unaffected by the multiplication of any τ_i by a positive scalar. Consequently the number of solutions of this system is unaffected if (N4) is replaced by another "normalization" that selects a representative point from each ray emanating from the origin in $I\!R^{T_i}$. Let

$$N := N_1 \times \cdots \times N_n \quad \text{and} \quad N^{++} := N_1^{++} \times \cdots \times N_n^{++}$$

where, for each i, N_i is the unit sphere in $I\!R^{T_i}$ and N_i^{++} is the intersection of N_i with the interior of the positive orthant. In very general terms, the advantage of these spaces is that N has a rich group of symmetries.

2.4. *Statement of the Main Result*

The most important analogue of the graph of the equilibrium correspondence will be

$$V = \{\, (u, \tau) \in M \times N : \text{(N1) and (N2)} \,\}.$$

Our strategy is to characterize the distribution of solutions of (N1) and (N2) in N, after which the distribution of solutions of (N1)-(N3) is obtained by restricting attention to N^{++}. Since the distribution of roots of (N1) and (N2) is a scalar multiple of $\mathbf{U}_N$, the expected number of roots of (N1) and (N2) in N^{++} is $2^{-(|T_1| + \cdots + |T_n|)}$ times the expected number of roots in N.

Let $\pi_1 : V \to M$ and $\pi_2 : V \to N$ be the restrictions to V of the natural projections from $M \times N$. Our goal is to characterize the measure ν on N defined, for measurable $E \subset N$, by

$$\nu(E) = \int_M \#\Big(\pi_1^{-1}(u) \cap \pi_2^{-1}(E)\Big)\, d\mathbf{U}_M(u).$$

We now define various objects that enter the statement of the result. For $i = 1, \ldots, n$ let $p_i := |T_i| - 1$ be the dimension of N_i, let $\mathbf{p} := (p_1, \ldots, p_n)$ be the vector of these dimensions, and let

$$p := p_1 + \cdots + p_n.$$

Effectively p is the number of degrees of freedom in the system of polynomial equations we are studying.

---

[4] The qualifier *strict* refers to the strictness of the inequalities in (N2). For generic utilities, all equilibria are strict (Harsanyi (1973)) so there is no difference between the expected number of Nash equilibria and the expected number of strict Nash equilibria.

The expectation of the absolute value of the determinant of a certain random matrix is the most complicated object entering the formula below. Let $D(\mathbf{p})$ be the $p \times n$ matrix with entries

$$D_{ki}(\mathbf{p}) = \begin{cases} 0, & \text{if } p_1 + \cdots + p_{i-1} < k \le p_1 + \cdots + p_i, \\ 1, & \text{otherwise.} \end{cases}$$

For example

$$D(1,2,3) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}.$$

Let $\tilde{Z}_{\mathbf{p}}$ be a $p \times p$ random matrix whose rows are indexed by the integers $k = 1, \ldots, p$, whose columns are indexed by the pairs ij for $i = 1, \ldots, n$ and $j = 1, \ldots, p_i$, and whose entries $\tilde{z}_k^{ij}$ are independently distributed normal random variables with mean zero and variance $D_{ki}(\mathbf{p})$. That is, $\tilde{Z}_{\mathbf{p}}$ is a $p \times p$ matrix of centered normal random variables with variance matrix $D(\mathbf{p})C(\mathbf{p})$ where

$$C(\mathbf{p}) := \begin{bmatrix} 1\ldots1 & 0\ldots0 & \ldots & 0\ldots0 \\ 0\ldots0 & 1\ldots1 & \ldots & 0\ldots0 \\ \vdots & \vdots & & \vdots \\ 0\ldots0 & 0\ldots0 & \ldots & 1\ldots1 \end{bmatrix}$$

is an $n \times p$ column-copying matrix that has exactly p_1 1's in the first row, p_2 1's in the second row, ... , p_n 1's in the n^{th} row. Continuing the example above,

$$D(1,2,3)C(1,2,3) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

For integers $a, b \ge 0$ let $r(a,b)$ be the probability that $\epsilon_0/\sqrt{a+1}$ is greater than each of $\epsilon_1, \ldots \epsilon_b$ when $\epsilon_0, \epsilon_1, \ldots, \epsilon_b$ are i.i.d. normal random variables with mean zero. Lemma 6.14 gives a geometric description of this quantity, showing that it is the fraction of the volume of a certain sphere occupied by a certain cone.

Let $\Gamma : (0,\infty) \to (0,\infty)$ be Euler's function: $\Gamma(s) \equiv \int_0^\infty \exp(-x)x^{s-1}\,dx$.

Main Theorem: *For all measurable $E \subset N$,*

$$\nu(E) = \mathbf{U}_N(E) \cdot \Bigl(\prod_{i=1}^n r(p_i, |S_i \setminus T_i|)\Bigr) \cdot 2^{n-\frac{p}{2}} \cdot \Bigl(\prod_{i=1}^n \frac{\pi/2}{\Gamma(\frac{p_i+1}{2})}\Bigr) \cdot \mathbf{E}\bigl(|\det \tilde{Z}_{\mathbf{p}}|\bigr).$$

To obtain the version of the result that is most natural from the point of view of game theory we restrict attention to roots whose components are all positive, as per (N3). Since ν is a multiple of the uniform distribution, to obtain the total number of Nash equilibria with support $\mathbf{T}$ we divide by 2^{p+n}. Under the natural map the uniform distribution on N_i^{++} induces a distribution on the unit simplex in $\mathbb{R}^{T_i}$.

Corollary: *For each $i = 1, \ldots, n$ let $\Delta^\circ(T_i) := \{\, \tau_i \in \mathbb{R}^{T_i}_{\geq} : \sum_{s_i \in T_i} \tau_i(s_i) = 1 \,\}$ be the open simplex of totally mixed strategy vectors with support T_i, and let*

$$\rho : N^{++} \to \Delta^\circ(T_1) \times \cdots \times \Delta^\circ(T_n) \ \text{ be the map } \ \rho(\tau) := \Big(\frac{\tau_1}{\|\tau_1\|_1}, \ldots, \frac{\tau_n}{\|\tau_n\|_1}\Big)$$

where $\|\tau_i\|_1 := \sum_{s_i \in T_i} |\tau_i(s_i)|$. For each measurable $E \subset \Delta^\circ(T_1) \times \cdots \times \Delta^\circ(T_n)$ the expected number of Nash equilibria in E is

$$\mathbf{U}_{N^{++}}(\rho^{-1}(E)) \cdot 2^{-\frac{3}{2}p} \cdot \Big(\prod_{i=1}^{n} r(p_i, |S_i \setminus T_i|)\Big) \cdot \Big(\prod_{i=1}^{n} \frac{\pi/2}{\Gamma(\frac{p_i+1}{2})}\Big) \cdot \mathbf{E}\big(|\det \tilde{Z}_{\mathbf{p}}|\big).$$

2.5. *Outline of the Argument*

The proof of the Main Theorem has three phases. The first is Proposition 4.1, which is a result from McLennan (2002) that characterizes the mean number of roots of a certain random system of polynomial equations. The Appendix reproduces, with slight adaptation, the proof from McLennan (2002). The next section illustrates the ideas of this argument by executing the calculation in a simple special case.

We refer to the special case of the Main Theorem given by $T_1 = S_1, \ldots, T_n = S_n$ as the *full support case.* In the full support case condition (N2) is vacuous, and the system of equations given by (N1) falls within the class of equations considered by Proposition 4.1. However, the distribution on coefficient vectors considered in Proposition 4.1 is *not* the one obtained by starting with our model of a random game and passing to the system of equations (N1). The second phase of the argument, given in Section 5, completes the proof of the full support case by showing that, nonetheless, the two random equation systems induce the same distribution of roots in N.

The third phase, which passes from the full support case to the general result, begins with the following observation. The spherical model, applied to the sets of pure strategies $T_1, \ldots, T_n$, is equivalent, from the point of view of the induced distribution of roots of (N1), to the distribution obtained from the spherical model, applied to the sets of pure strategies $S_1, \ldots, S_n$, by projection onto the space of payoffs associated with pure strategy vectors in $T_1 \times \cdots \times T_n$. (This is completely obvious from the point of view of the equivalent model of a random game in which the payoffs at the various pure strategy vectors are i.i.d. normal random variables.) Consequently it suffices to show that, for any measurable $E \subset N$, $\prod_i r(p_i, |S_i \setminus T_i|)$ is the ratio of the expected number of roots of (N1) and (N2) in E and the expected number of roots of (N1) in E. This is accomplished in Section 6.

3. An Illustrative Calculation

The main ideas in the proof of Proposition 4.1 are displayed in this section by following the computation of the mean number of real roots of a random univariate polynomial. Since Section 6 follows many of the same steps, this calculation is also illustrative of the argument presented there. As in Edelman and Kostlan (1995), which gives a much more extensive introduction to a broad range of related ideas, we present the ideas in a sequence that conveys some sense of the historical development of the subject.

First of all, consider the problem, first studied by Kac (1943), of determining the mean number of real roots of a quadratic polynomial at^2+bt+c where a, b, c are i.i.d. normal random variables with mean zero. Elementary properties of the normal distribution imply that the normalized coefficient vector $(a, b, c)/\|(a, b, c)\|$ is uniformly distributed in the unit sphere in $\mathbb{R}^3$. It seems that the problem is a matter of determining the area of the set of points in this sphere at which the discriminant $b^2 - 4ac$ is positive.

Economists are accustomed to thinking of "parameters," such as the coefficients a, b, c, as being exogenous and given. This seems to reinforce habits of thought that lead one to attempt to determine an average over a space of parameters by integrating the quantity of interest, here the number of solutions, across the given measure on the parameter space. From a mathematical point of view, however, this is usually an ill-behaved calculation. The set of solutions of an economic model is typically the set of fixed points of a function or correspondence. For polynomials it is relatively hard to pass from a given polynomial to its roots. *In contrast, if, for a particular point in the solution space, we ask what parameters have that point as a solution, the set of such parameters is often very well behaved.* For general equilibrium theory this point of view is emphasized in Balasko (1988) where the equilibrium manifold of an exchange economy is displayed as a vector bundle[5] in which the set of endowments that have a particular equilibrium allocation is the fiber.

For the quadratic polynomial the idea of integrating over the space of solutions leads us to define $\gamma : \mathbb{R} \to S^2 \subset \mathbb{R}^3$ by

$$\gamma(t) = \frac{(t^2, t, 1)}{\|(t^2, t, 1)\|}.$$

Observing that t is a root of the polynomial $aT^2 + bT + c$ if and only if $(a, b, c) \perp \gamma(t)$, we see that the probability that the polynomial will have a root in the interval $(t, t + \Delta t)$ is, for small Δt, approximately equal to the fraction of the sphere's area that lies between the

---

(5) A vector bundle is a type of fiber bundle. There are several closely related types of fiber bundles, the simplest of which is the topological notion: we say that a map $\pi : E \to B$ between topological spaces is a *fibration*, and that E is a *fiber bundle* with *base space* B, if there is a topological space F, called the *fiber*, such that for each $b \in B$ there is a neighborhood $U_b \subset B$ and a homeomorphism $\phi_b : U_b \times F \to \pi^{-1}(U_b)$ such that $\pi \circ \phi_b$ coincides with the standard projection $U_b \times F \to U_b$. This fibration is C^∞ if E, B, and F are C^∞ manifolds, π is a C^∞ map for which all points of E are regular, and each ϕ_b is a C^∞ diffeomorphism. A *vector bundle* is a fiber bundle in which F is a finite dimensional vector space, and, for all $b, b', \beta \in B$ such that $\beta \in U_b \cap U'_b$, the homeomorphism from F to itself that takes f to the second component of $\phi_{b'}^{-1}(\phi_b(\beta, f))$ is linear. If F has an inner product and this map is always an orthogonal transformation, there is an associated *sphere bundle* $\pi' : E' \to B$ constructed by letting F' be the unit sphere in F, letting $E' = \cup_{b\in B}\, \phi_b(U_b \times F')$, and letting π' be the restriction of π to E'.

two great circles consisting of the points orthogonal to $\gamma(t)$ and $\gamma(t+\Delta t)$ respectively, and this fraction is approximately $1/\pi$ times the distance from $\gamma(t)$ to $\gamma(t+\Delta t)$. (To see that $1/\pi$ is the appropriate factor consider that, for a given great circle, with two exceptions each point in the sphere is orthogonal to exactly two points on the great circle.) For small Δt the distance from $\gamma(t)$ to $\gamma(t+\Delta t)$ is approximately $\|\gamma'(t)\| \cdot \Delta t$. Therefore the mean number of real roots is the length of the curve γ divided by π. While this may not seem like an obvious increase in tractability, in fact the length of γ, and a host of related issues, have been studied extensively. (Cf., Edelman and Kostlan (1995).)

A different response to the complexities of the length of γ is to avoid them by finding a modified problem that is better behaved. In particular, we may ask whether a different distribution of coefficient vectors might be more tractable. In algebraic geometry it is often simplifying to replace an inhomogeneous polynomial such as $ax^2 + bx + c$ with the homogenized polynomial $ax^2 + bxy + cy^2$ obtained by multiplying each monomial by a new variable, here y, raised to a power chosen to give rise to a polynomial that is homogeneous: all monomials have the same total degree. If (x, y) is mapped to 0 by this polynomial function, then so is $(\alpha x, \alpha y)$ for any scalar α, so we may think of the "roots" of the polynomial as being one dimensional linear subspaces of $I\!R^2$, i.e., the elements of the projective space $\mathbf{P}(I\!R^2)$.

The distribution of coefficient vectors we have been assuming, that a, b, and c are i.i.d. normal random variables, is "natural" in the psychological sense of being the first to spring to mind, but there is a definite mathematical sense in which a different distribution of coefficient vectors is simpler. Consider, for $0 \le \theta < 2\pi$, the transformation of variables corresponding to a coordinate system that is rotated θ radians:

$$(x, y) = (w\cos\theta + z\sin\theta, -w\sin\theta + z\cos\theta).$$

Substituting in the quadratic, we obtain a quadratic polynomial in the variables (w, z):

$$\big(a\cos^2\theta - b\cos\theta\sin\theta + c\sin^2\theta\big)w^2 + \big(2(a-c)\cos\theta\sin\theta + b(\cos^2\theta - \sin^2\theta)\big)wz$$
$$+\big(a\sin^2\theta + b\cos\theta\sin\theta + c\cos^2\theta\big)z^2.$$

We will consider distributions on the space of coefficient vectors that are invariant under these transformations, meaning that, for any θ, the distribution of transformed coefficient vectors agrees with the given distribution.

We now digress, introducing relevant terminology and concepts from group theory. A group H with identity element e is said to *act* (from the left) on a space X if there is a rule associating an element $hx \in X$ to each pair $(h, x) \in H \times X$ that satisfies:

(a) $ex = x$ for all $x \in X$;
(b) $h(h'x) = (hh')x$ for all $h, h' \in H$ and $x \in X$.

A function f with domain X is *invariant* under the action if $f(hx) = f(x)$ for all $x \in X$ and $h \in H$. Various other usages of the term 'invariant' can be understood in terms of this definition being satisfied in relation to derived actions of H. For example, one says that a set is invariant, or an invariant, when it is really the set's characteristic function that satisfies this definition. When X is a linear space we say that an inner product is invariant under the action when, formally, we mean that the inner product is invariant under the derived action $h(x,x') = (hx, hx')$ on $X \times X$. When we say that a measure on X is invariant, what we mean formally is that the measure (which is a real valued function on a σ-algebra) is invariant under the derived action of H on the σ-algebra.

The action on X is said to be *transitive*[6] if, for any two elements $x, x' \in X$, there is some $h \in H$ with $hx = x'$. The only nontrivial result from group theory that we will use is the following famous theorem.

Proposition 3.1: (E.g., Royden (1988, ch. 14)) *If X is compact, the action on X is continuous in the sense that, for each $h \in H$, the map $x \mapsto hx$ is continuous, and H acts transitively, then (up to multiplication by a scalar) there is exactly one invariant measure (called* Haar measure*) on X.*

Most of the groups that appear in our work will be instances or products of the *orthogonal group* $O(\mathbb{R}^{m+1})$, which is the group of linear automorphisms of $\mathbb{R}^{m+1}$ that preserve the inner product. Of course the unit sphere S^m centered at the origin is invariant under the obvious action of $O(\mathbb{R}^{m+1})$ on $\mathbb{R}^{m+1}$, so there is an induced action on S^m which is easily shown to be transitive. The uniform distribution on S^m is the unique invariant measure of this action. Similarly, there is an induced action of $O(\mathbb{R}^{m+1})$ on m-dimensional projective space, and the unique invariant measure is the uniform distribution.

Our goal in the remainder of this section will be to compute the expected number of roots, in one dimensional projective space, of the polynomial

$$a_0x^n + a_1x^{n-1}y + \cdots + a_ny^n$$

with respect to a particular distribution of coefficient vectors specified below. The group $O(\mathbb{R}^2)$ acts on the space of homogeneous polynomials P of degree n in the variables (x, y) by the rule $OP = P \circ O^{-1}$. We would like to find a measure on the space of such polynomials that is invariant under this action, in the hope that the implied distribution of real roots might be well behaved. Automatically such a distribution must have the pleasant property that the implied distribution of roots in one dimensional projective space will be invariant under rotations of the plane, hence a multiple of the uniform distribution.

If we restrict attention to distributions in which the coefficients $a_0, \dots, a_n$ are independently and normally distributed, our problem is to solve for a system of variances with the desired property. Alternatively, we may look for an inner product on the space of coefficient vectors with respect to which the monomials $x^n, x^{n-1}y, \dots, y^n$ are orthogonal and the central multinormal distribution is invariant. The latter condition will hold if the transformation of polynomials (which is always linear) is an orthogonal transformation. It turns out that the inner product

$$\left\langle \sum_{i=0}^{n} a_i x^{n-i} y^i, \sum_{i=0}^{n} b_i x^{n-i} y^i \right\rangle = \sum_{i=0}^{n} \binom{n}{i}^{-1} a_i b_i \tag{1}$$

[6] This concept is unrelated to the usages of the term 'transitivity,' say for preferences, that are most common in economics.

is invariant[7], so that a desirable distribution on the space of coefficient vectors is obtained if each a_i is distributed normally with mean 0 and variance $\binom{n}{i}$. For the remainder of this section let $\mathcal{H}$ be the space of coefficient vectors $f = (a_0, \ldots, a_n)$ endowed with this inner product, let $\mathcal{M}$ denote the unit sphere relative to this inner product, and let N denote the unit sphere S^1 in $\mathbb{R}^2$.

Let $F : \mathcal{H} \times \mathbb{R}^2 \to \mathbb{R}$ be the *evaluation map*:

$$F(f, (x, y)) := f(x, y) = a_0 x^n + a_1 x^{n-1} y + \cdots + a_n y^n.$$

From the point of view of the distribution of roots, the central multinormal distribution on the space of coefficient vectors and the uniform distribution on the unit sphere $\mathcal{M}$ in this space are equivalent, and the latter is more convenient in certain respects. The *incidence variety* is

$$V := F^{-1}(0) \cap (\mathcal{M} \times N) = \{ (f, (x, y)) \in \mathcal{M} \times N : f(x, y) = 0 \}.$$

Note that F is invariant under the action

$$O(f, (x, y)) := (f \circ O^{-1}, O(x, y))$$

of $O(\mathbb{R}^2)$ on $\mathcal{H} \times \mathbb{R}^2$, so V is invariant under this action, and there is an action on V defined by restriction.

We would like to show that each point $(f, (x, y)) \in V$ is a regular point of the restriction of F to $\mathcal{M} \times N$, since then the regular value theorem (e.g., Guillemin and Pollack (1965)) would imply that V is an n-dimensional C^∞ manifold. Since $(x, y) \neq (0, 0)$, not all of the monomials $x^{n-i} y^i$ vanish, so f is a regular point of $F(\cdot, (x, y)) : \mathcal{H} \to \mathbb{R}$. Interpreting f as an element of $T_f\mathcal{H}$ and abusing notation by identifying f and $(f, 0) \in T_{(f,(x,y))}(\mathcal{H} \times \mathbb{R}^2)$, we have $DF(f, (x, y))f = 0$. (To see this note that (x, y) is a root of αf for all $\alpha \in \mathbb{R}$.) Since f is a regular point of $F(\cdot, (x, y))$, there must be a vector $v \in T_f\mathcal{M}$ such that $DF(f, (x, y))v \neq 0$. It follows that f is a regular point of the restriction of $F(\cdot, (x, y))$ to $\mathcal{M}$, which implies that $(f, (x, y))$ is a regular point of $F|_{\mathcal{M} \times N}$, as desired.

Let

$$\pi_1 : V \to \mathcal{M} \quad \text{and} \quad \pi_2 : V \to N$$

be the natural projections. Sard's theorem implies that the critical values of π_1 constitute a set of measure zero in $\mathcal{M}$, so they can be ignored in computing the expected number of real roots of the equation. A measure μ on V may be defined by requiring that if $U \subset V$ is an open set containing only regular points of π_1 and the restriction of π_1 to U is injective, then $\mu(U) = \mathbf{U}_{\mathcal{M}}(\pi_1(U))$. For any open $Z \subset V$ the expected number of real roots corresponding to points $(x, y) \in Z$ is

$$\int_{\mathcal{M}} \#(\pi_1^{-1}(x) \cap Z) \, d\mathbf{U}_{\mathcal{M}}(x) = \mu(Z).$$

---

(7) This invariance is not particularly easy to prove. Perhaps the most accessible and self contained account is Edelman and Kostlan (1995, pp. 15–17). There are other inner products that are invariant, but this is the only invariant inner product in which the monomials are orthogonal. These matters are discussed in detail in Kostlan (2000).

In turn there is a measure ν on N defined by requiring that, for each measurable $E \subset N$, $\nu(E) = \mu(\pi_2^{-1}(E))$, i.e., $\nu = \mu \circ \pi_2^{-1}$. If $Z = \pi_2^{-1}(\pi_2(Z))$, then $\mu(Z) = \nu(\pi_2(Z))$. In this sense ν is the distribution of roots.

It turns out that π_2 is the projection of a C^∞ sphere bundle. The *fiber* above $(x, y) \in N$ is

$$V_{(x,y)} := \{ f \in \mathcal{M} : (f, (x, y)) \in V \}.$$

Consider a particular $(x_0, y_0) \in N$. As the set of coefficient vectors in $\mathcal{M}$ that are orthogonal to $(x_0^n, x_0^{n-1}y_0, \ldots, y_0^n)$, $V_{(x_0,y_0)}$ is an $(n-1)$-dimensional sphere in $\mathcal{M}$. Varying (x, y) in a neighborhood of (x_0, y_0) can be thought of as inducing a motion of the sphere $V_{(x,y)}$ in $\mathcal{M}$, and, roughly speaking, the probability of having a root in a small neighborhood of (x_0, y_0) will be proportional to the speed at which $V_{(x,y)}$ moves as we vary (x, y) near (x_0, y_0).

This intuition is made rigorous, at the natural level of generality, by an integral formula of Shub and Smale (1993, p. 273). (Cf. Blum *et. al.* (1998, p. 240).) We now describe the consequence of this formula in the current context. For $(f, (x, y)) \in V$ let

$$A(f, (x, y)) : T_{(x,y)}N \to T_f\mathcal{M}$$

be the linear map whose graph is the orthogonal complement $\perp_{(f,(x,y))}$ of $T_{(f,(x,y))}V_{(x,y)}$ in $T_{(f,(x,y))}V$, and let $A^*(f, (x, y))$ be the adjoint[(8)] of this map. In this setting the Shub-Smale integral formula states that for any open set $Z \subset V$,

$$\int_{\mathcal{M}} \#(\pi_1^{-1}(f) \cap Z)\, df = \int_N \int_{V_{(x,y)} \cap Z} \det\left(A^*(f, (x, y))A(f, (x, y))\right)^{1/2} df\, d(x, y). \qquad (2)$$

The idea expressed in this formula is geometric, so that the assumed measures on $\mathcal{M}$, N, and $V_{(x,y)}$ are the natural notions of n-dimensional volume and length (which we have been denoting by $\text{vol}(\cdot)$) that are derived from the inner products of the ambient spaces. In particular, the expected number of roots is the integral of $\#(\pi_1^{-1}(f))$ with respect to the uniform distribution on $\mathcal{M}$, so

$$\int_{\mathcal{M}} \#(\pi_1^{-1}(f) \cap Z)\, d\mathbf{U}_{\mathcal{M}}(f) = \frac{1}{\text{vol}(\mathcal{M})} \int_{\mathcal{M}} \#(\pi_1^{-1}(f) \cap Z)\, df. \qquad (3)$$

When $Z = \pi_2^{-1}(Y)$ for some open $Y \subset N$, so that

$$\int_N \Big(\int_{V_{(x,y)} \cap Z} \cdots df \Big)\, d(x, y) = \int_Y \Big(\int_{V_{(x,y)}} \cdots df \Big)\, d(x, y),$$

the right hand side of the formula above can be further simplified by exploiting the invariances arising out of the action

$$O(f, (x, y)) = (f \circ O^{-1}, O(x, y))$$

---

(8) Recall that if V and W are inner product spaces, and $L : V \to W$ is linear, the *adjoint* of L is the linear transformation $L^* : W \to V$ defined implicitly by the requirement that $\langle v, L^*w \rangle = \langle Lv, w \rangle$ for all $v \in V$ and $w \in W$.

of $O(\mathbb{R}^2)$ on $\mathcal{M} \times N$. Without going into any detail at this point (examples of this sort of argument occurs in Section 6 and the Appendix) we simple assert that

$$\int_{V_{(x,y)}} \det \left(A^*(f,(x,y))A(f,(x,y))\right)^{1/2} df$$

is a constant function of $(x,y) \in N$. So, since the action of $O(\mathbb{R}^2)$ on N is transitive, for any open $Y \subset N$ and *any* $(x_0, y_0) \in N$ (including those not in Y)

$$\begin{aligned} &\int_Y \int_{V_{(x,y)}} \det \left(A^*(f,(x,y))A(f,(x,y))\right)^{1/2} df\, d(x,y) \\ &\qquad = \mathrm{vol}(Y) \cdot \int_{V_{(x_0,y_0)}} \det \left(A^*(f,(x_0,y_0))A(f,(x_0,y_0))\right)^{1/2} df. \end{aligned} \tag{4}$$

In evaluating the integral on the right hand side we are now free to choose (x_0, y_0) as we please, and it turns out to be simplest to work with $(1,0)$. Note that for $f = (a_0, a_1, \ldots, a_n) \in \mathcal{M}$ the condition $f \in V_{(1,0)}$ amounts to $a_0 = 0$.

Lemma 3.2: *For all* $f = (0, a_1, \ldots, a_n) \in V_{(1,0)}$,

$$\det \left(A^*(f,(1,0))A(f,(1,0))\right)^{1/2} = |a_1|.$$

Proof: We have

$$\begin{aligned} T_f\mathcal{M} &:= \{\, \phi = (\phi_0, \phi_1, \ldots, \phi_n) \in \mathcal{H} : \langle f, \phi \rangle = 0 \,\} \\ &= \{\, \phi \in \mathcal{H} : \binom{n}{1}^{-1} a_1\phi_1 + \cdots + \binom{n}{n}^{-1} a_n\phi_n = 0 \,\} \end{aligned}$$

and

$$T_{(1,0)}N = \{\, (\xi, \psi) \in \mathbb{R}^2 : (1,0) \cdot (\xi, \psi) = 0 \,\} = \{\, (0, \psi) : \psi \in \mathbb{R} \,\}.$$

Observe that

$$\left. \frac{\partial(a_1x^{n-1}y + \cdots + a_ny^n)}{\partial y} \right|_{(x,y)=(1,0)} = a_1$$

and consequently

$$DF(f,(1,0))(\phi,(0,\psi)) = Df(1,0)(0,\psi) + \phi(1,0) = a_1\psi + \phi_0.$$

(Here $\phi(1,0)$ is the polynomial ϕ evaluated at $(1,0)$.) Therefore

$$T_{(f,(1,0))}V = \{\, (\phi,(0,\psi)) \in T_f\mathcal{M} \times T_{(1,0)}N : a_1\psi + \phi_0 = 0 \,\}.$$

A vector $(\phi, (0, \psi)) \in T_{(f,(1,0))}V$ is in $T_{(f,(1,0))}V_{(1,0)}$ if and only if $\psi = 0$, in which case $\phi_0 = 0$. That is,

$$T_{(f,(1,0))}V_{(1,0)} = \{ ((0, \phi_1, \ldots, \phi_n), (0,0)) \in T_f\mathcal{M} \times T_{(1,0)}N \}$$

A vector $(\phi, (0, \psi)) \in T_{(f,(1,0))}V$ that is orthogonal to $T_{(f,(1,0))}V_{(1,0)}$ must satisfy

$$0 = \langle \phi, \phi' \rangle = \binom{n}{1}^{-1} \phi_1 \phi_1' + \cdots + \binom{n}{n}^{-1} \phi_n \phi_n'$$

for all $\phi' \in T_f\mathcal{M}$ such that $\phi_0' = 0$, and $0 = \binom{n}{1}^{-1} a_1 \phi_1 + \cdots + \binom{n}{n}^{-1} a_n \phi_n$ since $\phi \in T_f\mathcal{M}$, so $\phi_1 = \cdots = \phi_n = 0$. This means that

$$\perp_{(f,(1,0))} = \{ ((\phi_0, 0, \ldots, 0), (0, \psi)) : a_1 \psi + \phi_0 = 0 \},$$

and $A(f, (1,0))$ is the linear map taking $(0,1) \in T_{(1,0)}N$ to $(-a_1, 0, \ldots, 0) \in T_f\mathcal{M}$. The adjoint $A^*(f, (1,0))$ is the map taking $(\phi_0, \phi_1, \ldots, \phi_n) \in T_f\mathcal{M}$ to $(0, -a_1\phi_0) \in T_{(1,0)}N$. We conclude that $A^*(f,(1,0))A(f,(1,0))$ is the map taking $(0,1) \in T_{(1,0)}N$ to $(0, a_1^2) \in T_{(1,0)}N$, and its determinant is a_1^2, as desired. ∎

We now endow $V_{(1,0)}$ with the geometrically natural coordinate system. Recalling that the inner product (1) is not the usual one, let

$$z_0 = \frac{a_0}{\sqrt{\binom{n}{0}}},\ z_1 = \frac{a_1}{\sqrt{\binom{n}{1}}},\ \ldots,\ z_n = \frac{a_n}{\sqrt{\binom{n}{n}}}$$

be the system of coordinates for $\mathcal{H}$ in which the i^{th} standard unit basis vector is the vector of unit length, relative to (1), that is a positive multiple of the coefficient of the monomial $x^{n-i}y^i$. Then $V_{(1,0)}$ is the $(n-1)$-dimensional unit sphere in the coordinate subspace given by $z_0 = 0$, so

$$\int_{V_{(1,0)}} |a_1|\, dz = \sqrt{n} \int_{V_{(1,0)}} |z_1|\, dz. \tag{5}$$

Summarizing the work to this point, (2), (3), (4), and (5) combine to imply that, for any open $Y \subset N$:

$$\int_{\mathcal{M}} \#(\pi_1^{-1}(f) \cap \pi_2^{-1}(Y))\, d\mathbf{U}_{\mathcal{M}}(f) = \frac{\text{vol}(Y)}{\text{vol}(\mathcal{M})} \cdot \sqrt{n} \int_{V_{(1,0)}} |z_1|\, df. \tag{6}$$

The calculation may be completed in numerous ways. We choose one that illustrates in simplified form the ideas underlying Lemmas 6.11 and 6.12 and the notation employed there. Let $D = (-1, 1)$ be the open unit disk in $\mathbb{R}^1$, let $E = S^{n-2}$ and $F = S^{n-1}$ be the $(n-2)$-dimensional and $(n-1)$-dimensional unit spheres. The fiber $V_{(1,0)}$ corresponds to F, and we will arrive at a more tractable version of the right hand side of (6) by means of the change of variables

$$\gamma : D \times E \to F \quad \text{defined by} \quad \gamma(p, r) = (p, (1 - p^2)^{1/2} r).$$

The determinant of the Jacobean of this function is computed as follows. The partial derivative of γ with respect to p is $(1, -p(1-p^2)^{-1/2}r)$ and the norm of this vector is $(1-p^2)^{-1/2}$. Evaluating $D\gamma(p,r)$ at the various elements of an orthonormal basis of $T_{(p,r)}(D \times E) = T_pD \times T_rE$ whose first element is in T_pD, one finds that the image is a pairwise orthogonal basis of $T_{\gamma(p,r)}F$ whose elements other than the first all have length $(1-p^2)^{-1/2}$. Therefore

$$|\det D\gamma(p,r)| = (1-p^2)^{-1/2}((1-p^2)^{1/2})^{n-2} = (1-p^2)^{\frac{n-3}{2}}.$$

Noting that the indefinite integral of $p(1-p^2)^{\frac{n-3}{2}}$ is $-(1-p^2)^{\frac{n-1}{2}}/(n-1)$, we combine these calculations in the change of variables formula, computing that

$$\begin{aligned}\int_F |z_1|\, df &= \int_{D\times E} |p| \cdot |\det D\gamma(p,r)|\, d(p,r)\\ &= \mathrm{vol}(E) \cdot \int_D |p| \cdot (1-p^2)^{\frac{n-3}{2}}\, dp\\ &= 2\,\mathrm{vol}(E) \int_0^1 p(1-p^2)^{\frac{n-3}{2}}\, dp\\ &= \frac{2\,\mathrm{vol}(E)}{n-1}.\end{aligned}$$

Substituting this in (6) and applying the formula

$$\mathrm{vol}(S^{m-1}) = 2\frac{\pi^{m/2}}{\Gamma(\frac{m}{2})} \quad (m \geq 1) \tag{7}$$

for the volume of the $(m-1)$-dimensional unit sphere (e.g., Federer (1969) p. 251) we may finally conclude that, for any open $Y \subset N$,

$$\int_{\mathcal{M}} \#(\pi_1^{-1}(f) \cap \pi_2^{-1}(Y))\, d\mathbf{U}_{\mathcal{M}}(x) = \mathbf{U}_N(Y) \cdot \frac{\mathrm{vol}(N)\mathrm{vol}(E)}{\mathrm{vol}(\mathcal{M})} \cdot \frac{2\sqrt{n}}{(n-1)}$$

$$= \mathbf{U}_N(Y) \cdot \frac{2\pi \cdot \Gamma(\frac{n+1}{2}) \cdot 2\pi^{\frac{n-1}{2}}}{2\pi^{\frac{n+1}{2}} \cdot \Gamma(\frac{n-1}{2})} \cdot \frac{2\sqrt{n}}{(n-1)} = \mathbf{U}_N(Y) \cdot 2\sqrt{n},$$

where the last equality derives from the formula $\Gamma(s+1) = s\Gamma(s)$. Since there are two roots in the circle corresponding to each point in one dimensional projective space, the mean number of roots in one dimensional projective space is $\sqrt{n}$.

4. A Random Multihomogeneous System

The calculation of the last section has been extended in a sequence of papers (Kostlan (1993), Shub and Smale (1993), Rojas (1996), McLennan (2002)) to increasingly general systems of multivariate polynomial equations. The last of these results will be applied in the proof of the full support case. The statement of the formula developed in McLennan (2002) is the main objective of this section.

The most general class of systems of polynomial equations considered here are the so-called sparse systems. We will work with polynomials in the variables $\mathbf{x} = (x_1, \ldots, x_r)$. For an *exponent vector* $a \in \mathbf{N}^r$ let $\mathbf{x}^a$ denote the monomial $x_1^{a_1} \cdots x_r^{a_r}$. A *sparse system of polynomials* is given by specifying *supports* $\mathcal{A}_1, \ldots, \mathcal{A}_p \subset \mathbf{N}^r$, where each $\mathcal{A}_k$ is nonempty and finite, the interpretation being that we are studying the polynomial equation systems of the form $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_p(\mathbf{x})) = 0$ where, for $k = 1, \ldots, p$,

$$f_k(\mathbf{x}) = \sum_{a \in \mathcal{A}_k} f_{ka} \mathbf{x}^a$$

is a polynomial that is a weighted sum of the monomials corresponding to the exponent vectors in $\mathcal{A}_k$. The system is said to be *unmixed* if $\mathcal{A}_1 = \cdots = \mathcal{A}_p$; otherwise it is said to be *mixed*. For $k = 1, \ldots, p$ let $\mathcal{H}_k := \mathbb{R}^{\mathcal{A}_k}$ be the space of real coefficient vectors $f_k = (f_{ka})_{a \in \mathcal{A}_k}$, and let $\mathcal{H} = \mathcal{H}_1 \times \cdots \times \mathcal{H}_p$.

The theory of sparse systems is of relatively recent origin. Classical theory was dominated by the view that any particular system of polynomial equations could be homogenized (using the generalization of the procedure employed in the last section) and then viewed as an instance of the general homogeneous system (this will be defined below) so that results concerning this system were universally applicable. Bernshtein's (1975) theorem (stated below) gave an important impetus to the development of the theory of sparse systems because it displayed a property that was generic in the space of systems for the given supports that differed from the corresponding generic property of the general homogeneous system of which the sparse system was a specialization. In addition, increasing attention to computational efficiency has led to interest in the development of algorithms (e.g., Emiris and Canny (1995), Huber and Sturmfels (1995), Verschelde, Verlinden, and Cools (1994)) for solving systems of polynomial equations that take advantage of the sparseness that is typically present in applications. Algorithms of this sort are being incorporated into the *Gambit*[9] package of software for analysis of noncooperative games.

The sparse system is *multihomogeneous* if the following description is satisfied: the variables in $\mathbf{x}$ are grouped in n blocks $(\mathbf{y}_1, \ldots, \mathbf{y}_n)$, where $\mathbf{y}_i = (y_{i0}, y_{i1}, \ldots, y_{ip_i})$, and, for a given $p \times n$ matrix D of nonnegative integers, each $\mathcal{A}_k$ represents the set of polynomials that are homogeneous of degree d_{ki} as a function of $\mathbf{y}_i$, $i = 1, \ldots, n$. In combinatoric terms this means that

$$\mathcal{A}_k = \mathcal{A}_{k1} \times \cdots \times \mathcal{A}_{kn}$$

where $\mathcal{A}_{ki}$ is the simplex (in the integer lattice) of coefficient vectors in $\mathbf{N}^{p_i+1}$ whose components sum to d_{ki}.

If $(\mathbf{y}_1, \ldots, \mathbf{y}_n)$ is a root of the multihomogeneous system $f = 0$, then, for any scalars $\phi_1, \ldots, \phi_n$, the point $(\phi_1 \mathbf{y}_1, \ldots, \phi_n \mathbf{y}_n)$ is also a solution, so it is natural to look for solutions in the space $S^{p_1} \times \cdots \times S^{p_n}$, where S^{p_i} is the unit sphere in $\mathbb{R}^{p_i+1}$. In this section we denote this space by $N := N_1 \times \cdots \times N_n$. (It is even more natural to count solutions in

[9] `http://www.hss.caltech.edu/~gambit/Gambit.html`.

the cartesian product of projective spaces $\mathbf{P}(\mathbb{R}^{p_1+1}) \times \cdots \times \mathbf{P}(\mathbb{R}^{p_n+1})$, but this would be less convenient in the analysis.)

The multihomogeneous system is said to be *exactly determined* if there are, effectively, as many equations as degrees of freedom:

$$p = p_1 + \cdots + p_n = r - n.$$

Henceforth we consider only this case. Note that, with this restriction, the system is determined by the vector of dimensions $\mathbf{p} := (p_1, \ldots, p_n)$ and the matrix D. When $n = 1$ we say that the system is an instance of *the general homogeneous system.*

The product group

$$G := O(\mathbb{R}^{p_1+1}) \times \cdots \times O(\mathbb{R}^{p_n+1})$$

acts on $\mathbb{R}^{p_1+1} \times \cdots \times \mathbb{R}^{p_n+1}$ in the obvious way: $g\mathbf{x} := (g_1\mathbf{y}_1, \ldots, g_n\mathbf{y}_n)$. Clearly N is an invariant of this action. Since the composition of a linear function with a homogeneous function is homogeneous, of the same degree, for each $k = 1, \ldots, p$ there is also an action of G on each $\mathcal{H}_k$ defined by $gf = f \circ g^{-1}$.

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

$$\langle f_k, f'_k \rangle_k := \sum_{a \in \mathcal{A}_k} \eta(a) f_{ka} f'_{ka}$$

where

$$\eta(a) := \frac{a_{10}! \cdot \ldots \cdot a_{1p_1}!}{(a_{10} + \cdots + a_{1p_1})!} \cdot \ldots \cdot \frac{a_{n0}! \cdot \ldots \cdot a_{np_n}!}{(a_{n0} + \cdots + a_{np_n})!}. \tag{8}$$

McLennan (2002) shows that this is the unique (up to multiplication by a scalar) inner product on $\mathcal{H}_k$ that is invariant under the action of G and which has distinct monomials orthogonal. (The extension from the two dimensional case mentioned in the last section to this level of generality is easy: by virtue of that result, $\langle \cdot, \cdot \rangle_k$ is the unique inner product that is invariant under the action of those elements of G that rotate one 2-dimensional coordinate subspace while leaving all other variables fixed. But such elements generate G, in the sense that the smallest subgroup of G containing all finite products of such elements is G itself.)

Let $\mathcal{M} := \mathcal{M}_1 \times \cdots \times \mathcal{M}_p$ where each $\mathcal{M}_k$ is the unit sphere, with respect to this inner product, in $\mathcal{H}_k$. We study the model of a random multihomogeneous system given by the uniform distribution $\mathbf{U}_\mathcal{M}$. The distributution of roots of the random system is given by the measure ν on N defined, for measurable $E \subset N$, by

$$\nu(E) = \int_\mathcal{M} \#(\{\,\zeta \in E : f(\zeta) = 0\,\})\, d\mathbf{U}_\mathcal{M}(f).$$

Let $\tilde{Z}_{(\mathbf{p},D)}$ be a random $p \times p$ matrix whose entries $\tilde{z}_k^{ij}$ are independently distributed normal random variables with mean zero and variance d_{ki}. That is, $\tilde{Z}_{(\mathbf{p},D)}$ is a $p \times p$ matrix of centered normal random variables with variance matrix $DC(\mathbf{p})$ where $C(\mathbf{p})$ is the $n \times p$ column copying matrix from Section 2.

Proposition 4.1: (McLennan (2002))

$$\nu(N) = 2^{n-p/2} \cdot \Big(\prod_{i=1}^{n} \frac{\pi/2}{\Gamma(\frac{p_i+1}{2})}\Big) \cdot \mathbf{E}\big(|\det \tilde{Z}_{(\mathbf{p},D)}|\big),$$

and $\nu(E) = \mathbf{U}_N(E) \cdot \nu(N)$ *for all measurable* $E \subset N$.

Proof: Appendix A. ∎

For each i here are 2 points in N_i for each point in $\mathbf{P}(\mathbb{R}^{p_i+1})$, so the mean number of roots in $\mathbf{P}(\mathbb{R}^{p_1+1}) \times \cdots \times \mathbf{P}(\mathbb{R}^{p_n+1})$ is obtained by dividing by 2^n. For the unmixed general homogeneous system consisting of p homogeneous equations of degree d Kostlan (1993) showed that the mean number of real roots in p-dimensional projective space is $\sqrt{d^p}$. Rojas (1996) gives a formula that generalizes this to the case of an unmixed multihomogeneous system. For the mixed general homogeneous system consisting of homogeneous equations of degrees $d_1, \ldots, d_p$ Shub and Smale (1993) show that the mean number of real roots in $\mathbf{P}(\mathbb{R}^{p+1})$ is $\sqrt{\prod_{k=1}^{p} d_k}$. As is explained in McLennan (2002), all these results can be derived without difficulty from the formula above.

The classical result known as Bezout's theorem asserts that, for generic systems of complex coefficients, the general homogeneous system has $\prod_{k=1}^{p} d_k$ roots in p-dimensional complex projective space. Bernshtein's (1975) theorem is a generalization for an exactly determined sparse system in p variables characterized by supports $\mathcal{A}_1, \ldots, \mathcal{A}_p \subset \mathbf{N}^p$. For $k = 1, \ldots, p$ the *Newton polytope* of equation k, denoted by Q_k, is the convex hull of $\mathcal{A}_k$. The *mixed volume* of the tuple $(Q_1, \ldots, Q_p)$ is the coefficient of $\lambda_1 \cdots \lambda_p$ in the polynomial[10] $\mathrm{vol}(\lambda_1 Q_1 + \cdots + \lambda_p Q_p)$. Let $\mathbf{C}^* = \mathbf{C} \setminus \{0\}$. Bernshtein's theorem asserts that, for generic systems of complex coefficients, the number of roots of the system in $(\mathbf{C}^*)^p$ is equal to the mixed volume of $(Q_1, \ldots, Q_p)$. McLennan (2002) uses Proposition 4.1 to generalize the relation, given by the Shub-Smale theorem, between the mean number of roots of the general homogeneous system and the Bezout number: the mean number of roots of the multihomogeneous system in $\mathbf{P}(\mathbb{R}^{p_1+1}) \times \cdots \times \mathbf{P}(\mathbb{R}^{p_n+1})$ is greater than or equal to the square root of the number given by Bernshtein's theorem for the exactly determined system obtained by "demultihomogenizing" by setting $y_{10} = \cdots = y_{n0} = 1$.

McKelvey and McLennan (1997) show that for a given normal form there are payoffs for which there are as many totally mixed (i.e., full support) regular Nash equilibria as are permitted by Bernshtein's theorem. For each root in $\mathbf{P}(\mathbb{R}^{p_1+1}) \times \cdots \times \mathbf{P}(\mathbb{R}^{p_n+1})$ there are 2^n roots in N. Since the roots in N are uniformly distributed, the fraction lying in N^{++} is $\prod_i 2^{-(p_i+1)} = 2^{-(p+n)}$. In the full support case the Main Theorem asserts that the mean number of roots of (N1) is the same as the mean number of roots in N given

(10) See Ewald (1996) for a proof that $\mathrm{vol}(\lambda_1 Q_1 + \cdots + \lambda_p Q_p)$ is, in fact, a polynomial function of $\lambda_1, \ldots, \lambda_p \geq 0$.

by Proposition 4.1 for the corresponding multihomogeneous system. Putting these facts together leads to the conclusion that *the mean number of totally mixed Nash equilibria is greater than or equal to* 2^{-p} *times the square root of the maximal number of regular totally mixed Nash equilibria.*

This lower bound may be used to establish exponential rates of growth of the mean number of totally mixed equilibria as the size of the game grows in various ways. McKelvey and McLennan (1997) use a recursive characterization of the number of roots given by Bernshtein's theorem to establish that $\prod_{i=1}^{n-1} (n-i)^{p_i}$ is a lower bound on the maximal number of totally mixed Nash equilibria, so the mean number of totally mixed Nash equilibria is greater than or equal to

$$2^{-p}\sqrt{\prod_{i=1}^{n-1} (n-i)^{p_i}}.$$

(Note that $\prod_{i=1}^{n-1} (n-i)^{p_i}$ is maximized if the agents are indexed so that $p_1 \geq \cdots \geq p_n$.) For example, for normal forms in which $p_1 = \cdots = p_n = k$, the bound reduces to $\left((n-1)!/2^{2n}\right)^{k/2}$. For integral n, $(n-1)!/2^{2n} \geq 1$ if and only if $n \geq 12$, so that the mean number of totally mixed equilibria of the random game in which all agents have the same number of pure strategies grows exponentially when there are twelve or more agents.

5. The Full Support Case

This section proves the full support case of the Main Theorem. For $i = 1, \ldots, n$ we identify $\mathbb{R}^{p_i+1}$ with $\mathbb{R}^{T_i}$, and we specialize the multihomogeneous system of the last section to the system (N1) by setting

$$\mathcal{A}_{p_1+\cdots+p_{i-1}+1} = \cdots = \mathcal{A}_{p_1+\cdots+p_i} = \mathcal{B}_i.$$

Equivalently, $D = D(\mathbf{p})$. The argument is a matter of comparing the random multihomogeneous system for this data with the random game model. Naively one might hope that the two distributions of coefficient vectors are the same, possibly after rescaling, but the situation is not so simple.

The geometric basis of our argument has the following informal description. Suppose several equations in a sparse system have the same support. Each equation amounts to a requirement that the vector of monomials in the support is orthogonal to the coefficient vector, so collectively the equations amount to the requirement that the vector of monomials in the support is orthogonal to the plane spanned by the coefficient vectors of the various equations. Any distribution on the space of coefficient vectors induces a distribution on the relevant space of planes, and it is possible that different distributions on the space of coefficient vectors induce the same distribution on the space of planes, in which case they must induce the same distribution of roots. Let $\Gamma := \Gamma_1 \times \cdots \times \Gamma_n$ where, for each i, Γ_i is the Grassman manifold(11) of p_i-dimensional linear subspaces of $\mathbb{R}^{\mathcal{B}_i}$.

(11) This manifold has the following description. The elements of Γ_i are the linear subspaces of dimension

Beginning with the construction for the random multihomogeneous system, we decompose a coefficient vector $\mathbf{f} = (f_1, \ldots, f_p)$ as $\mathbf{f} = (\mathbf{g}_1, \ldots, \mathbf{g}_n)$ by setting

$$\mathbf{g}_i = (g_{i1}, \ldots, g_{ip_i}) := (f_{p_1+\cdots+p_{i-1}+1}, \ldots, f_{p_1+\cdots+p_i}).$$

Let $\mathcal{M}^* := \mathcal{G}_1^* \times \cdots \times \mathcal{G}_n^* \subset \mathcal{M}$ where, for each i,

$$\mathcal{G}_i := \mathcal{M}_{p_1+\cdots+p_{i-1}+1} \times \cdots \times \mathcal{M}_{p_1+\cdots+p_i}$$

and $\mathcal{G}_i^*$ is the set of $\mathbf{g}_i = (g_{i1}, \ldots, g_{ip_i}) \in \mathcal{G}_i$ such that $g_{i1}, \ldots, g_{ip_i}$ are linearly independent. Of course each $\mathcal{G}_i^*$ has full measure in $\mathcal{G}_i$, so $\mathcal{M}^*$ has full measure in $\mathcal{M}$. Define the map $\varphi : \mathcal{M}^* \to \Gamma$ by setting $\varphi(\mathbf{f}) := (\varphi_1(\mathbf{g}_1), \ldots, \varphi_n(\mathbf{g}_n))$ where each $\varphi_i : \mathcal{G}_i^* \to \Gamma_i$ is the map

$$\varphi_i(\mathbf{g}_i) := \operatorname{span}\{g_{i1}, \ldots, g_{ip_i}\}.$$

Comparing the definition of $\mathcal{B}_i$ with (8) shows that $\eta(a) = 1$ for all $a \in \mathcal{B}_i$, so the inner product defined there coincides with the usual one. This means that there is no ambiguity in saying that for each i and $h = 1, \ldots, p_i$, $\mathbf{x} \in \mathbb{R}^{T_1} \times \cdots \times \mathbb{R}^{T_n}$ is a root of g_{ih} if and only if $\mathbf{x}^{\mathcal{B}_i}$ is orthogonal to g_{ih}. Therefore $\mathbf{f}(\mathbf{x}) = 0$ if and only if

$$\mathbf{x}^{\mathcal{B}_1} \perp \varphi_1(\mathbf{g}_1), \ldots, \mathbf{x}^{\mathcal{B}_n} \perp \varphi_n(\mathbf{g}_n),$$

and the distribution of roots of a randomly distributed $\mathbf{f} \in \mathcal{M}$ is the same as the distribution of points in N satisfying

$$\mathbf{x}^{\mathcal{B}_1} \perp \gamma_1, \ldots, \mathbf{x}^{\mathcal{B}_n} \perp \gamma_n \tag{9}$$

when $\gamma \in \Gamma$ is a random n-tuple of planes with distribution $\mathbf{U}_{\mathcal{M}^*} \circ \varphi^{-1}$.

We now give a similar description of the distribution of solutions of (N1) for the random game model. To avoid systems with redundant equations we restrict attention to $M^* := M_1^* \times \cdots \times M_n^*$ where, for each i,

$$M_i^* := \{\, u_i \in M_i : \kappa_i^1(u_i), \ldots, \kappa_i^{p_i}(u_i) \text{ are linearly independent} \,\}.$$

It is easy to show that each M_i^* has full measure in M_i, so M^* has full measure in M. Define $\psi : M^* \to \Gamma$ by setting $\psi(u) = (\psi_1(u_1), \ldots, \psi_n(u_n))$ where each $\psi_i : M_i^* \to \Gamma_i$ is the map

$$\psi_i(u_i) := \operatorname{span}\{\kappa_i^1(u_i), \ldots, \kappa_i^{p_i}(u_i)\}.$$

Then $(u, \tau) \in V$ if and only if

$$\tau^{\mathcal{B}_1} \perp \psi_1(u_1), \ldots, \tau^{\mathcal{B}_n} \perp \psi_n(u_n),$$

---

p_i in $\mathbb{R}^{\mathcal{B}_i}$. Consider a pair of subspaces $V, W \subset \mathbb{R}^{\mathcal{B}_i}$ of dimensions p_i and $|\mathcal{B}_i| - p_i$ respectively that are complementary, in the sense that $V \cap W = \{0\}$ and $V + W = \mathbb{R}^{\mathcal{B}_i}$. The map that takes an element $\lambda \in L(V, W)$ to the plane $\{\, v + \lambda(v) : v \in V \,\}$ is a parameterization of a subset of Γ_i, and it is not hard to check that two such parameterizations have C^∞ (in fact analytic) overlap, so the collection of such parameterizations defines a differentiable structure on Γ_i.

and consequently the distribution of solutions of (N1) in the full support case coincides with the distribution of $\tau \in N$ satisfying (9) when $\gamma \in \Gamma$ is a random n-tuple of planes with distribution $\mathbf{U}_{M^*} \circ \psi^{-1}$.

The result will follow once we show that $\mathbf{U}_{\mathcal{M}^*} \circ \varphi^{-1} = \mathbf{U}_{M^*} \circ \psi^{-1}$. To begin with, observe that since $\mathbf{g}_1, \ldots, \mathbf{g}_n$ and $u_1, \ldots, u_n$ are statistically independent, the first of these is the product of the measures $\mathbf{U}_{\mathcal{G}_i} \circ \varphi_i^{-1}$, while the second is the product of the measures $\mathbf{U}_{M_i^*} \circ \psi_i^{-1}$. Therefore it suffices to show that $\mathbf{U}_{\mathcal{G}_i^*} \circ \varphi_i^{-1} = \mathbf{U}_{M_i^*} \circ \psi_i^{-1}$ for each i.

In general, if a group H acts on two spaces X and Y, a map $F : X \to Y$ is said to be H-*equivariant* (or simply *equivariant* if H is clear) if $F(hx) = hF(x)$ for all $x \in X$ and $h \in H$. If ρ is an invariant measure on X and $F : X \to Y$ is H-equivariant, then $\rho \circ F^{-1}$ is an invariant measure on Y since, for measurable $E \subset Y$ and $h \in H$, we have

$$(\rho \circ F^{-1})(hE) = \rho(F^{-1}(hE)) = \rho(hF^{-1}(E)) = \rho(F^{-1}(E)) = (\rho \circ F^{-1})(E).$$

For $i = 1, \ldots, n$ there is an obvious action of $O(\mathbb{R}^{\mathcal{B}_i})$ on Γ_i given by setting $OV := \{\, Ov : v \in V \,\}$ for $V \in \Gamma_i$ and $O \in O(\mathbb{R}^{\mathcal{B}_i})$. It is easy to show that the action of $O(\mathbb{R}^{\mathcal{B}_i})$ on Γ_i is transitive: choose orthonormal bases for $V, V' \in \Gamma_i$ and construct an orthogonal transformation taking the basis of V to the basis of V'. Consequently Proposition 3.1 implies that Γ_i has a unique invariant measure.

Let $O(\mathbb{R}^{\mathcal{B}_i})$ act on $\mathcal{G}_i$ by the rule $O(g_{i1}, \ldots, g_{ip_i}) := (Og_{i1}, \ldots, Og_{ip_i})$. Clearly $\mathcal{G}_i^*$ is invariant under this action, so there is induced (by restriction) action on $\mathcal{G}_i^*$. The equivariance of φ_i follows directly from its definition. Each of the cartesian factors of $\mathcal{G}_i$ is the unit sphere in $\mathbb{R}^{\mathcal{B}_i}$, and the uniform distribution on this sphere is invariant under the action of $O(\mathbb{R}^{\mathcal{B}_i})$. In general, if μ is an invariant measure for the action of H on a space X and $k \geq 1$ is an integer, then the k-fold product $\mu \times \cdots \times \mu$ is an invariant measure for the action $h(x_1, \ldots, x_k) := (hx_1, \ldots, hx_k)$ of H on X^k. Therefore $\mathbf{U}_{\mathcal{G}_i}$ is an invariant measure under the action on $\mathcal{G}_i$. Since $\mathcal{G}_i^*$ has full measure in $\mathcal{G}_i$, $\mathbf{U}_{\mathcal{G}_i^*}$ is an invariant measure for the induced action on $\mathcal{G}_i^*$. We conclude that $\mathbf{U}_{\mathcal{G}_i^*} \circ \varphi_i^{-1}$ is the invariant measure on Γ_i.

For $O \in O(\mathbb{R}^{\mathcal{B}_i})$ let

$$T_O^i := \iota_i^{-1} \circ O \circ \iota_i : \mathbb{R}^{T_{-i}} \to \mathbb{R}^{T_{-i}},$$

where $\iota_i : \mathbb{R}^{T_{-i}} \to \mathbb{R}^{\mathcal{B}_i}$ is the function introduced in Section 2. Since $S_1 = T_1, \ldots, S_n = T_n$, the decomposition of u_i introduced in Section 2 reduces to $u_i = (v_i^{t_i^0}, \ldots, v_i^{t_i^{p_i}})$. There is an action of $O(\mathbb{R}^{\mathcal{B}_i})$ on $\mathbb{R}^S = \mathbb{R}^T$ given by setting

$$Ou_i := \left(T_O^i(v_i^{t_i^0}), \ldots, T_O^i(v_i^{t_i^{p_i}})\right). \tag{10}$$

Since T_O^i is linear, $\kappa_i^1, \ldots, \kappa_i^{p_i}$ are equivariant, and consequently ψ_i is equivariant. Since $O \in O(\mathbb{R}^{\mathcal{B}_i})$ and ι_i is a relabelling of coordinates, hence inner product preserving, we have $T_O^i \in O(\mathbb{R}^{T_{-i}})$. Therefore the map $u_i \mapsto Ou_i$ is an orthogonal transformation, so that M_i and M_i^* are invariant, and these spaces have actions induced by restriction. The uniform distribution on M_i is invariant and M_i^* has full measure in M_i, so $\mathbf{U}_{M_i^*}$ is an invariant measure. Therefore $\mathbf{U}_{M_i^*} \circ \psi_i^{-1}$ is also the invariant measure on Γ_i, hence equal to $\mathbf{U}_{\mathcal{G}_i^*} \circ \varphi_i^{-1}$ as desired.

6. Extension to the General Case

The passage from the full support case to the general case is a matter, for given measurable $E \subset N$, of comparing the following three quantities: (a) the expected number of roots in E of (N1) and (N2); (b) the expected number of roots in E of (N1); (c) the expected number of roots in E of (N1) for the spherical model applied to the truncated game obtained by removing all pure strategies in $S_1 \setminus T_1, \ldots, S_n \setminus T_n$. The full support case characterizes (c). As we explained in Section 2, the distribution of coefficient vectors given by (N1) is equivalent, from the point of view of the induced distribution of roots, to the distribution for the truncated game obtained by eliminating pure strategies outside $T_1, \ldots, T_n$, so (b) and (c) are the same. Thus it suffices to prove that the ratio of (a) and (b) is $\prod_{i=1}^{n} r(p_i, |S_i \setminus T_i|)$.

Let

$$\tilde{V} := \{ (u, \tau) \in M \times N : (\mathrm{N1}) \}, \quad \text{so that} \quad V = \{ (u, \tau) \in \tilde{V} : (\mathrm{N2}) \}.$$

Let $\tilde{\pi}_1 : \tilde{V} \to M$ and $\tilde{\pi}_2 : \tilde{V} \to N$ be the restrictions to $\tilde{V}$ of the natural projections from $M \times N$. Then π_1 and π_2 are the restrictions of $\tilde{\pi}_1$ and $\tilde{\pi}_2$ to V. Let $\tilde{\nu}$ be the measure on N defined, for measurable $E \subset N$, by

$$\tilde{\nu}(E) := \int_M \#\Big(\tilde{\pi}_1^{-1}(u) \cap \tilde{\pi}_2^{-1}(E)\Big)\, d\mathbf{U}_M(u).$$

The remainder is devoted to the proof of:

Proposition 6.1: *For all measurable $E \subset N$ with $\tilde{\nu}(E) > 0$,*

$$\frac{\nu(E)}{\tilde{\nu}(E)} = \prod_{i=1}^{n} r(p_i, |S_i \setminus T_i|).$$

Since the Borel σ-algebra is generated by the open sets, it suffices to prove this for E open. The proof is fairly lengthy, paralleling many of the steps in the proof of Proposition 4.1 in the Appendix. We divide it into a number of shorter steps.

6.1. *Manifold Theoretic Properties*

Proposition 6.2: *$\tilde{V}$ is a smooth manifold of the same dimension as M.*

Proof: By the regular value theorem (e.g., Guillemin and Pollack (1965)) it suffices to show that 0 is a regular value of the restriction to $M \times N$ of the map

$$(u, \tau) \mapsto \big(\langle \kappa_1^1(u_1), \tau^{\mathcal{B}_1}\rangle, \ldots, \langle \kappa_1^{p_1}(u_1), \tau^{\mathcal{B}_1}\rangle, \ldots, \langle \kappa_n^1(u_n), \tau^{\mathcal{B}_n}\rangle, \ldots, \langle \kappa_1^{p_n}(u_n), \tau^{\mathcal{B}_n}\rangle\big).$$

It is easy to verify that all points in $(\mathbb{R}^S)^n \times \prod_i (\mathbb{R}^{T_i} \setminus \{0\})$ are regular points of this function: if no component τ_i of τ is vanishing, then $\tau^{\mathcal{B}_i} \neq 0$, and each component $\langle \kappa_i^j(u_i), \tau^{\mathcal{B}_i}\rangle$ can be freely varied, without affecting any other component, by varying $u_i(t_i^j, \cdot)$.

We wish to show that each point $(u, \tau) \in \tilde{V}$ is also a regular point of the restriction of the function to $M \times N$. Let (v, ξ) denote a vector in $(I\!R^S)^n \times \prod_i I\!R^{S_i}$, thought of as a tangent vector at (u, τ). For each i and j, $\langle \kappa_i^j(u_i), \tau^{\mathcal{B}_i} \rangle$ is linear as a function of u_i and multilinear as a function of τ. If $(u, \tau) \in V$, so that $\langle \kappa_i^j(u_i), \tau^{\mathcal{B}_i} \rangle = 0$, each v_i is a scalar multiple of u_i and each ξ_i is scalar multiple of τ_i, then the derivative of $\langle \kappa_i^j(u_i), \tau^{\mathcal{B}_i} \rangle$ along the vector (v, ξ) is zero. Consequently, if $Df(u, \tau)$ is surjective, so is the restriction of $Df(u, \tau)$ to $T_{(u,\tau)}(M \times N)$, which consists of those vectors (v, ξ) with v_i orthogonal to u_i and ξ_i orthogonal to τ_i for each i. ∎

The set of utilities that have a particular $\tau \in N$ as a solution of (N1) is well behaved, and varies in a smooth manner as τ varies. The *fiber* over τ is

$$\tilde{V}_\tau := \{ u \in M : (u, \tau) \in \tilde{V} \}.$$

It will sometimes be convenient to abuse notation by not distinguishing explicitly between $\tilde{V}_\tau$ and $\tilde{\pi}_2^{-1}(\tau) \subset M \times N$. The fiber over τ is a cartesian product of subspheres of the spheres M_i given by the equations defining $\tilde{V}$, so its topology is independent of τ, suggesting the following result.

Lemma 6.3: *$\tilde{\pi}_2 : \tilde{V} \to N$ is a C^∞ fibration.*

Proof: Standard methods (e.g., Milnor and Stasheff (1974, §3)) can be used to prove that for each i,

$$\{ (u_i, \tau) \in I\!R^S \times N : \langle \kappa_i^j(u_i), \tau^{\mathcal{B}_i} \rangle = 0 \text{ for all } j = 1, \dots, p_i \}$$

is a C^∞ vector bundle over N, after which $\tilde{V}$ is seen to be the product (in the sense of taking the cartesian product of the fibers over each base point) of the associated sphere bundles[12]. ∎

Application of the Shub-Smale integral formula (Shub and Smale (1993, p. 273), Blum *et. al.* (1998, p. 240)) yields:

Proposition 6.4: *For any open set $Z \subset \tilde{V}$,*

$$\int_M \#(\tilde{\pi}_1^{-1}(u) \cap Z)\, du = \int_N \int_{\tilde{V}_\tau \cap Z} \det\left(A^*(u, \tau) A(u, \tau)\right)^{1/2} du\, d\tau \qquad (11)$$

where $A(u, \tau) : T_\tau(N) \to T_u(M)$ is the linear map whose graph is the orthogonal complement $\perp_{(u,\tau)}$ of $T_{(u,\tau)}\tilde{V}_\tau$ in $T_{(u,\tau)}\tilde{V}$ and $A^*(u, \tau)$ is the adjoint of $A(u, \tau)$.

---

(12) Recall footnote 5.

6.2. *More Equivariant Group Actions*

The next goal is to show that when $Z = V$ or $Z = \tilde{V}$, the inner integral on the right hand side of (11) does not depend on τ. This will be done by exploiting the symmetry that we built into the mathematical apparatus with the choice of the distribution of payoffs and the displacement of the space of mixed strategies for each agent from the unit simplex to the positive orthant of the unit sphere. This symmetry is expressed in various actions of the group

$$G := O(\mathbb{R}^{T_1}) \times \cdots \times O(\mathbb{R}^{T_n}).$$

To begin with, there is the action of G on $\mathbb{R}^{T_1} \times \cdots \times \mathbb{R}^{T_n}$ given by the rule

$$g\tau := (g_1\tau_1, \ldots, g_n\tau_n).$$

Of course N is invariant under this action, and there is consequently an action of G on N given by restriction.

In the following we use a general result about group actions. Suppose that the groups H and H' both act on a space X. Then the rule $(h, h')x = h(h'x)$ defines an action of the product group $H \times H'$ on X if $h'(hx) = h(h'x)$ for all $x \in X$, $h \in H$ and $h' \in H$. (In fact this condition is necessary as well as sufficient: if $h'(hx) \neq h(h'x)$ then the effect of $(e, h')(h, e')$ would not agree with the effect of (h, h').)

Lemma 6.5: *For each i there is a unique action of G on $\mathbb{R}^{\mathcal{B}_i}$, by orthogonal transformations, with respect to which the map $\tau \mapsto \tau^{\mathcal{B}_i}$ is G-equivariant.*

Proof: For each $h \neq i$ we may think of $\mathbb{R}^{\mathcal{B}_i}$ as a cartesian product of copies of $\mathbb{R}^{T_h}$ indexed by the various elements of $\prod_{a \neq h,i} T_a$. Equivariance clearly requires that the action of any $O_h \in O(\mathbb{R}^{T_h})$ is to simultaneously transform each copy according to O_h. To show that this rule does, indeed, define a group action, we need to show that the effect of $O_h \in O(\mathbb{R}^{T_h})$ commutes with the effect of $O_{h'} \in O(\mathbb{R}^{T_{h'}})$ when $h \neq h'$.

It is probably easier to understand the idea in a general and abstract setting. Let A and B be nonempty finite sets, and denote the standard unit basis vectors of $\mathbb{R}^{A \times B}$ by $\mathbf{e}_{(\alpha,\beta)}$. Let the action of $O(\mathbb{R}^A)$ (resp. $O(\mathbb{R}^B)$) on $\mathbb{R}^{A\times B}$ be given by construing $\mathbb{R}^{A\times B}$ as a cartesian product of copies of $\mathbb{R}^A$ indexed by B (resp. $\mathbb{R}^B$ indexed by A) with the action given by simultaneously transforming each copy. If Y and Z are elements of $O(\mathbb{R}^A)$ and $O(\mathbb{R}^B)$ respectively, with matrices $(y_{\gamma\alpha})$ and $(z_{\delta\beta})$, then for each $(\alpha, \beta) \in A \times B$ we have

$$Z(Y\mathbf{e}_{(\alpha,\beta)}) = Z\Big(\sum_{\gamma\in A} y_{\gamma\alpha}\mathbf{e}_{(\gamma,\beta)}\Big) = \sum_{\gamma\in A} y_{\gamma\alpha}\Big(Z\mathbf{e}_{(\gamma,\beta)}\Big) = \sum_{\gamma\in A} y_{\gamma\alpha}\Big(\sum_{\delta\in B} z_{\delta\beta}\mathbf{e}_{(\gamma,\delta)}\Big)$$

$$= \sum_{\delta\in B} z_{\delta\beta}\Big(\sum_{\gamma\in A} y_{\gamma\alpha}\mathbf{e}_{(\gamma,\delta)}\Big) = \sum_{\delta\in B} z_{\delta\beta}\Big(Y\mathbf{e}_{(\alpha,\delta)}\Big) = Y\Big(\sum_{\delta\in B} z_{\delta\beta}\mathbf{e}_{(\alpha,\delta)}\Big) = Y(Z\mathbf{e}_{(\alpha,\beta)}). \blacksquare$$

As in Section 2 we write

$$u_i = \big((v_i^{t_i^0}, w_i^{t_i^0}), \ldots, (v_i^{t_i^{p_i}}, w_i^{t_i^{p_i}}), (v_i^{s_i^1}, w_i^{s_i^1}), \ldots, (v_i^{s_i^{q_i}}, w_i^{s_i^{q_i}})\big).$$

Generalizing the action (10) for the full support case, there is an action of $O(\mathbb{R}^{\mathcal{B}_i})$ on $\mathbb{R}^S$ given by setting

$$Ou_i := \big((T^i_O(v_i^{t_i^0}), w_i^{t_i^0}), \ldots, (T^i_O(v_i^{t_i^{p_i}}), w_i^{t_i^{p_i}}), (T^i_O(v_i^{s_i^1}), w_i^{s_i^1}), \ldots, (T^i_O(v_i^{s_i^{q_i}}), w_i^{s_i^{q_i}})\big),$$

where $T^i_O := \iota_i^{-1} \circ O \circ \iota_i$ is the function introduced in the last section. The last result may be reinterpreted as saying that there is a homomorphism

$$\epsilon_i : G \to O(\mathbb{R}^{\mathcal{B}_i}) \text{ satisfying } (g\tau)^{\mathcal{B}_i} = \epsilon_i(g)\tau^{\mathcal{B}_i}$$

for all $\tau \in \mathbb{R}^{T_1} \times \cdots \times \mathbb{R}^{T_n}$ and $g \in G$. In general, whenever the range of a homomorphism acts on a space, there is an obvious induced action of the domain. In this sense we may define an action of G on the space $\mathbb{R}^S$ of utilities for agent i by setting $gu_i := \epsilon_i(g)u_i$.

For $j = 1, \ldots, p_i$, $u_i \in \mathbb{R}^S$, and $g \in G$ we now have

$$\begin{aligned}
\langle \kappa_i^j(gu_i), (g\tau)^{\mathcal{B}_i} \rangle &= \langle \iota_i(T^i_{\epsilon_i(g)} v_i^{t_i^j} - T^i_{\epsilon_i(g)} v_i^{t_i^0}), (g\tau)^{\mathcal{B}_i} \rangle \\
&= \langle \epsilon_i(g)(\iota_i(v_i^{t_i^j} - v_i^{t_i^0})), \epsilon_i(g)\tau^{\mathcal{B}_i} \rangle \\
&= \langle \iota_i(v_i^{t_i^j} - v_i^{t_i^0}), \tau^{\mathcal{B}_i} \rangle \\
&= \langle \kappa_i^j(u_i), \tau^{\mathcal{B}_i} \rangle.
\end{aligned}$$

Here the first and last equalities derive from the definitions of κ_i^j and gu_i, the second combines the definition of $T^i_{\epsilon_i(g)}$ with the last result, and the third follows from $\epsilon_i(g) \in O(\mathbb{R}^{\mathcal{B}_i})$. Similarly,

$$\langle \lambda_i^k(gu_i), (g\tau)^{\mathcal{B}_i} \rangle = \langle \lambda_i^k(u_i), \tau^{\mathcal{B}_i} \rangle$$

for all $k = 1, \ldots, q_i$, $u_i \in \mathbb{R}^S$, and $g \in G$.

There is now an action of G on M given by $gu = (gu_1, \ldots, gu_n)$ and an action on $M \times N$ given by $g(u, \tau) = (gu, g\tau)$. Of course an important consequence of the equations above is:

Lemma 6.6: *$\tilde{V}$ and V are invariants of the action of G on $M \times N$: for all $g \in G$ we have $g\tilde{V} = \tilde{V}$ and $gV = V$.*

The next two results show that the inner integral of the right hand side of (11) is independent of τ, so that the right hand side of (11) is the inner integral at an arbitrary point times the volume of N.

Lemma 6.7: *Let $Z \subset \tilde{V}$ be an open set that is invariant under the action of G: $gZ = Z$ for all $g \in G$. Then for any $\tau \in N$,*

$$\int_{\tilde{V}_\tau \cap Z} \det(A^*(u,\tau)A(u,\tau))^{1/2}\, du = \int_{\tilde{V}_{g\tau} \cap Z} \det(A^*(u,g\tau)A(u,g\tau))^{1/2}\, du.$$

Proof: As a general matter, if P is a submanifold of $\mathbb{R}^k$ that is invariant under the action of an orthogonal transformation $O \in O(\mathbb{R}^k)$, then the tangent space T_pP at a point $p \in P$ (understood concretely as a subspace of $\mathbb{R}^k$) is invariant in the sense that $T_{Op}P = O(T_pP)$. If P is invariant under the action of a subgroup $H \subset O(\mathbb{R}^k)$, then in this sense there is an induced action on the tangent bundle

$$TP := \{ (p,v) : p \in P, v \in T_pP \}$$

given by restricting the action of H on $\mathbb{R}^k \times \mathbb{R}^k$.

For all $g \in G$ we have

$$T_{(gu,g\tau)}\tilde{V}_{g\tau} = g(T_{(u,\tau)}\tilde{V}_\tau) \quad \text{and} \quad T_{(gu,g\tau)}\tilde{V} = g(T_{(u,\tau)}\tilde{V}),$$

so $\perp_{(gu,g\tau)}= g(\perp_{(u,\tau)})$. Now the relations

$$A(gu,g\tau) = g \circ A(u,\tau) \circ g^{-1} \quad \text{and} \quad A^*(gu,g\tau) = g \circ A^*(u,\tau) \circ g^{-1}$$

follow directly from the definitions of $A(u,\tau)$ and its adjoint. The map $(u,\tau) \mapsto (gu,g\tau)$ is an isometry between $\tilde{V}_\tau \cap Z$ and $\tilde{V}_{g\tau} \cap Z$, so the change of variables formula, followed by application of the invariances above and elementary properties of the determinant, yields the desired equation. ∎

For each m the action of the orthogonal group $O(\mathbb{R}^{m+1})$ on the unit sphere S^m is transitive. As the cartesian product of transitive actions, the action of G on N is also transitive. Therefore the last result implies:

Lemma 6.8: *For any open set $Z \subset \tilde{V}$ that is invariant under the action of G, any open $Y \subset N$, and any $\tau \in N$,*

$$\int_M \#(\tilde{\pi}_1^{-1}(u) \cap \tilde{\pi}_2^{-1}(Y) \cap Z)\, du = \mathrm{vol}(Y) \cdot \int_{\tilde{V}_\tau \cap Z} \det(A^*(u,\tau)A(u,\tau))^{1/2}\, du.$$

The following result extracts the relevant consequences of our work up to this point.

Lemma 6.9: *For all open $Y \subset N$ with $\tilde{\nu}(Y) > 0$ it is the case, for any $\tau \in N$, that*

$$\frac{\nu(Y)}{\tilde{\nu}(Y)} = \frac{\int_{\tilde{V}_\tau \cap V} \det(A^*(u,\tau)A(u,\tau))^{1/2}\, du}{\int_{\tilde{V}_\tau} \det(A^*(u,\tau)A(u,\tau))^{1/2}\, du}. \tag{12}$$

Proof: Applying Proposition 6.4 to evaluate the definitions of $\tilde{\nu}(Y)$ and $\nu(Y)$ yields

$$\begin{aligned}\frac{\nu(Y)}{\tilde{\nu}(Y)} &= \frac{\int_N \int_{\tilde{V}_\tau \cap V \cap \tilde{\pi}_2^{-1}(Y)} \det(A^*(u,\tau)A(u,\tau))^{1/2}\, du\, d\tau}{\int_N \int_{\tilde{V}_\tau \cap \tilde{\pi}_2^{-1}(Y)} \det(A^*(u,\tau)A(u,\tau))^{1/2}\, du\, d\tau} \\ &= \frac{\int_Y \int_{\tilde{V}_\tau \cap V} \det(A^*(u,\tau)A(u,\tau))^{1/2}\, du\, d\tau}{\int_Y \int_{\tilde{V}_\tau} \det(A^*(u,\tau)A(u,\tau))^{1/2}\, du\, d\tau}.\end{aligned}$$

Since $\tilde{V}$ and V are invariant under the action of G, Lemma 6.8 implies that the inner integrals of both the numerator and the denominator are independent of τ. ∎

6.3. *Reduction to a Smaller Set of Variables*

In the further evaluation of the right hand side of (12) we are free to choose τ as we please. The choice of $\tau = t^0 = (t_1^0, \ldots, t_n^0)$ has the advantage that the integrand depends on only a few of the variables $u_i(s)$ defining u. For each i let

$$R_{-i} = \{\, t_{-i} \in T_{-i} : \ t_j \neq t_j^0 \text{ for exactly one } j\,\},$$

and let $R_i = T_i \times R_{-i} \subset S$.

Proposition 6.10: *Suppose that $u, u' \in \tilde{V}_{t^0}$, and that, for each i, the projections of u_i and u_i' to $\mathbb{R}^{R_i}$ agree. Then $A(u, t^0) = A(u', t^0)$.*

Proof: In order to be able to express its derivative using conventional notation, for $i = 1, \ldots, n$ we introduce the function $\theta_i(\tau) := \tau^{\mathcal{B}_i}$ from N to $\mathbb{R}^{\mathcal{B}_i}$. Then $(u, \tau) \in M \times N$ is in $\tilde{V}$ if and only if $\langle \kappa_i^j(u_i), \theta_i(\tau)\rangle = 0$ for all $i = 1, \ldots, n$ and $j = 1, \ldots, p_i$.

Fix $u \in \tilde{V}_{t^0}$. Now $T_{(u,t^0)}\tilde{V}$ is the set of $(w, \psi) \in (\mathbb{R}^S)^n \times \mathbb{R}^{T_1} \times \cdots \times \mathbb{R}^{T_n}$ such that $w_i \perp u_i$ and $\psi_i \perp t_i^0$ for all i (so that $(w, \psi) \in T_{(u,t^0)}(M \times N)$) and

$$0 = \langle D\kappa_i^j(u_i)w_i, \theta_i(t^0)\rangle + \langle \kappa_i^j(u_i), D\theta_i(t^0)\psi\rangle$$

for all i and $j = 1, \ldots, p_i$. In view of the linearity of κ_i^j, the latter condition reduces to

$$0 = \langle \kappa_i^j(w_i), \theta_i(t^0)\rangle + \langle \kappa_i^j(u_i), D\theta_i(t^0)\psi\rangle.$$

In turn $T_{(u,t^0)}\tilde{V}_{t^0} = \{\, (w, \psi) \in T_{(u,t^0)}\tilde{V} : \psi = 0\,\}$ is seen to be the set of $(w, 0) \in (\mathbb{R}^S)^n \times \mathbb{R}^{T_1} \times \cdots \times \mathbb{R}^{T_n}$ such that for all i, $w_i \perp u_i$ and $0 = \langle \kappa_i^j(w_i), \theta_i(t^0)\rangle$ for all $j = 1, \ldots, p_i$. Therefore $\perp_{(u,t^0)}$ is the set of $(w, \psi) \in (\mathbb{R}^S)^n \times \mathbb{R}^{T_1} \times \cdots \times \mathbb{R}^{T_n}$ such that for all i:

(a) $w_i \perp u_i$;
(b) $\psi_i \perp t_i^0$;
(c) $0 = \langle \kappa_i^j(w_i), \theta_i(t^0)\rangle + \langle \kappa_i^j(u_i), D\theta_i(t^0)\psi\rangle$ for all $j = 1, \ldots, p_i$;
(d) w_i is perpendicular to all w_i' that are perpendicular to u_i and which satisfy $0 = \langle \kappa_i^j(w_i'), \theta_i(t^0)\rangle$ for all $j = 1, \ldots, p_i$.

Note that the equilibrium condition implies that $0 = \langle \kappa_i^j(u_i), \theta_i(t^0)\rangle$ for all i, so that (a) and (d) together are equivalent to

(d′) $w_i \perp w_i'$ for all $w_i' \in \mathbb{R}^S$ satisfying $0 = \langle \kappa_i^j(w_i), \theta_i(t^0)\rangle$ for all $j = 1, \ldots, p_i$.

The characterization of $\perp_{(u,t^0)}$ given by (b), (c), and (d′) refers to u only through the terms $\langle \kappa_i^j(u_i), D\theta_i(t^0)\psi\rangle$. Direct differentiation of the function θ_i shows that these terms depend only on the components of u_i for the pure strategy profiles in R_i. ∎

The fiber $\tilde{V}_{t^0}$ is a cartesian product over i of the subset of the sphere M_i given by the conditions

$$0 = \langle \kappa_i^1(u_i), \theta_i(t^0)\rangle = \cdots = \langle \kappa_i^{p_i}(u_i), \theta_i(t^0)\rangle,$$

i.e., $u_i(t^0) = u_i(t^0|t_i^1) = \cdots = u_i(t^0|t_i^{p_i})$. The next step is to simplify by applying a change of variables that is derived from a diffeomorphism between this subset of M_i and a cartesian product in which one factor is the space of the variables that do not affect the integrand in (12), by virtue of the last result. In the reparameterized integral the effect of these variables on the integral reduces to multiplication by the volume of the relevant portion of this space.

Although the next two lemmas are notationally cumbersome, the underlying idea is a computation at the level of elementary calculus, and has already appeared in the final calculation of Section 3. The complex appearance is due entirely to the complexity of the intended application's environment.

Lemma 6.11: *For $i = 1, \ldots, n$ suppose that:*

(a) a_i, b_i, c_i are nonnegative integers;

(b) $\tilde{D}^i$ is the open unit disk in $\mathbb{R}^{a_i}$;

(c) $\tilde{E}^i$ is the subsphere of the unit sphere in $\mathbb{R}^{b_i+c_i}$ consisting of those unit vectors $(q^i, r^i) \in \mathbb{R}^{b_i} \times \mathbb{R}^{c_i}$ such that $q_1^i = \cdots = q_{b_i}^i$;

(d) $\tilde{F}^i$ is the subsphere of the unit sphere in $\mathbb{R}^{a_i+b_i+c_i}$ consisting of those unit vectors $(p^i, q^i, r^i) \in \mathbb{R}^{b_i} \times \mathbb{R}^{c_i}$ such that $q_1^i = \cdots = q_{b_i}^i$.

For each i define $\gamma^i : \tilde{D}^i \times \tilde{E}^i \to \tilde{F}^i$ by

$$\gamma^i(p^i, (q^i, r^i)) = \big(p^i, (1 - \|p^i\|^2)^{1/2} q^i, (1 - \|p^i\|^2)^{1/2} r^i\big).$$

Let $\tilde{D} := \prod_i \tilde{D}^i$ and $\tilde{F} := \prod_i \tilde{F}^i$, let

$$\tilde{E} := \{\, (q, r) \in \prod_i \mathbb{R}^{b_i} \times \prod_i \mathbb{R}^{c_i} : (q^1, r^1) \in \tilde{E}^1, \ldots, (q^n, r^n) \in \tilde{E}^n \,\},$$

and let $\gamma : \tilde{D} \times \tilde{E} \to \tilde{F}$ be the function

$$\gamma(p, (q, r)) := \big(\gamma_1(p^1, (q^1, r^1)), \ldots, \gamma_n(p^n, (q^n, r^n))\big).$$

Then for all $(p, (q, r)) \in \tilde{D} \times \tilde{E}$,

$$|\det D\gamma(p, (q, r))| = \prod_{i=1}^{n} (1 - \|p^i\|^2)^{\frac{c_i - 1}{2}}.$$

Proof: First we compute that

$$\left\|D\gamma^i(p^i, (q^i, r^i))\Big(\frac{p^i}{\|p^i\|}, 0\Big)\right\| = \left\| \frac{d(x, (1 - x^2)^{1/2})}{dx}\Big|_{x = \|p^i\|} \right\| = (1 - \|p^i\|^2)^{-1/2}.$$

Consider an orthonormal basis for $T_{p_i}\tilde{D}_i$ that includes $p^i/\|p^i\|$. The images of its elements under $D\gamma^i(p^i, (q^i, r^i))$ are pairwise orthogonal, and except for the image of $p^i/\|p^i\|$, each has unit length. The image under $D\gamma^i(p^i,(q^i,r^i))$ of an orthonormal basis of $T_{(q_i,r_i)}\tilde{E}_i$ will be pairwise orthogonal, and each will have length $(1-\|p^i\|^2)^{\frac{1}{2}}$. Furthermore, the images of the two bases are orthogonal, and $\dim \tilde{E}_i = c_i$, so

$$|\det D\gamma^i(p^i,(q^i,r^i))| = (1-\|p^i\|^2)^{\frac{c_i-1}{2}}.$$

With respect to such bases, the matrix of $D\gamma$ is block diagonal with the matrices of the $D\gamma^i$ as the blocks, so that

$$|\det D\gamma(p,(q,r))| = |\det D\gamma^1(p^1,(q^1,r^1))| \cdot \ldots \cdot |\det D\gamma^n(p^n,(q^n,r^n))|. \blacksquare$$

Lemma 6.12: *In addition to the hypotheses of the last result, suppose that, for each $i=1,\ldots,n$, E^i is an open subset of $\tilde{E}^i$. Let*

$$E = \{\,(q,r) \in \tilde{E} : (q^i,r^i) \in E^i \text{ for all } i=1,\ldots,n\,\},$$

and set $F := \prod_i F_i$ where

$$F^i = \{\,(p^i,q^i,r^i) \in \tilde{F}^i : (q^i,r^i) \neq (0,0) \text{ and } \frac{(q^i,r^i)}{\|(q^i,r^i)\|} \in E^i\,\}.$$

If $h : \tilde{D} \to I\!R$ is a bounded measurable function, then

$$\int_F h(p^1,\ldots,p^n)\,dF = \mathrm{vol}(E) \cdot \int_{\tilde{D}} h(p)\Big(\prod_{i=1}^n (1-\|p^i\|^2)^{\frac{c_i-1}{2}}\Big)\,dp.$$

Proof: Applying the last result, the change of variables formula gives

$$\int_F h(p^1,\ldots,p^n)\,dF = \int_{\tilde{D}\times E} h(p)|\det D\gamma\big(p,(q,r)\big)|\,d(\tilde{D}\times E)$$

$$= \int_{\tilde{D}\times E} h(p)\Big(\prod_{i=1}^n (1-\|p^i\|^2)^{\frac{c_i-1}{2}}\Big)\,d(\tilde{D}\times E)$$

and the asserted equation follows from Fubini's theorem. $\blacksquare$

The application of these results is as follows.

Lemma 6.13: *For each $i=1,\ldots,n$ let $\tilde{E}^i$ be the set of elements of the unit sphere in $I\!R^{S\setminus R^i}$ with*

$$u_i(t^0) = u_i(t^0|t_i^1) = \cdots = u_i(t^0|t_i^{p_i}),$$

and let

$$E^i := \{\,u_i \in \tilde{E}^i :\ u_i(t^0) > u_i(t^0|s_i) \text{ for all } s_i \in S_i \setminus T_i\ \}.$$

If $Y \subset N$ is open with $\mathrm{vol}(Y) > 0$, then

$$\frac{\nu(Y)}{\tilde{\nu}(Y)} = \frac{\mathrm{vol}(E)}{\mathrm{vol}(\tilde{E})} = \prod_{i=1}^n \frac{\mathrm{vol}(E^i)}{\mathrm{vol}(\tilde{E}^i)}.$$

Proof: Let $\tilde{D}^i$ be the unit disk in $I\!R^{R_i}$, and let $\tilde{F}_i$ be the set of all $u_i \in M_i$ satisfying this equation. To evaluate the numerator and denominator, respectively, of (12), we apply Lemma 6.12 twice, once as indicated by the notation and once with E^i replaced by $\tilde{E}^i$. $\blacksquare$

6.4. *Reinterpretation of the Function r*

Proposition 6.1 now follows from the result below, which interprets the function r as the fraction of the volume of a certain sphere (such as $\tilde{E}^i$ above) contained in the intersection (such as E^i above) of the sphere with a particular cone.

Lemma 6.14: *For nonnegative integers a, b, c with $a, b \geq 0$, and $c \geq a + b$,*

$$r(a, b) = \frac{\mathrm{vol}(\{\, x \in S^c : x_0 = x_1 = \cdots = x_a > \max\{x_{a+1}, \ldots, x_{a+b}\}\,\})}{\mathrm{vol}(\{\, x \in S^c : x_0 = x_1 = \cdots = x_a \,\})}.$$

Proof: Let $\mathbf{e}_0, \mathbf{e}_1, \ldots, \mathbf{e}_c$ be the unit basis vectors of $\mathbb{R}^{c+1}$. Let $\epsilon_0, \epsilon_{a+1}, \epsilon_{a+2}, \ldots, \epsilon_c$ be i.i.d. normal random variables, let

$$\tilde{x} = \epsilon_0 \cdot \frac{\mathbf{e}_0 + \mathbf{e}_1 + \cdots + \mathbf{e}_a}{\sqrt{a+1}} + \epsilon_{a+1} \cdot \mathbf{e}_{a+1} + \cdots + \epsilon_c \cdot \mathbf{e}_c,$$

and let $\tilde{y} = \tilde{x}/\|\tilde{x}\|$. By virtue of standard properties of the multivariate normal distribution, $\tilde{y}$ is a uniformly distributed random point in the sphere $\{\, x \in S^c : x_0 = x_1 = \cdots = x_a \,\}$. The ratio of volumes in the assertion is now seen to be the probability of the event $\epsilon_0/\sqrt{a+1} > \max\{\epsilon_{a+1}, \ldots \epsilon_{a+b}\}$. ∎

References

1. Y. Balasko, *Foundations of the Theory of General Equilibrium*, Academic Press, London, (1988).
2. B. D. Bernheim, Rationalizable strategic behavior, *Econometrica* **52** (1984), 1007–1028.
3. J. Berg and A. McLennan, The asymptotic expected number of Nash equilibria of two player normal form games, mimeo, University of Minnesota, (2002).
4. D. N. Bernstein, The number of roots of a system of equations, *Functional Analysis and its Applications* **9** (1975), 183–185.
5. L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and Real Computation*, Springer-Verlag, New York, (1998).
6. M. Dresher, Probability of a pure equilibrium point in n-person games, *Journal of Combinatorial Theory* **8** (1970), 134–145.
7. A. Edelman and E. Kostlan, How many zeros of a random polynomial are real?, *Bulletin of the American Mathematical Society* **32** (1995), 1–37.
8. I.Z. Emiris and J.F. Canny, Efficient incremental algorithms for the sparse resultant and the mixed volume, *Journal of Symbolic Computation* **20** (1995), 117–149.
9. G. Ewald, *Combinatorial Convexity and Algebraic Geometry*, Springer-Verlag, New York, (1996).
10. H. Federer, *Geometric Measure Theory*, Springer, New York, (1969).
11. D. Fudenberg and D. K. Levine, *The Theory of Learning in Games*, MIT Press, Cambridge, MA (1998).
12. P. Gritzman and V. Klee, Mathematical programming and convex geometry, *Handbook of Convex Geometry*, vol. I, P. M. Gruber and J. M. Wills (eds) (1993).
13. V. Guillemin and A. Pollack, *DifferentialTopology*, Prentice-Hall, Englewood Cliffs, (1965).
14. J. C. Harsanyi, Oddness of the number of equilibrium points: a new proof, *Int. J. Game Theory,* **2** (1973), 235–250.
15. B. Huber and B. Sturmfels, A polyhedral method for solving sparse polynomial systems, *Mathematics of Computation* **64** (1995), 1541–1555.
16. 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.
17. H. Keiding, On the maximal number of Nash equilibria in a bimatrix game. *Games and Economic Behavior* **21** (1995), 148-160.
18. E. Kohlberg and J.-F. Mertens, On the strategic stability of equilibria, *Econometrica* **54** (1986), 1003–1038.
19. E. Kostlan, On the distribution of roots of random polynomials, *From Topology to Computation: Proceedings of the Smalefest*, Hirsch, M., Marsden, J., and Shub, M. (eds) Springer-Verlag (1993).
20. E. Kostlan, On the expected number of real roots of a system of random polynomial equations, in *Foundations of Computational Mathematics: Proceedings of the Smalefest 2000*, Cucker, F., and Rojas, M. (eds) World Scientific (2002).
21. H. Kuhn, et. al., The work of John Nash in game theory: Nobel seminar, December 8, 1994, *Journal of Economic Theory* **69** (1996), 153–185.
22. R. D. McKelvey and A. McLennan, The maximal number of regular totally mixed Nash equilibria, *Journal of Economic Theory* **72** (1997), 411–425.

23. A. McLennan, The maximal generic number of pure Nash equilibria, *Journal of Economic Theory* **72** (1997), 408–410.
24. A. McLennan, The maximal number of real roots of a multihomogeneous system of polynomial equations, *Beiträge zur Algebra und Geometrie* **40**-2 (1999), 343-350.
25. A. McLennan, The expected number of real roots of a multihomogeneous system of polynomial equations, forthcoming in the *American Journal of Mathematics*, (2002).
26. A. McLennan and I.-U. Park, Generic 4x4 games have at most 15 Nash equilibria, *Games and Economic Behavior* **26** (1999), 111-130.
27. J.W. Milnor and J.D. Stasheff, *Characteristic Classes*, Annals of Mathematics Studies 76, Princeton University Press, Princeton (1974).
28. D. G. Pearce, Rationalizable strategic behavior and the problem of perfection, *Econometrica* **52** (1984), 1029–1050.
29. I. Powers, Limiting distributions of the number of pure strategy Nash equilibria, *International Journal of Game Theory* **19** (1990), 277–286.
30. T. Quint and M. Shubik, A theorem on the number of Nash equilibria in a bimatrix game, *International Journal of Game Theory* **26** (1997), 353–360.
31. 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.
32. H.L. Royden, *Real Analysis*, Third Edition, Macmillan, New York, 1988.
33. L. Samuelson, *Evolutionary Games and Equilibrium Selection*, MIT Press, Cambridge, MA 1997.
34. 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.
35. S. Smale, On the average number of steps of the simplex method of linear programming, *Mathematical Programming* **27** (1983), 241-262.
36. W. Stanford, A note on the probability of k pure Nash equilibria in matrix games, *Games and Economic Behavior* **9** (1995), 238-246.
37. J. Verschelde, P. Verlinden, and R. Cools, Homotopies exploiting Newton polytopes for solving sparse polynomial systems, *SIAM Journal of Numerical Analysis* **31** (1994), 915–930.
38. J. W. Weibull, *Evolutionary Game Theory*, MIT Press, Cambridge, MA (1995).

Appendix A. The Proof of Proposition 4.1

This Appendix reproduces the proof of Proposition 4.1 from McLennan (2002), with some minor modifications arising from differences in notational system and other adjustments to the current context.

Recall, from Section 4, how we defined an action of G on each $\mathcal{H}_k$ by setting $gf := f \circ g^{-1}$, and how $\langle f_k, f'_k \rangle_k := \sum_{a \in \mathcal{A}_k} \eta(a) f_{ka} f'_{ka}$ is the unique inner product that is both invariant under this action and has distinct monomials orthogonal. We denote the orthogonality relation derived from $\langle \cdot, \cdot \rangle_k$ by $\perp_k$. For each $k = 1, \ldots, p$ let $\theta_k : N \to \mathcal{H}_k$ be the function whose a-component, for $a \in \mathcal{A}_k$, is

$$\theta_{ka}(\tau) := \eta(a)^{-1} \tau^a.$$

Note that $f_k(\tau) = \langle f_k, \theta_k(\tau) \rangle_k$, so that $\tau \in N$ is a root of $f_k \in \mathcal{M}_k$ if and only if $f_k \perp_k \theta_k(\tau)$.

Lemma A.1: Each θ_k is G-equivariant with respect to the actions of G on N and $\mathcal{H}_k$. The image of θ_k is contained in the unit sphere of $\mathcal{H}_k$.

Proof: For fixed τ and g, $\theta_k(g\tau) = g\theta_k(\tau)$ follows from the fact that, for all $f_k \in \mathcal{H}_k$,

$$\langle gf_k, g\theta_k(\tau) \rangle_k = \langle f_k, \theta_k(\tau) \rangle_k = f_k(\tau) = f_k(g^{-1}(g\tau)) = \langle gf_k, \theta_k(g\tau) \rangle_k.$$

Here the first equality is the invariance of the inner product established in McLennan (2002), and the other three equalities are essentially matters of definition.

Clearly $\theta_k(\tau)$ is a standard basis vector of $\mathcal{H}_k$ if $\tau_1, \ldots, \tau_n$ are all standard basis vectors in $\mathbb{R}^{p_1+1}, \ldots, \mathbb{R}^{p_n+1}$ respectively. For such a τ the first claim implies that any vector of the form $\theta_k(g\tau)$ is contained in the unit sphere, so the second claim follows from the transitivity of the action of G on N. ■

Let $F : \mathcal{M} \times N \to \mathbb{R}^p$ be the *evaluation map* with components

$$F_k(f, \tau) := f_k(\tau) = \langle f_k, \theta_k(\tau) \rangle_k.$$

In this context the *incidence variety* is $V = F^{-1}(0)$.

Since $f_k(\tau) = 0$ if and only if $f_k \perp_k \theta_k(\tau)$, for $(f, \tau) \in V$ we may construe $\theta_k(\tau)$ as a tangent vector in $T_{f_k}\mathcal{M}_k$. Consider the effect on F of perturbing f_k in the direction of $\theta_k(\tau)$: since $||\theta_k(\tau)|| = 1$ the image of

$$((0, \ldots, \theta_k(\tau), \ldots, 0), 0) \in T_{(f,\tau)}(\mathcal{M} \times N)) \approx \big(\prod_k T_{f_k}\mathcal{M}_k\big) \times T_\tau N$$

under $DF(f, \tau)$ is the k^{th} unit basis vector of $\mathbb{R}^p$. Thus the image of $DF(f, \tau)$ is all of $\mathbb{R}^p$. Since (f, τ) was an arbitrary point in V, every point in V is a regular point of F, i.e., 0 is a regular value of F, so the regular value theorem (e.g., Guillemin and Pollack (1965)) implies:

Lemma A.2: V is a C^∞ submanifold of $\mathcal{M} \times N$ with $\dim V = \dim \mathcal{M}$.

Let π_1 and π_2 be the projections from V to $\mathcal{M}$ and N respectively. Abusing notation, we let V_τ denote both of the fibers

$$\pi_2^{-1}(\tau) \subset V \subset \mathcal{M} \times N \quad \text{and} \quad \{\, f \in \mathcal{M} : (f,\tau) \in \pi_2^{-1}(\tau)\,\}$$

over a point $\tau \in N$, with the appropriate interpretation to be inferred from context. For each k let $V_{\tau,k}$ be the set of $f_k \in \mathcal{M}_k$ with $f_k(\tau) = 0$. As the intersection of $\mathcal{M}_k$ with a hyperplane, this set is a subsphere of $\mathcal{M}_k$ of codimension one. Thus the topology of $V_\tau = V_{\tau,1} \times \cdots \times V_{\tau,p}$ is independent of τ.

Lemma A.3: $\pi_2 : V \to N$ is a C^∞ fibration.

As in the proof of Lemma 6.3, standard methods (e.g., Milnor and Stasheff (1974, §3)) can be used to prove that each $\{\, (f_k,\tau) \in \mathcal{H}_k \times N : f_k(\tau) = 0 \,\}$ is a C^∞ vector bundle over N, after which V is seen to be the product (in the sense of taking the cartesian product of the fibres over each base point) of the associated sphere bundles.

A.1 An Integral Formula

By definition

$$\nu(N) = \frac{1}{\mathrm{vol}(\mathcal{M})} \int_{\mathcal{M}} \#(\pi_1^{-1}(f))\, d\mathcal{M}. \tag{13}$$

Sard's theorem implies that almost all points of $\mathcal{M}$ are regular values of π_1, so we need only consider such points in computing the average number of roots. Consider a regular point (f,τ) of π_1. Since $T_{(f,\tau)}V$ is mapped surjectively onto $T_f\mathcal{M}$ by $D\pi_1(f,\tau)$, the restriction of $DF(f,\tau)$ to $T_\tau N \subset T_{(f,\tau)}(\mathcal{M} \times N)$ must be nonsingular, else (f,τ) would not be a regular point of F. The implicit function theorem implies that there is a neighborhood $U \subset \mathcal{M}$ of f for which there is a smooth $J : U \to N$ with $J(f) = \tau$ whose graph is contained in V. The *condition matrix* at (f,τ) is the matrix of $DJ(f)$ which, by the implicit function theorem, is

$$C(f,\tau) := -\Big(\frac{\partial F}{\partial \tau}(f,\tau)\Big)^{-1} \frac{\partial F}{\partial f}(f,\tau) : T_f\mathcal{M} \to T_\tau N.$$

This linear transformation gives a description of the way polynomial systems f are associated with their roots near (f,τ). Let $C^*(f,\tau) : T_\tau N \to T_f\mathcal{M}$ be the adjoint of $C(f,\tau)$.

Proposition A.4: (Blum, *et. al.* (1998, p. 240)) *For any open* $U \subset V$,

$$\int_{\mathcal{M}} \#(\pi_1^{-1}(f) \cap U)\, d\mathcal{M} = \int_N \int_{V_y \cap U} \det\big(C(f,\tau)C^*(f,\tau)\big)^{-1/2} dV_\tau dN.$$

Lemma A.5: If $(f,\tau) \in V$ is a regular point of π_1, then

$$\det\big(C(f,\tau)C^*(f,\tau)\big)^{-1/2} = |\det Df(\tau)|.$$

Proof: The adjoint $\frac{\partial F}{\partial f}(f,\tau)^*$ of $\frac{\partial F}{\partial f}(f,\tau)$ is the map $v \mapsto (v_1\theta_1(\tau), \ldots, v_p\theta_p(\tau))$. To see this we compute that, for $v \in T_0\mathbb{R}^p$ and $\phi \in T_f\mathcal{M} = T_{f_1}\mathcal{M}_1 \times \cdots \times T_{f_p}\mathcal{M}_p$,

$$\begin{aligned}\Big\langle \frac{\partial F}{\partial f}(f,\tau)\phi, v\Big\rangle &= \Big\langle \big(\langle\phi_1,\theta_1(\tau)\rangle_1, \ldots, \langle\phi_p,\theta_p(\tau)\rangle_p\big), v\Big\rangle \\ &= \langle \phi_1, v_1\theta_1(\tau)\rangle_1 + \cdots + \langle\phi_p, v_p\theta_p(\tau)\rangle_p \\ &= \Big\langle \phi, \big(v_1\theta_1(\tau), \ldots, v_p\theta_p(\tau)\big)\Big\rangle.\end{aligned}$$

Lemma A.1 implies that each $\theta_k(\tau)$ is a unit vector, so $\frac{\partial F}{\partial f}(f,\tau)^*$ is an isometric embedding of $\mathbb{R}^p$ in $\mathcal{H}$. In general, if V and W are inner product spaces and $i : V \to W$ is linear, then $(i^*)^* = i$, and if i is an isometric embedding, then i^* is the orthogonal projection onto $i(V)$ followed by i^{-1}, so that $i^* \circ i = \mathrm{Id}_V$. Thus $\frac{\partial F}{\partial f}(f,\tau)\frac{\partial F}{\partial f}(f,\tau)^*$ is the identity on $T_0\mathbb{R}^p$. Since the matrix of the adjoint of a linear transformation is the transpose of the transformation's matrix, substituting the definition of the condition matrix leads to

$$\begin{aligned}\det\big(C(f,\tau)C^*(f,\tau)\big)^{-1/2} &= \det\Big(\frac{\partial F}{\partial \tau}(f,\tau)^{-1}\frac{\partial F}{\partial f}(f,\tau)\frac{\partial F}{\partial f}(f,\tau)^*\big(\frac{\partial F}{\partial \tau}(f,\tau)^{-1}\big)^*\Big)^{-1/2} \\ &= \det\Big(\frac{\partial F}{\partial \tau}(f,\tau)^{-1}\big(\frac{\partial F}{\partial \tau}(f,\tau)^{-1}\big)^*\Big)^{-1/2} \\ &= \Big|\det\frac{\partial F}{\partial \tau}(f,\tau)\Big| = |\det Df(\tau)|. \ \blacksquare\end{aligned}$$

The last two results imply that for any open $U \subset V$ we have

$$\int_{\mathcal{M}} \#(\pi_1^{-1}(f)\cap U)\, d\mathcal{M} = \int_N \int_{V_\tau \cap U} |\det Df(\tau)|\, dV_\tau dN. \tag{14}$$

A.2 Invariance

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

$$gf := (f_1 \circ g^{-1}, \ldots, f_p \circ g^{-1}).$$

We will exploit this symmetry to further simplify the right hand side of (14). Each $\mathcal{M}_k$ is invariant under the action of G on $\mathcal{H}_k$ since the inner product of $\mathcal{H}_k$ is invariant. Thus $\mathcal{M}$ is invariant under the action of G on $\mathcal{H}$, and the restriction of this action to $\mathcal{M}$ is an action of G on $\mathcal{M}$. Of course N is invariant under the usual action of G on $\prod_{i=1}^n \mathbb{R}^{p_i+1}$. Combining these actions, we derive an action of G on $\mathcal{M} \times N$ given by $g(f,\tau) := (gf, g\tau)$. For any $g \in G$, $f \in \mathcal{M}$, and $\tau \in N$ we have $gf(g\tau) = f \circ g^{-1}(g\tau) = f(\tau)$, so:

Lemma A.6: The variety V is invariant under the action of G on $\mathcal{M} \times N$: $gV = V$ for all $g \in G$. Consequently $g(V_\tau) = V_{g\tau}$ for all τ and g.

Proposition A.7: The quantity $\int_{V_\tau} |\det Df(\tau)|\, dV_\tau$ is independent of τ.

Proof: Observe that

$$D(gf)(g\tau) = D(f \circ g^{-1})(g\tau) = Df(\tau) \circ g^{-1}$$

(the second equality is from the chain rule) so that $|\det D(gf)(g\tau)| = |\det Df(\tau)|$. We now have the calculation that

$$\int_{V_{g\tau}} |\det Df(g\tau)|\, dV_{g\tau} = \int_{V_\tau} |\det D(gf)(g\tau)|\, dV_\tau = \int_{V_\tau} |\det Df(\tau)|\, dV_\tau.$$

Here the first equality is an application of the change of variables formula to the change of variables $g : V_\tau \to V_{g\tau}$, which is an isometry, so that the Jacobian is identically one. The claim now follows from the fact that the action of G on N is transitive. ∎

Applying this to (14), for any open $Y \subset N$ and any $\tau \in N$ we have

$$\int_{\mathcal{M}} \#(\pi_1^{-1}(f) \cap \pi_2^{-1}(Y))\, d\mathcal{M} = \mathrm{vol}(Y) \cdot \int_{V_\tau} |\det Df(\tau)|\, dV_\tau. \tag{15}$$

Clearly (b) of Proposition 4.1 follows directly from this. The remaining task is to prove (a) of that result.

A.3 The Final Calculations

It is now convenient to introduce the model of a random multihomogeneous system in which the coefficient vectors of the various equations are statistically independent, with the coefficient vector of the k^{th} equation centrally normally distributed in $\mathcal{H}_k$ relative to $\langle \cdot, \cdot \rangle_k$. Concretely this means that the coefficients $\tilde{f}_{ka}$ are independent Gaussian random variables with mean 0 and variance $\eta(a)^{-1}$. Let μ_k be the probability measure on $\mathcal{H}_k$ that is the distribution of $\tilde{f}_k$, and let

$$\mu := \mu_1 \times \cdots \times \mu_p$$

be the distribution of $\tilde{f} := (\tilde{f}_1, \ldots, \tilde{f}_p)$.

Fixing $\tau \in N$, let $\tilde{f}_\tau = (\tilde{f}_{\tau,1}, \ldots, \tilde{f}_{\tau,p})$ be the orthogonal projection of $\tilde{f}$ onto the subspace of polynomial systems for which τ is a root. For each k, $\|\tilde{f}_{\tau,k}\|$ and $\tilde{f}_{\tau,k}/\|\tilde{f}_{\tau,k}\|$ are statistically independent, and the normalized vector is uniformly distributed in $V_{\tau,k}$, so

$$\begin{aligned}\int_{\mathcal{H}} |\det D\tilde{f}_\tau(\tau)|\, d\mu &= \int_{\mathcal{H}} \Big(\prod_{k=1}^{p} \|\tilde{f}_{\tau,k}\|\Big) \cdot \Big|\det D\Big(\frac{\tilde{f}_{\tau,1}}{\|\tilde{f}_{\tau,1}\|}, \ldots, \frac{\tilde{f}_{\tau,p}}{\|\tilde{f}_{\tau,p}\|}\Big)(\tau)\Big|\, d\mu \\ &= \Big(\prod_{k=1}^{p} \mathbf{E}\big(\|\tilde{f}_{\tau,k}\|\big)\Big) \frac{1}{\mathrm{vol}(V_\tau)} \int_{V_\tau} |\det Df(\tau)|\, dV_\tau.\end{aligned}$$

Combining this with (13) and (15), we now obtain

$$\nu(N) = \frac{\mathrm{vol}(N) \cdot \mathrm{vol}(V_\tau)}{\mathrm{vol}(\mathcal{M}) \cdot \prod_{k=1}^{p} \mathbf{E}\big(\|\tilde{f}_{\tau,k}\|\big)} \int_{\mathcal{H}} |\det D\tilde{f}_\tau(\tau)|\, d\mu. \tag{16}$$

Lemma A.8: Let $\tilde{\epsilon} = (\tilde{\epsilon}_1, \ldots, \tilde{\epsilon}_m)$ where $\tilde{\epsilon}_1, \ldots, \tilde{\epsilon}_m$ are independent identically distributed normal random variables with mean zero and unit variance. Then

$$\mathbf{E}(\|\tilde{\epsilon}\|) = \frac{\sqrt{2}\cdot\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})}. \tag{17}$$

Proof: We compute that

$$\begin{aligned}
\mathbf{E}(\|\tilde{\epsilon}\|) &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\|x\|\cdot(\frac{1}{\sqrt{2\pi}}e^{-x_1^2/2}\,dx_1)\cdots(\frac{1}{\sqrt{2\pi}}e^{-x_m^2/2}\,dx_m)\\
&= (2\pi)^{-m/2}\int_{I\!R^m}\|x\|\cdot e^{-\|x\|^2/2}\,dx\\
&= (2\pi)^{-m/2}\int_0^{\infty}(re^{-r^2/2})\cdot \mathrm{vol}(S^{m-1})\cdot r^{m-1}\,dr.
\end{aligned}$$

The change of variables $t := r^2/2$ gives

$$\int_0^{\infty} r^m e^{-r^2/2}\,dr = \int_0^{\infty}(2t)^{m/2}e^{-t}\frac{dt}{\sqrt{2t}} = 2^{\frac{m-1}{2}}\int_0^{\infty} t^{\frac{m-1}{2}}e^{-t}\,dt = 2^{\frac{m-1}{2}}\Gamma(\frac{m+1}{2}),$$

so the asserted formula is now obtained from the formula (7). ∎

The formula (7) for sphere volume implies

$$\mathrm{vol}(N_i) = 2\frac{\pi^{\frac{p_i+1}{2}}}{\Gamma(\frac{p_i+1}{2})},\quad \mathrm{vol}(\mathcal{M}_k) = 2\frac{\pi^{\frac{\dim\mathcal{H}_k}{2}}}{\Gamma(\frac{\dim\mathcal{H}_k}{2})},\quad \mathrm{vol}(V_{\tau,k}) = 2\frac{\pi^{\frac{\dim\mathcal{H}_k-1}{2}}}{\Gamma(\frac{\dim\mathcal{H}_k-1}{2})},$$

and Lemma 6.8 yields

$$\mathbf{E}(\|\tilde{f}_{\tau,k}\|) = \frac{\sqrt{2}\cdot\Gamma(\frac{\dim\mathcal{H}_k}{2})}{\Gamma(\frac{\dim\mathcal{H}_k-1}{2})}.$$

Since $\mathrm{vol}(\mathcal{M}) = \mathrm{vol}(\mathcal{M}_1)\times\cdots\times\mathrm{vol}(\mathcal{M}_p)$, and similarly for N and V_τ, substituting into (16) and simplifying yields

$$\nu(N) = 2^{n-p/2}\cdot\Big(\prod_{i=1}^{n}\frac{\sqrt{\pi}}{\Gamma(\frac{p_i+1}{2})}\Big)\cdot\int_{\mathcal{H}}|\det D\tilde{f}_\tau(\tau)|\,d\mu. \tag{18}$$

In the further evaluation of this quantity we are free to let τ be any convenient point in N. We will compute at $\tau_0 = (\mathbf{e}_{10},\ldots,\mathbf{e}_{n0}) \in N$ where, for $1 \le i \le n$, $\mathbf{e}_{i0}, \mathbf{e}_{i1}, \ldots, \mathbf{e}_{ip_i}$ are the standard unit basis vectors of $I\!R^{p_i+1}$. Note that $T_{\tau_0}N$ is spanned by the p vectors

$$\mathbf{b}_{ih} := (0,\ldots,\mathbf{e}_{ih},\ldots,0)\qquad(1\le i\le n, 1\le h\le p_i),$$

For each k let

$$a_k^0 = (a_{k1}^0,\ldots,a_{kn}^0)\in\mathcal{A}_k$$

where, for each i, $a_{ki}^0 = (d_{ki}, 0, \ldots, 0) \in \mathcal{A}_{ki}$. Then $\tau_0^{a_k^0} = 1$ and $\tau_0^a = 0$ for all $a \in \mathcal{A}_k \setminus \{a_k^0\}$, so the polynomials f_k having τ_0 as a root are those with $f_{ka_k^0} = 0$, and the map $f \mapsto f_{\tau_0}$ is the projection taking each $f_{ka_k^0}$ to zero. Elementary calculus yields

$$D\tilde{f}_{\tau_0,k}(\tau_0)\mathbf{b}_{ih} = \begin{cases} \tilde{f}_{ka_k^{ih}} & \text{if } d_{ki} > 0, \\ 0 & \text{if } d_{ki} = 0, \end{cases}$$

where, for each k and i such that $d_{ki} > 0$ and each $h = 1, \ldots, p_i$, a_k^{ih} is a_k^0 with a_{ki}^0 replaced by $(d_{ki} - 1, 0, \ldots, 0, 1, 0, \ldots, 0)$ (the '1' is at the h^{th} component). In this way we obtain a description of $D\tilde{f}_{\tau_0}(\tau_0)$ as an $p \times p$ matrix with rows indexed by $f_1, \ldots, f_p$, columns indexed by the pairs (i, h), and this (k, ih)–entry. Evaluating (8) with $a = a_k^{ih}$ shows that the variance of $\tilde{f}_{ka_k^{ih}}$ is $\eta(a_k^{ih})^{-1} = d_{ki}$, so the matrix of $D\tilde{f}_{\tau_0}(\tau_0)$ has the same distribution as $\tilde{Z}$. In view of (18) this observation completes the proof of Proposition 4.1.