COMPLEXITY OF BEZOUT'S THEOREM II
VOLUMES AND PROBABILITIES

Michael Shub* and Steve Smale**

May 7, 1992

Abstract. In this paper we study volume estimates in the space of systems of n homegeneous polynomial equations of fixed degrees d_i with respect to a natural Hermitian structure on the space of such systems invariant under the action of the unitary group. We show that the average number of real roots of real systems is $\mathcal{D}^{1/2}$ where $\mathcal{D} = \prod d_i$ is the Bezout number. We estimate the volume of the subspace of badly conditioned problems and show that volume is bounded by a small degree polynomial in n, N and $\mathcal{D}$ times the reciprocal of the condition number to the fourth power. Here N is the dimension of the space of systems.

Section 1. Introduction.

This paper can be read independently of **Shub-Smale** hereafter referred to as [I], but is closely related to it. Here we confine ourselves to homogeneous polynomials and projective spaces although some extensions to the affine case may be dealt with as in [I].

The paper [I] can be read for background and more references.

First consider a real polynomial system $f : \mathbb{R}^{n+1} \to \mathbb{R}^n$ so that $f(z) = (f_1(z_0, \ldots, z_n), \ldots, f_n(z_0, \ldots, z_n))$ and each f_i is a homogeneous polynomial of degree $d_i > 0$. Let $\mathcal{H}_{(d)}^{\mathbb{R}}$ be the linear space of all such f where $d = (d_1, \ldots, d_n)$ (permitting $f_i \equiv 0$).

There is a natural inner product on $\mathcal{H}_{(d)}^{\mathbb{R}}$ invariant under the induced action of the orthogonal group $O(n+1)$ acting on $\mathbb{R}^{n+1}$ (so that $\langle f \circ O, g \circ O \rangle = \langle f, g \rangle$ for $O \in O(n+1)$). See Section 2 for this. This inner product defines a Riemannian structure and volume element on the corresponding projective space $P(\mathcal{H}_{(d)}^{\mathbb{R}})$ of lines (Fubini-Study). In turn this volume element defines a probability measure so that the following makes sense.

Theorem A. *The average number of real zeros in $P_n(\mathbb{R})$ of $f \in P(\mathcal{H}_{(d)}^{\mathbb{R}})$ is $\mathcal{D}^{1/2}$ where $\mathcal{D} = \prod_{i=1}^n d_i$.*

According to Bezout's theorem, the (average) number of zeros over $\mathbb{C}$ is $\mathcal{D}$.

*Some of this work was carried out when Shub was visiting the Berkeley Math Department for 2 months in 1992,

**Supported partially by NSF funds .

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Eric **Kostlan-1991** proved this result earlier in the case that all of the d_i's are the same. In this paper Kostlan has a similar result for underdetermined systems, or gives average volumes of real varieties. Moreover, **Kostlan-1987** suggested using an orthogonally (or unitarily over $\mathbb{C}$) invariant metric in these complexity matters. Such a metric was already used in the theory of group representations and harmonic analysis (see e.g. **Stein-Weiss**).

In 1943, M. **Kac** had found in one variable, the expected number of real roots to be asymptotic to $\frac{2}{\pi}\log d$ using the traditional measure.

Next consider the complex case, $f : \mathbb{C}^{n+1} \to \mathbb{C}^n$, with each coordinate function f_i a homogeneous polynomial of degree d_i. Let $\mathcal{H}_{(d)}$ be the linear space of such f, and $P(\mathcal{H}_{(d)})$ the corresponding complex projective space. Define

$$V_{(d)} = \{(f,\zeta) \in P(\mathcal{H}_{(d)}) \times P(\mathbb{C}^{n+1}) \mid f(\zeta) = 0\}.$$

This complex non-singular subvariety of codimension n plays a central role in our work.

Inspired by **Wilkinson** we define the condition number $\mu : V_{(d)} \to \mathbb{R}^+ \cup \infty$ by

$$\mu(f,\zeta) = \|Df(\zeta)|_{N_\zeta}^{-1}\Delta(d_i^{1/2}\|\zeta\|^{d_i-1})\|\,\|f\|.$$

Here $\Delta(y_i)$ is the diagonal matrix with y_i as the (i,i) entry and

$$N_\zeta = \{w \in \mathbb{P}(\mathbb{C}^{n+1}) \mid \langle w,\zeta\rangle = 0\}.$$

The careful reader will have noted our customary practice of identifying objects in linear spaces and their quotient projective spaces. But appropriate homogenization gives sense to our definitions as $\mu(f,\zeta)$ above. If $Df(\zeta)|_{N_\zeta}$ is singular then $\mu(f,\zeta) = \infty$. In [I] μ was defined and called μ_{proj}.

Example. Let $e_0 = (1,0,\ldots,0)$ and $f_i(z) = z_0^{d_i-1}z_i$, $i = 1,\ldots,n$. We claim that $\mu(f,e_0) = \left(\sum_{i=1}^n \frac{1}{d_i}\right)^{1/2} D^{1/2}$, $D = \max_i d_i$, and thus (f,e_0) is a very well conditioned set of pairs varying over d.

Observe $N_{e_0} = \{(0,v_1,\ldots,v_n), v_i \in \mathbb{C}\}$ and that $Df(e_0)|_{N_{e_0}}$ is represented by the identity matrix. Moreover it is checked that $\|f\| = \left(\sum_{i=1}^n \frac{1}{d_i}\right)^{1/2}$. This yields our statement.

Typical numerical algorithms follow paths in $V_{(d)}$ and loss of precision due to round-off errors can be controlled by upper bounds on the corresponding condition number μ. Moreover μ plays a primary role in the (exact arithmetic) complexity analysis of such algorithms (see [I]). Thus it is important to understand the probability distribution of μ. We will show:

Theorem B. *If $n > 1$, and $D \geq 1$, then the probability that $\mu(f,\zeta) > \mu_0$ is less than $K\frac{nN}{\mu_0^2}$. More precisely*

$$\frac{\text{Vol}\{(f,\zeta) \in V_{(d)} \mid \mu(f,\zeta) \geq \mu_0\}}{\text{Vol}\, V_{(d)}} \leq K\frac{nN}{\mu_0^2}.$$

Here N is the dimension of $\mathcal{H}_{(d)}$.

The constant K is a universal constant less than 25. The case $n = 1$ will be dealt with in Theorem D.

Note that as a consequence of Theorem B, we see that most (f, ζ) in $V_{(d)}$ are well conditioned. The bound is independent of $\mathcal{D}$ and a low polynomial in n and N.

Numerical analysis has a useful tradition of relating the condition number to the distance ρ of the nearest ill-posed problem (see e.g. **Eckart-Young, Demmel**). In our setting the ill-conditioned pairs $\Sigma' \subset V_{(d)}$ are described by

$$\Sigma' = \{(f, \zeta) \in V_{(d)} \mid Df(\zeta)|_{N_\zeta} \text{ is singular}\}.$$

It can be shown that Σ' is a non-singular hypersurface in $V_{(d)}$ by a transversality argument. For $\zeta \in \mathbb{C}^{n+1}$ let $\hat{V}_\zeta = \{f \in \mathcal{H}_{(d)} \mid f(\zeta) = 0\}$ and $V_\zeta = P(\hat{V}_\zeta)$. Then V_ζ can be naturally identified with $\pi_2^{-1}(\zeta) \subset V_{(d)}$ where $\pi_2 : V_{(d)} \to P_n$ is the restriction of the projection $P(\mathcal{H}_{(d)}) \times P_n \to P_n$, and $P_n = P(\mathbb{C}^{n+1})$.

Define $\rho(f, \zeta)$ to be the distance in V_ζ of (f, ζ) to $\Sigma' \cap V_\zeta$. This projective space distance is taken for convenience to be $d_P = \sin d_R$ where d_R is the Riemannian (or Fubini-Study) distance. The diameter of projective space is one.

A main result of [I] is:

Condition Number Theorem.

$$\mu(f, \zeta) = \frac{1}{\rho(f, \zeta)}.$$

A sketch of the proof is given in Section 2.

Thus Theorem B may be interpreted as giving distribution estimates of pairs (f, ζ) close to ill-conditioned ones.

We now pass to the more subtle situation corresponding to all roots of a given $f \in P(\mathcal{H}_{(d)})$. Define the condition number $\mu(f)$ of $f \in P(\mathcal{H}_{(d)})$, $\mu : P(\mathcal{H}_{(d)}) \to \mathbb{R}^+ \cup \infty$ by

$$\mu(f) = \max_{\substack{\zeta \\ f(\zeta)=0}} \mu(f, \zeta).$$

Example. $n = 1$, $f_d(z) = z_0^d - z_1^d$. The zeros of f_d are $(1, q)$, whre q is a d^{th} root of unity. Take typically $q = 1$. Then $\|\zeta\| = \sqrt{2}$, $\zeta = (1, 1)$ and $\|f_d\| = \sqrt{2}$. Also $N_\zeta = \{(v, -v), v \in \mathbb{C}\}$, $\|Df_d(\zeta)|_{N_\zeta}^{-1}\| = \frac{1}{2d}$ and it follows that $\mu(f_d) = \frac{2^{d/2}}{4d^{1/2}}$.

Note that for this series f_d the condition number grows exponentially in d, and so is extremely ill-conditioned. Yet this f_d and its variations with exponential behavior are used in numerical analysis as a model starting polynomial whose roots are known. We will prove the existence of well-conditioned sequences $\{g_d\}$ (even for general $n \geq 1$), yet are unable to exhibit them even for the 1-variable case. In a further account we hope to develop this subject which in the case $n = 1$, is intimately connected to elliptic capacity and transfinite diameter (see **Tsuji**).

Defining $\rho(f) = \min_{\zeta, f(\zeta)=0} \rho(f,\zeta)$ we have:

Corollary of the Condition Number Theorem.

$$\mu(f) = \frac{1}{\rho(f)}.$$

Theorem C. *Let* $N_\rho = \{f \in P(\mathcal{H}_{(d)}) \mid \rho(f) < \rho\}$. *Then for* $\rho < \frac{1}{\sqrt{n}}$, $n > 1$, $D \geq 1$

$$\frac{\operatorname{Vol} N_\rho}{\operatorname{Vol} P(\mathcal{H}_{(d)})} \leq \frac{\rho^4 n^2 (n+1)(N-1)(N-2)\mathcal{D}}{4}$$

where $N = \dim \mathcal{H}_{(d)}$.

The closest previous results in the direction of Theorem 3 are due to **Renegar**.

Remark. The exponent 4 in Theorem C is somewhat surprising. Since $\Sigma \subset P(\mathcal{H}_{(d)})$ is a hypersurface, dimensional considerations say that the volume of the tube of radius ρ around Σ should vary like ρ^2.

Problem. Is $d(f,\Sigma) = O(\rho^2_{(f)})$?
The theorem says that on the average it is. There is one exceptional case where $d = (1,\ldots,1)$ in which case $\Sigma \subset P(\mathcal{H}_{(1,1,\ldots,1)})$ is of complex codimension 2. In this exceptional case the ρ^4 could be expected.

If we consider $x^2 - 2\varepsilon xy = f(x,y)$ $f \in \mathcal{H}_{(2)}$, $x^2 - 2\varepsilon xy + \varepsilon^2 y^2 \in \Sigma$ but $\rho(f) = O(\varepsilon)$.

Corollary. *For* $n > 1$ *and each* $d = (d_1, \ldots, d_n)$ *there is* $f_{(d)} \in P(\mathcal{H}_{(d)})$ *such that* $\mu(f_{(d)}) \leq \left(\frac{n^2(n+1)(N-1)(N-2)\mathcal{D}}{4}\right)^{1/4}$.

This is a consequence of Theorem 3 and the Corollary of the Condition Number Theorem.

As we have suggested, it is an interesting open problem to exhibit such $f_{(d)}$. These could serve as better starting points of numerical algorithms than those currently used.

The probabilistic estimates given in Theorems B and C, together with the results of [I] have implications on the efficiency of algorithms for solving non-linear systems. We hope to develop this point in a future paper.

In Theorem D we give the results of Theorem B and C for the case $n = 1$.

Theorem D. *For* $n = 1$ *and* $d > 1$

i) $$\frac{\operatorname{Vol}\{(f,\xi) \in V_{(d)} | \mu(f,\xi) \geq \mu_0 \geq 1\}}{\operatorname{Vol} V_{(d)}} = 1 - \frac{(d+1)}{2}\left(1 - \frac{1}{\mu_0^2}\right)^{d-1} + \frac{(d-1)}{2}\left(1 - \frac{1}{\mu_0^2}\right)^{d}$$

ii) $$\frac{\operatorname{Vol} N_\rho}{\operatorname{Vol} P(\mathcal{H}_{(d)})} \leq d(1 - (1-\rho^2)^{d-1}(1 + (d-1)\rho^2)) \quad \text{for } 0 \leq \rho \leq 1.$$

Section 2. Background.

The following diagram motivates the more abstract treatment of the next section.

$$\begin{array}{ccccc} & & P(\mathcal{H}_{(d)}) \times P_n & & \\ & & \bigcup & & \\ & \pi_1 & V_{(d)} & \pi_2 & \\ & \swarrow & & \searrow & \\ P(\mathcal{H}_{(d)}) & & & & P_n \end{array}$$

The map $\pi_1 : V_{(d)} \to P(\mathcal{H}_{(d)})$ is a branched covering, branched along Σ' so that on $V_{(d)} - \pi_1^{-1}(\Sigma)$, where $\Sigma = \pi_1(\Sigma')$, π_1 is a covering map. The fiber $\pi_1^{-1}(f)$, $f \in P(\mathcal{H}_{(d)})$, $f \notin \Sigma$ consists of $\mathcal{D}$ points, corresponding to the zeros of f.

Moreover, $\pi_2 : V_{(d)} \to P_n$ is a fiber map with fiber $\pi_2^{-1}(\zeta) = V_\zeta$ over $\zeta \in P_n$.

The Unitary Group $U(n+1)$ is the group of linear automorphisms of $\mathbb{C}^{n+1}$ preserving the standard Hermitian inner product. This induces an action $x \to ux$, $u \in U(n+1)$ on P_n as well as the action $f \to fu^{-1}$ on $P(\mathcal{H}_{(d)})$. Moreover $U(n+1)$ acts on $P(\mathcal{H}_{(d)}) \times P_n$ by $(f,x) \to (fu^{-1}, ux)$ leaving $V_{(d)}$ invariant. The fiber map $\pi_2 : V_{(d)} \to P_n$ is also invariant under these actions.

The invariant norm on $\mathcal{H}_{(d)}$ is given by $\|f\|^2 = \sum_i \sum_\alpha |a^i_\alpha|^2 \begin{pmatrix} d_i \\ \alpha_1, \ldots, \alpha_n \end{pmatrix}^{-1}$ where $f^i_\alpha(z) = \sum_\alpha a^i_\alpha z^\alpha$, $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a multi-index and $\begin{pmatrix} d_i \\ \alpha_1, \ldots, \alpha_n \end{pmatrix}$ is the multinomial coefficient, $\frac{d_i!}{\alpha_1! \ldots \alpha_n!}$.

Here is a sketch of the proof of the condition number theorem. The condition number $\mu : V_{(d)} \to \mathbb{R}$ is invariant under $U(n+1)$ using the chain rule (as in Lemma 1 of Section III-1 of [I]).

The variety $\Sigma' \subset V_{(d)}$ is unitarily invariant as well as the distance in V_ζ. This implies that $\rho : V_{(d)} \to \mathbb{R}$ is invariant under $U(n+1)$. Given $(f,x) \in V_{(d)}$ pick $u \in U(n+1)$ with $ux = e_0 = (1, 0, \ldots, 0)$. Then

$$\rho(f,x) = \rho(fu^{-1}, e_0), \text{ and } \frac{1}{\mu(fu^{-1}, e_0)} = \frac{1}{\mu(f,x)}.$$

Thus it is sufficient to prove the condition number theorm for $x = e_0$.

For $f \in \mathcal{H}_{(d)}$, write $f = (f_1, \ldots, f_n)$

$$(*) \qquad f_i(z) = a_i z_0^{d_i} + z_0^{d_i - 1} \Sigma a_{ij} z_j + \ldots$$

If $f \in \hat{V}_{e_0}$ then the a_i are all zero, and note thus that $K = \{f \in \mathcal{H}_{(d)} \mid f_i(z) = a_i z_0^{d_i}\}$ is the orthogonal space to $\hat{V}_{e_0}$. Let $L(d) = \{f \in \mathcal{H}_{(d)} \mid f_i(z) = z_0^{d_i - 1} \Sigma a_{ij} z_j\}$, $J(d) =$ the orthogonal complement of $L(d)$ in $\hat{V}_{e_0}$, and $\pi : V_{e_0} \to L(d)$ the projection.

Note that $\mathrm{Null}(e_0) = \{u \in \mathbb{C}^{n+1} \mid \langle u, e_0 \rangle = 0\}$ is $\{(0, u_1, \ldots, u_n) \in \mathbb{C}^{n+1}\}$ and can be identified with $\mathbb{C}^n$. In this way, $Df(e_0)$ may be regarded as a linear map from $\mathbb{C}^n \to \mathbb{C}^n$ and as a matrix. This is the matrix (a_{ij}) when f has the form of $(*)$.

Let $\mathcal{M}(n)$ be the space of $n \times n$ matrices endowed with the Frobenius norm $\|M\|^2 = \Sigma|m_{ij}|^2$. Recall that $\Delta(y_i)$ denotes the diagonal matrix with entries $(y_1, \ldots, y_n)$.

Lemma. *The map $L(d) \to \mathcal{M}(n)$ sending $f \to \Delta(d_i^{-1/2})Df(e_0)$ is a norm preserving linear isomorphism.*

The proof follows from the definition of the norm on $\mathcal{H}_{(d)} \supset L(d)$.

Now suppose $\|f\| = 1$. Then $\mu(f, e_0) = \|Df(e_0)^{-1}\Delta d_i^{1/2}\|$ using the operator norm. By the theorem of **Eckardt-Young** (see [I]), $\mu(f, e_0) = d(\Delta(d_i^{-1/2})Df(e_0), S)^{-1}$ in $\mathcal{M}(n)$ where S is the set of singular matrices. It follows using the previous lemma that $\mu(f, e_0)$ is $\frac{1}{d(f, \Sigma' \cap \hat{V}_{e_0})}$ in $\mathcal{H}_{(d)}$. In the projective space this translates into our $\frac{1}{\rho(f, e_0)}$ completing the sketch of our proof.

Remark 1. In [I] it was part of the definition that $\mu(f, \zeta) \geq 1$. In fact it is a consequence of the original definition that $\mu(f, \zeta) \geq \sqrt{n}$. One uses the condition number theorem and the fact that $\rho \leq \sqrt{n}$ for any any matrix on the unit sphere in $\mathcal{M}(n)$.

Remark 2. There is a real version of this section where the unitary group is replaced by the orthogonal group, $O(n+1)$, acting on $\mathbb{R}^{n+1}$, etc.

We end this section by stating some standard facts that we use in our proofs. The volume $\mathrm{Vol}(S^{n-1})$ of the unit $(n-1)$-sphere S^{n-1} is $\frac{2\pi^{n/2}}{\Gamma(n/2)}$ using the gamma function. Let $P_n(\mathbb{R})$ be real projective space of dimension n and P_n complex projective space of (complex) dimension n. Then

$$\mathrm{Vol}\, P_n(\mathbb{R}) = \frac{1}{2}\mathrm{Vol}(S^n)$$
$$\mathrm{Vol}\, P_n = \frac{1}{2\pi}\mathrm{Vol}(S^{2n+1}).$$

We also use the following integral formula:

$$\int_0^1 t^p(1-t^2)^q dt = \frac{1}{2}\frac{\Gamma\left(\frac{p+1}{2}\right)\Gamma(q+1)}{\Gamma\left(\frac{p+1}{2}+q+1\right)}.$$

Section 3. Some General Integral Formulae.

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

Theorem 1.

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

Proof.

$$\int_{\pi_1 U} \#(x) dM = \int_U |\det(D\pi_1)| dV = \int_N \int_{V_y \cap U} |\det(D\pi_1)| \frac{1}{N_J \pi_2} dV_y dN.$$

Here $N_J\pi_2 = |\det(D\pi_2 \mid T(V_y)^\perp)|$ is the normal Jacobian and $T(V_y)^\perp$ is the orthogonal complement of TV_y in $TV(x,y)$. Here the first equality is a version of the usual change of variable formula for integrals. The second equality is the coarea formula (**Morgan**).

By Sard's Theorem it is sufficient to show that

$$(*) \qquad |\det(D\pi_1)| \frac{1}{N_J \pi_2} = (\det A^* A)^{1/2}$$

where both sides are evaluated at $(x,y) \in M \times N$ and x is a regular value of $\pi_1 : V \to M$.

Let H_1 and H_2 be finite dimensional real vector spaces (complex vector spaces) with inner product (Hermitian product). Let $A : H_1 \to H_2$ be linear and define the graph of A as $\Gamma(A) = \{(x, A(x)) \mid x \in H_1\}$. Let $\pi : \Gamma(A) \to H_1$ be the restriction of the projection. Let $\Gamma(A)$ inherit the inner product structure of the product.

Lemma 1.

$$|\det \pi| = \frac{1}{\det(I + A^*A)^{1/2}}$$

where A^* is the adjoint of A.

Proof. Let $\begin{pmatrix} I \\ A \end{pmatrix} : H_1 \to H_1 \times H_2$ be the map $x \to (x, Ax)$. There is an orthogonal (unitary in the complex case) automorphism) of H_1 such that

$$O\left(\left(\begin{pmatrix} I \\ A \end{pmatrix}^* \begin{pmatrix} I \\ A \end{pmatrix}\right)^{1/2}\right) = \begin{pmatrix} I \\ A \end{pmatrix}$$

and hence $|\det \begin{pmatrix} I \\ A \end{pmatrix}| = \det(I + A^*A)^{1/2}$. Also $\operatorname{Det} \pi = \frac{1}{\operatorname{Det}\begin{pmatrix} I \\ A \end{pmatrix}}$ since $\pi \circ \begin{pmatrix} I \\ A \end{pmatrix} = I$.

These two equalities give the proof.

Lemma 2. *Let $B(x,y) : TM(x) \to TN(y)$ be the linear map whose graph $\Gamma(B(x,y))$ is $TV(x,y)$ for a regular value x of π_1. Then*

$$\det(I + B^*(x,y)B(x,y)) = \det(I + A^{-1*}(x,y)A^{-1}(x,y)).$$

Proof. $TV_y(x,y)$ is contained in the $TM(x)$ factor of $TM(x) \times TN(y)$. As $TV(x,y)$ is the orthogonal direct sum of $TV_y(x,y)$ and $\Gamma(A(x,y))$, it follows that $TV(x,y)$ is the graph of the linear map $B : TM(x) \to TN(y)$ which is zero on $TV_y(x,y)$ and $A^{-1}(x,y)$ on $(D\pi_1(x,y))(\Gamma(A(x,y)))$ which in turn is orthogonal to $TV_y(x,y)$ in $TM(x)$. Consequently $\det(I + B^*(x,y)B(x,y)) = \det(I + A^{-1*}(x,y)A^{-1}(x,y))$.

Now we return to the proof of $(*)$ and the theorem. By Lemmas 1 and 2, using $A = A(x,y)$,

$$|\det(D\pi_1)(x,y)| = \det(I + A^{-1*}A^{-1})^{-1/2}.$$

By Lemma 1 and the definition of A,

$$\frac{1}{N_J\pi_2(x,y)} = \det(I + A^*A)^{1/2}.$$

Thus

$$|\det(D\pi_1(x,y))|\frac{1}{N_J\pi_2(x,y)} = \frac{\det(I + A^*A)^{1/2}}{\det(I + A^{-1*}A^{-1})^{1/2}}.$$

The next lemma finishes the proof.

Lemma 3. *Let $A : V \to W$ be a linear isomorphism of finite dimensional vector spaces with inner product, then*

$$\frac{\det(I + A^*A)}{\det(I + A^{-1*}A^{-1})} = \det(A^*A).$$

Proof. Multiply numerator and denominator by $\det(AA^*)$

$$\begin{aligned}\frac{\det(I + A^*A)}{\det(I + A^{*-1}A^{-1})} &= \frac{\det(AA^*)\det(I + A^*A)}{\det(AA^*)\det(I + A^{*-1}A^{-1})} \\ &= \frac{\det(AA^*)\det(I + A^*A)}{\det(AA^* + I)}.\end{aligned}$$

But then $\det(AA^* + I) = \det(I + A^*A)$ since AA^* and A^*A have the same eigenvalues.

Let U, V, M, π_1, π_2 etc. be as in Theorem 2.

Theorem 2.

$$\operatorname{Vol} U = \int_N \int_{V_y \cap U} \det(I + A^* A)^{1/2} dV_y dN.$$

Proof.

$$\operatorname{Vol} U = \int_U 1 dV = \int_N \int_{V_y \cap U} \frac{1}{N_J \pi_2} dV_y dN$$
$$= \int_N \int_{V_y \cap U} \det(I + A^* A)^{1/2} DV_y dN$$

by Lemma 1 and the definition of A.

Remark. The complex versions of Theorems 1 and 2 are true with the same proof. Then the exponents of $\frac{1}{2}$ on $\det(I + A^* A)^{1/2}$, $(\det A^* A)^{1/2}$ musts be removed as we pass to the real determinant. That is $|\text{real determinant}| = |\text{determinant}|^2$ for a complex matrix.

Section 4. Integration Formulae in $V_{(d)}$.

We use the notations introduced in Section 1, 2 and specialize the complex versions of Theorems 1 and 2 of Section 3 (cf. the remark at the end of Section 3) with $M = P(\mathcal{H}_{(d)})$ and $N = P_n$. Section 2 provides the background for the following.

Theorem 1. *Let U be an open set in $V_{(d)}$ which is unitarily invariant, $e_0 = (1, 0, \ldots, 0) \in P_n$ and $\#(f)$ be the number of points in $\pi_1^{-1}(f) \cap U$ (i.e. the number of zeros of f in U). Then*

(a) $\operatorname{Vol} U = \operatorname{Vol} P(n) \int_{V_{e_0} \cap U} \det(I + Df(e_0^*) Df(e_0))$

(b) $\int_{\pi_1 U} \#(f) = \operatorname{Vol} P(n) \int_{V_{e_0} \cap U} \det(Df(e_0)^* Df(e_0))$.

We will use:

Proposition. $\det A^*(f, x) A(f, x)$ *is invariant under the unitary group acting on $V_{(d)}$, and*

$$(*) \qquad A^*(f, e_0) A(f, e_0) = Df(e_0)^* Df(e_0).$$

Postponing the proof of the proposition for the moment, we will prove Theorem 1(b). Use Theorem 1 of the previous section and both parts of this proposition to obtain

$$\int_{\pi_1 U} \#(f) = \int_{P_n} \int_{V_{e_0} \cap U} \det(Df(e_0)^* Df(e_0))$$
$$= \operatorname{Vol} P_n \int_{V_{e_0} \cap U} \det(Df(e_0)^* Df(e_0)).$$

The proof of Theorem 1(a) is similar.

Since $\pi_2 : V_{(d)} \to P_n$ is unitarily invariant, so is the orthogonal complement of $TV_x(f,x)$ in $TV_{(d)}(f,x)$ where $V_x = \pi_2^{-1}(x)$. Then by definition $A(f,x)$ transforms by unitary compositions and thus $\det A^*(f,x)A(f,x)$ is unitarily invariant. This proves the first part of the proposition.

Working in the corresponding vector spaces, write as in Section 2, $\mathcal{H}_{(d)} = K + L(d) + J(d)$. Also

$$T(V_{(d)})(f,e_0) = \{(h,w) \mid Df(e_0)w = h(e_0)\}$$

so that $A : \mathbb{C}^n \to K \subset T_f(P(\mathcal{H}_{(d)}))$ is characterized by $Aw = h \in K$ and $h(e_0) = Df(e_0)(w)$. Here we are using the notation of Section 2 and the definition of $A = A(f,e_0)$. Then

$$A(w)_i = h_i = \sum_j a_{ij} w_j z_0^{d_i}$$

recalling $Df(e_0)(w) = \sum_j a_{ij} w_j$ and the Hermitian structure on K. Then $A^*A = Df(e_0)^* Df(e_0)$.

Section 5. Proof of Theorem A.

We start with:

Proposition 1. *Let $\pi : S_1^l \to D^k$ be orthogonal projection of the unit sphere $S_1^l \subset \mathbb{R}^{l+1}$ on the unit disk $D^k \subset R^k$ a subspace of $\mathbb{R}^l$, $0 < k < l$. Let $\phi : D^k \to \mathbb{R}$ be continuous and $U \subset D^k$ be open then*

$$\int_{S^l} (\phi \circ \pi) \mathcal{X}(\pi^{-1}(U)) = \int_{D^k} \mathcal{X}(U)(1 - \|x\|^2)^{\frac{l-k-1}{2}} \phi(x) \operatorname{Vol} S_1^{l-k}.$$

For the proof we use the following lemma.

Lemma. *The normal Jacobian of $\pi : S^l \to D^k$ at $x \in S^l$ is $(1 - \|\pi(x)\|^2)^{1/2}$.*

Proof. Let $\|\pi(x)\| = r$. Let S_r^{k-1} be the sphere of radius r about 0 in $\mathbb{R}^k$. Then $\pi^{-1}(S_r^{k-1}) = S_r^{k-1} \times S^{l-k}_{\sqrt{1-r^2}}$ which leaves only one normal direction to $\pi^{-1}(\pi(x))$, namely the one which maps to the day in D^k. Thus the problem reduces to $S^1 \subset \mathbb{R}^2$ and the norm of the projection is easily seen to be $\sqrt{1-r^2}$.

Proof of Proposition 1.

$$\begin{aligned}\int_{S^N}(\phi\circ\pi)\mathcal{X}(\pi^{-1}(U)) &= \int_{\pi^{-1}(U)}\phi\circ\pi\\ &= \int_U\int_{\pi^{-1}(x)}(\phi\circ\pi)(\pi^{-1}(x))\frac{1}{NJ(x)}\\ &= \int_U\int_{S_1^{l-k}}\phi(x)(1-\|x\|^2)^{\frac{l-k}{2}}\frac{1}{(1-\|x\|^2)^{1/2}}\\ &= \operatorname{Vol}S_1^{l-k}\int_U(1-\|x\|^2)^{\frac{l-k-1}{2}}\phi(x)\\ &= \operatorname{Vol}S_1^{l-k}\int_{D^k}\mathcal{X}(U)(1-\|x\|^2)^{\frac{l-k-1}{2}}\phi(x).\end{aligned}$$

Let $d=(d_1,\ldots,d_n)$, $\mathcal{D}=\pi_1^n d_i$. Then

$$A_d=\frac{\int_{P(\mathcal{H}_{(d)}^{\mathbb{R}})}\#(f)}{\operatorname{Vol}P(\mathcal{H}_{(d)}^{\mathbb{R}})}$$

is by definition the average number of real roots. We will prove that $A_d=\mathcal{D}^{1/2}$.

To prove this we will show

Lemma 1. *$A_d=\mathcal{D}^{1/2}G(n)$ where G is a function of n.*

Apply the lemma to the case $d=(d_1,\ldots,d_n)=(1,\ldots,1)$. Then clearly $A_d=1$ and $\mathcal{D}=1$. Therefore $G(n)$ is identically 1, and $A_d=\mathcal{D}^{1/2}$.

Thus it is sufficient to prove Lemma 1.

For this we use the real analogues of results of the previous sections, with orthogonal invariance replacing unitary invariance. Thus the real version of Theorem 1(b) of Section 4 gives (with $U=P(\mathcal{H}_{(d)}^{\mathbb{R}})$):

Lemma 2.

$$A_d=\frac{1}{2}\frac{\operatorname{Vol}P_n(\mathbb{R})}{\operatorname{Vol}P(\mathcal{H}_{(d)}^{\mathbb{R}})}\int_{S_{e_0}}(\det Df(e_0)^*Df(e_0))^{1/2}.$$

Here $e_0=(1,0,\ldots,0)$, $Df(e_0):\mathbb{R}^n\to\mathbb{R}^n$ is the derivative and $Df(e_0)^*$ its adjoint. Moreover S_{e_0} is the unit sphere in $\hat{V}_{e_0}^{\mathbb{R}}=\{f\in\mathcal{H}_{(d)}^{\mathbb{R}}\mid f(e_0)=0\}$ and $V_{e_0}=P(\hat{V}_{e_0}^{\mathbb{R}})$ has $\frac{1}{2}$ the volume of S_{e_0}. Let $N=\dim\mathcal{H}_{(d)}^{\mathbb{R}}$ so that $\dim S_{e_0}=N-n-1$.

Lemma 3.

$$\int_{S_{e_0}}(\det Df(e_0)^*Df(e_0))^{1/2}=\mathcal{D}^{1/2}\frac{\pi^{N/2}}{\Gamma\left(\frac{N}{2}\right)}H(n)$$

where $H(n)$ depends only on n.

Since $\operatorname{Vol}P(\mathcal{H}_{(d)}^{\mathbb{R}})=\frac{1}{2}\operatorname{Vol}(S^{N-1})=\frac{\pi^{N/2}}{\Gamma(N/2)}$, Lemmas 2 and 3 yield that $A_d=\mathcal{D}^{1/2}\operatorname{Vol}P_n(\mathbb{R})H(n)$ proving Lemma 1. Thus it remains to prove only Lemma 3.