ELSEVIER

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Theoretical Computer Science 315 (2004) 525–555

Theoretical Computer Science

www.elsevier.com/locate/tcs

High probability analysis of the condition number of sparse polynomial systems

Gregorio Malajovich[a,*,1], J. Maurice Rojas[b,2]

[a] Departamento de Matemática Aplicada, Universidade Federal do Rio de Janeiro, Caixa Postal 68530, CEP 21945-970, Rio de Janeiro, RJ, Brazil

[b] Department of Mathematics, Texas A&M University, TAMU 3368, College Station, TX 77843-3368, USA

Abstract

Let $f := (f^1, \ldots, f^n)$ be a random polynomial system with fixed n-tuple of supports. Our main result is an upper bound on the probability that the condition number of f in a region U is larger than $1/\varepsilon$. The bound depends on an integral of a differential form on a toric manifold and admits a simple explicit upper bound when the Newton polytopes (and underlying variances) are all identical.

We also consider polynomials with real coefficients and give bounds for the expected number of real roots and (restricted) condition number. Using a Kähler geometric framework throughout, we also express the expected number of roots of f inside a region U as the integral over U of a certain *mixed volume* form, thus recovering the classical mixed volume when $U = (\mathbb{C}^*)^n$.

Keywords: Mixed volume; Condition number; Polynomial systems; Sparse; Random

1. Introduction

From the point of view of numerical analysis, it is not only the number of complex solutions of a polynomial system which make it hard to solve numerically but the sensitivity of its roots to small perturbations in the coefficients. This is formalized in

* Corresponding author. Tel.: 5521-2562-7515; fax: 5521-2290-1095.

E-mail addresses: gregorio@ufrj.br (G. Malajovich), rojas@math.tamu.edu (J.M. Rojas).

URLs: http://www.labma.ufrj.br/~gregorio, http://www.math.tamu.edu/~rojas

[1] Partially supported by CNPq grant 303472/02-2, by CERG (Hong Kong) grants 9040393, 9040402, and 9040188, by FAPERJ and by Fundação José Pelúcio Ferreira.

[2] Partially supported by Hong Kong UGC grant #9040402-730, Hong Kong/France PROCORE Grant #9050140-730, and NSF Grants DMS-0138446 and DMS-0211458.

0304-3975/$ - see front matter
doi:10.1016/j.tcs.2004.01.006

the condition number, $\mu(f,\zeta)$ (cf. Definition 4 of Section 1.1), which dates back to work of Alan Turing [39]. In essence, $\mu(f,\zeta)$ measures the sensitivity of a solution ζ to perturbations in a problem f, and a large condition number is meant to imply that f is intrinsically hard to solve numerically. Such analysis of numerical conditioning, while having been applied for decades in numerical linear algebra (see, e.g., [11]), has only been applied to computational algebraic geometry toward the end of the twentieth century (see, e.g., [33]).

Here we use Kähler geometry to analyze the numerical conditioning of sparse polynomial systems, thus setting the stage for more realistic complexity bounds for the numerical solution of polynomial systems. Our bounds generalize some earlier results of Kostlan [20] and Shub and Smale [36] on the more restricted dense case, and also yield new formulae for the expected number of roots (real and complex) in a region. The appellations "sparse" and "dense" respectively refer to either (a) taking into account the underlying monomial term structure or (b) ignoring this finer structure and simply working with degrees of polynomials. Since many polynomial systems occurring in practice have rather restricted monomial term structure, sparsity is an important consideration and we therefore strive to state our complexity bounds in terms of this refined information.

To give the flavor of our results, let us first make some necessary definitions. We must first formalize the spaces of polynomial systems we work with and how we measure perturbations in the spaces of problems and solutions.

Definition 1. Given any finite subset $A \subset \mathbb{Z}^n$, let $\mathscr{F}_{\mathbb{C}}(A)$ (resp. $\mathscr{F}_{\mathbb{R}}(A)$) denote the vector space of all polynomials in $\mathbb{C}[x_1,\ldots,x_n]$ (resp. $\mathbb{R}[x_1,\ldots,x_n]$) of the form $\sum_{a\in A} c_a x^a$ where the notation $x^a := x_1^{a_1}\cdots x_n^{a_n}$ is understood. For any finite subsets $A_1,\ldots,A_n \subset \mathbb{Z}^n$ we then let $\mathscr{A} := (A_1,\ldots,A_n)$ and $\mathscr{F}_{\mathbb{C}}(\mathscr{A}) := \mathscr{F}_{\mathbb{C}}(A_1)\times\cdots\times\mathscr{F}_{\mathbb{C}}(A_n)$ (resp. $\mathscr{F}_{\mathbb{R}}(\mathscr{A}) := \mathscr{F}_{\mathbb{R}}(A_1)\times\cdots\times\mathscr{F}_{\mathbb{R}}(A_n)$).

The n-tuple $\mathscr{A}$ will thus govern our notion of sparsity as well as the perturbations allowed in the coefficients of our polynomial systems. It is then easy to speak of random polynomial systems and the distance to the nearest degenerate system. Recall that a *degenerate* root of f is simply a root of f having Jacobian of rank $<n$.

Definition 2. By a *complex* (*resp. real*) *random sparse polynomial system* we will mean a choice of $\mathscr{A} := (A_1,\ldots,A_n)$ and an assignment of a probability measure to each $\mathscr{F}_{\mathbb{C}}(A_i)$ (resp. $\mathscr{F}_{\mathbb{R}}(A_i)$) as follows: endow $\mathscr{F}_{\mathbb{C}}(A_i)$ (resp. $\mathscr{F}_{\mathbb{R}}(A_i)$) with an independent complex (resp. real) Gaussian distribution having mean $\mathbf{O}$ and a (positive definite and diagonal) variance matrix C_i. Finally, let the *discriminant variety*, $\Sigma(\mathscr{A})$, denote the set of all $f \in \mathscr{F}_{\mathbb{C}}(\mathscr{A})$ (resp. $f \in \mathscr{F}_{\mathbb{R}}(\mathscr{A})$) with a degenerate root and define $\mathscr{F}_{\zeta}(\mathscr{A}) := \{f \in \mathscr{F}_{\mathbb{C}}(\mathscr{A}) \mid f(\zeta) = \mathbf{O}\}$ (resp. $\mathscr{F}_{\zeta}(\mathscr{A}) := \{f \in \mathscr{F}_{\mathbb{R}}(\mathscr{A}) \mid f(\zeta) = \mathbf{O}\}$) and $\Sigma_{\zeta}(\mathscr{A}) := \mathscr{F}_{\zeta}(\mathscr{A}) \cap \Sigma(\mathscr{A})$.

Theorem 1. *Suppose $A \subset \mathbb{Z}^n$ is a finite set with a convex hull of positive volume and $\mathscr{A} := (\underbrace{A,\ldots,A}_{n})$. Then there is a natural metric $d(\cdot,\cdot)$ on $\mathscr{F}_{\mathbb{C}}(\mathscr{A})$ such that*

$\mu(f,\zeta)=1/d(f,\Sigma_\zeta(\mathscr{A}))$. *Furthermore*,

$$\text{Prob}\left[\mu(f,\zeta) \geqslant \frac{1}{\varepsilon} \text{ for some root } \zeta \in (\mathbb{C}^*)^n \text{ of } f\right] \leqslant n^3(n+1)\,\text{Vol}(A)(\#A-1)(\#A-2)\varepsilon^4,$$

where f *is a complex random sparse polynomial system*, $\#A$ *denotes the number of points in* A, *and* $\text{Vol}(A)$ *denotes the volume of the convex hull of* A (*normalized so that* $\text{Vol}(\mathbf{O},e_1,\dots,e_n)=1$).

The above theorem is in fact a simple corollary of two much more general theorems (Theorems 4 and 5) which also include as a special case an analogous result of Shub and Smale in the dense case [6, Theorem 1, p. 237]. We also note that theorems such as the one above are natural precursors to explicit bounds on the number of steps required for a homotopy algorithm [33] to solve f. We will pursue the latter topic in a future paper. Indeed, one of our long term goals is to provide a rigorous and explicit complexity analysis of the numerical homotopy algorithms for sparse polynomial systems developed by Verschelde et al. [40], Huber and Sturmfels [17], and Li and Li [21].

The framework underlying our first main theorem involves Kähler geometry, which is the intersection of Riemannian metrics and symplectic and complex structures on manifolds. On a more concrete level, we can give new formulae for the expected number of roots of f in a region U. For technical reasons, we will mainly work with *logarithmic coordinates*. That is, we will let $\mathscr{T}^n$ be the n-fold product of cylinders $(\mathbb{R}\times(\mathbb{R} \bmod 2\pi))^n \subset \mathbb{C}^n$, and use coordinates $p+\mathrm{i}q := (p_1+\mathrm{i}q_1,\dots,p_n+\mathrm{i}q_n) \in \mathscr{T}^n$ to stand for a root $\zeta := \exp(p+\mathrm{i}q) := (\mathrm{e}^{p_1+\mathrm{i}q_1},\dots,\mathrm{e}^{p_n+\mathrm{i}q_n})$ of f. Roots with zero coordinates can be handled by then working in a suitable *toric compactification* and this is made precise in Section 2. The idea of working with roots of polynomial systems in logarithmic coordinates seems to be extremely classical, yet it gives rise to interesting and surprising connections (see the discussions in [24,25,41]).

Theorem 2. *Let* $A_1,\dots,A_n$ *be finite subsets of* $\mathbb{Z}^n$ *and* $U\subseteq\mathscr{T}^n$ *be a measurable region. Pick positive definite diagonal variance matrices* $C_1,\dots,C_n$ *and consider a complex random polynomial system as in Definition* 2, *for some* $(A_1,C_1,\dots,A_n,C_n)$. *Then there are natural real* 2-*forms* $\omega_{A_1},\dots,\omega_{A_n}$ *on* $\mathscr{T}^n$ *such that the expected number of roots of* f *in* $\exp U\subseteq(\mathbb{C}^*)^n$ *is exactly*

$$\frac{(-1)^{n(n-1)/2}}{\pi^n}\int_U \omega_{A_1}\wedge\cdots\wedge\omega_{A_n}.$$

In particular, *when* $U=(\mathbb{C}^*)^n$, *the above expression is exactly the mixed volume of the convex hulls of* $A_1,\dots,A_n$ (*normalized so that the mixed volume of n standard n-simplices is* 1).

See [7,31] for the classical definition of mixed volume and its main properties. The result above generalizes the famous connection between root counting and mixed volumes discovered by David N. Bernshtein [5]. The special case of unmixed systems with identical coefficient distributions ($A_1 = \cdots = A_n$, $C_1 = \cdots = C_n$) recovers a particular case of Theorem 8.1 in [12]. However, comparing Theorem 2 and [12, Theorem 8.1], this is the only overlap since neither theorem generalizes the other. The very last assertion of Theorem 2 (for uniform variance $C_i = I$ for all i) was certainly known to Gromov [14], and a version of Theorem 2 was known to Kazarnovskii [18, p. 351] and Khovanskii [19, Proposition 1, Section 1.13]. In [18], the supports A_i are even allowed to have complex exponents. However, uniform variance is again assumed. His method may imply this special case of Theorem 2, but the indications given in [18] were insufficient for us to reconstruct a proof. Also, there is some intersection with a result by Passare and Rullgård (Theorem 5 in [29] and Theorem 20 in [28]). However, this result is about a more restrictive choice of the domain U and a more general class of functions (holomorphic, not polynomials) under a different averaging process.

As a consequence of our last result, we can also give a coarse estimate on the expected number of real roots in a region.

Theorem 3. *Let U be a measurable subset of $\mathbb{R}^n$ with Lebesgue volume $\lambda(U)$. Then, following the notation above, suppose instead that f is a real random polynomial system. Then the average number of real roots of f in $\exp U \subset \mathbb{R}^n_+$ is bounded above by*

$$(4\pi^2)^{-n/2}\sqrt{\lambda(U)}\sqrt{\int_{(p,q)\in U\times[0,2\pi)^n}(-1)^{n(n-1)/2}\omega_{A_1}\wedge\cdots\wedge\omega_{A_n}}.$$

This bound is of interest when n and U are fixed, in which case the expected number of positive real roots grows as the square root of the mixed volume.

1.1. Stronger results via mixed metrics

Our remaining new results, which further sharpen the preceding bounds and formulae, will require some additional notation.

Definition 3. We define a norm on $\mathscr{F}_{\mathbb{C}}(A_i)$ by $\|f^i\|_{C_i^{-1}} := c^i C^{-1}(c^i)^H$ where c_i is the row vector of coefficients of f_i and $(\cdot)^H$ denotes the usual Hermitian conjugate transpose. Finally, we define a norm on $\mathscr{F}_{\mathbb{C}}(\mathscr{A})$ by $\|f\|^2 := \sum_{i=1}^n \|f^i\|^2_{C_i^{-1}}$, and a metric $d_{\mathbb{P}}$ on the product of projective spaces $\mathbb{P}(\mathscr{F}_{\mathbb{C}}(\mathscr{A})) := \mathbb{P}(\mathscr{F}_{\mathbb{C}}(A_1)) \times \cdots \times \mathbb{P}(\mathscr{F}_{\mathbb{C}}(A_n))$ by $d_{\mathbb{P}}(f,g) := \sum_{i=1}^n \min_{\lambda\in\mathbb{C}^*} \|f^i - \lambda g^i\|/\|f^i\|$, where we implicitly use the natural embedding of $\mathbb{P}(\mathscr{F}_{\mathbb{C}}(A_i))$ into the unit hemisphere of $\mathscr{F}_{\mathbb{C}}(A_i)$.

Each of the terms in the sum above corresponds to the square of the sine of the Fubini (or angular) distance between f^i and g^i. Therefore, $d_{\mathbb{P}}$ is never larger than the Hermitian distance between points in $\mathscr{F}_{\mathbb{C}}(\mathscr{A})$, but is a correct first-order approximation of the distance when $g \to f$ in $\mathbb{P}(\mathscr{F}_{\mathbb{C}}(\mathscr{A}))$ (compare with [6, Chapter 12]).

Recall that T_pM denotes the tangent space at p of a manifold M.

Definition 4. Define the *evaluation map*, $ev_{\mathscr{A}}$, as follows:

$$\begin{aligned} ev_{\mathscr{A}}: \quad & \mathscr{F} \times \mathscr{T}^n \to \mathbb{C}^n \\ & ((f^1,\ldots,f^n), p+\mathrm{i}q) \mapsto (f^1(\exp(p+\mathrm{i}q)),\ldots,f^n(\exp(p+\mathrm{i}q))). \end{aligned}$$

Given any root $\exp(p+\mathrm{i}q)$ of an f in $\mathscr{F}_{\mathbb{C}}(\mathscr{A})$, the *condition number of* f *at* $p+\mathrm{i}q$, $\boldsymbol{\mu}(f,p+\mathrm{i}q)$, is then defined to be the operator norm

$$\|DG|_f\| := \max_{\|g\|=1} \|DG|_f\|,$$

where G is the unique branch of the implicit function which satisfies $G(f)=p+\mathrm{i}q$ and $ev_{\mathscr{A}}(g,G(g))=\mathbf{O}$ for all g sufficiently near f, and $DG: T_f\mathscr{F}_{\mathbb{C}}(\mathscr{A}) \to T_{p+\mathrm{i}q}\mathscr{T}^n$ is the derivative of G. (We set the condition number $\mu(f,p+\mathrm{i}q) := +\infty$ in the event that Df does not have full rank and G thus fails to be uniquely defined.)

Note that the implied norm on $T_f\mathscr{F}_{\mathbb{C}}(\mathscr{A})$ was detailed in the previous definition, while the implied norm on $T_{p+\mathrm{i}q}\mathscr{T}^n$ has intentionally been left unspecified. This is because while $\mathscr{F}_{\mathbb{C}}(\mathscr{A})$ admits a natural Hermitian structure, the solution-space $\mathscr{T}^n$ admits n different natural Hermitian structures (one from each support A_i, as we shall see in the next section). Nevertheless, we can give useful bounds on the condition number and give an unambiguous definition in certain cases.

Theorem 4 (Condition Number Theorem). *If* $(p,q) \in \mathscr{T}^n$ *is a non-degenerate root of* f *then*

$$\max_{\|\dot f\| \leqslant 1} \min_i \|DG_f \dot f\|_{A_i} \leqslant \frac{1}{d_{\mathbb{P}}(f, \Sigma_{(p,q)})} \leqslant \max_{\|\dot f\| \leqslant 1} \max_i \|DG_f \dot f\|_{A_i}.$$

In particular, if $A_1 = \cdots = A_n$ *and* $C_1 = \cdots = C_n$, *then*

$$\max_{\|\dot f\| \leqslant 1} \min_i \|DG_f \dot f\|_{A_i} = \max_i \max_{\|\dot f\| \leqslant 1} \|DG_f \dot f\|_{A_i} = \frac{1}{d_{\mathbb{P}}(f, \Sigma_{(p,q)})}$$

and we can define $\boldsymbol{\mu}(f;(p,q))$ *to be any of the three preceding quantities.*

This generalizes [6, Theorem 3, p. 234] which is essentially equivalent to the last assertion above, in the special case where A_i is an n-column matrix whose rows $\{A_i^\alpha\}_\alpha$ consist of all partitions of d_i into n non-negative integers and $C_i = \mathrm{Diag}_\alpha((d_i-1)!/(A_i)_1^\alpha!(A_i)_2^\alpha!\cdots(A_i)_n^\alpha!(d_i - \sum_{j=1}^n (A_i)_j^\alpha)!)$—in short, the case where one considers complex random polynomial systems with f^i a degree d_i polynomial and the underlying probability measure is invariant under a natural action of the unitary group $U(n+1)$ on the space of roots. The last assertion of Theorem 4 also bears some similarity to Theorem D of [9] where the notion of metric is considerably loosened to give a statement which applies to an even more general class of equations. However, our philosophy is radically different: we consider the inner product in $\mathscr{F}_{\mathbb{C}}(\mathscr{A})$ as the starting point of our investigation and we do *not* change the metric in the fiber $\mathscr{F}_{(p,q)}$.

Theorem 4 thus gives us some insight about reasonable intrinsic metric structures on $\mathscr{T}^n$.

In view of the preceding theorem, we can define a restricted condition number with respect to any measurable sub-region $U \subset \mathscr{T}^n$ as follows:

Definition 5. We let $\boldsymbol{\mu}(f;U) := 1/\min_{(p,q)\in U} d_{\mathbb{P}}(f,\Sigma_{(p,q)})$. Also, via the natural $GL(n)$-action on $T_{(p,q)}\mathscr{T}^n$ defined by $(\dot{p},\dot{q}) \mapsto (L\dot{p},L\dot{q})$ for any $L \in GL(n)$, we define the *mixed dilation* of the tuple $(\omega_{A_1},\ldots,\omega_{A_n})$ as

$$\kappa(\omega_{A_1},\ldots,\omega_{A_n};(p,q)) := \min_{L\in GL(n)} \max_i \frac{\max_{\|u\|=1} (\omega_{A_i})_{(p,q)}(Lu,JLu)}{\min_{\|u\|=1} (\omega_{A_i})_{(p,q)}(Lu,JLu)},$$

where $J : T\mathscr{T}^n \to T\mathscr{T}^n$ is canonical complex structure of $\mathscr{T}^n$. Finally, we define $\kappa_U := \sup_{(p,q)\in U} \kappa(\omega_{A_1},\ldots,\omega_{A_n};(p,q))$, provided the supremum exists, and $\kappa_U := +\infty$ otherwise.

We can then bound the expected number of roots with condition number $\boldsymbol{\mu} > \varepsilon^{-1}$ on U in terms of the mixed volume form, the mixed dilation κ_U and the expected number of ill-conditioned roots in the *linear case*. The linear case corresponds to the point sets and variance matrices below:

$$A_i^{\mathrm{Lin}} = \begin{bmatrix} 0 & \cdots & 0 \\ 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} \quad C_i^{\mathrm{Lin}} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}.$$

Theorem 5 (Expected value of the condition number). *Let $v^{\mathrm{Lin}}(n,\varepsilon)$ be the probability that a complex random system of n polynomial in n variables has condition number larger than ε^{-1}. Let $v^A(U,\varepsilon)$ be the probability that $\boldsymbol{\mu}(f,U) > \varepsilon^{-1}$ for a complex random polynomial system f with supports $A_1,\ldots,A_n$ and variance matrices $C_1,\ldots,C_n$. Then*

$$v^A(U,\varepsilon) \leqslant \frac{\int_U \bigwedge \omega_{A_i}}{\int_U \bigwedge \omega_{A_i^{\mathrm{Lin}}}} v^{\mathrm{Lin}}(n,\sqrt{\kappa_U}\,\varepsilon).$$

Our final main result concerns the distribution of the real roots of a real random polynomial system. Let $v_{\mathbb{R}}(n,\varepsilon)$ be the probability that a real random linear system of n polynomials in n variables has condition number larger than ε^{-1}.

Theorem 6. *Let $A = A_1 = \cdots = A_n$ and $C = C_1 = \cdots = C_n$ and let $U \subseteq \mathbb{R}^n$ be measurable. Let f be a real random polynomial system. Then,*

$$\mathrm{Prob}[\boldsymbol{\mu}(f,U) > \varepsilon^{-1}] \leqslant E(U)v_{\mathbb{R}}(n,\varepsilon),$$

where $E(U)$ is the expected number of real roots on U.

Note that $E(U)$ depends on C, so even if we make $U = \mathbb{R}^n$ we may still obtain a bound depending on C. Shub and Smale showed in [32] that the expected number of real roots in the dense case (with a particular choice of probability measure) is exactly the square root of the expected number of complex roots. The sparse analogue of this result seems hard to prove even in the general unmixed case: Explicit formulæ for the unmixed case are known only in certain special cases, e.g., certain systems of bounded multi-degree [30,27]. Hence our last theorem can be interpreted as another step toward a fuller generalization.

2. Symplectic geometry and polynomial systems

2.1. Some basic definitions and examples

For the standard definitions and properties of symplectic structures, complex structures, Riemannian manifolds, and Kähler manifolds, we refer the reader to [26,8]. A treatment focusing on toric manifolds can be found in [15, Appendix A]. We briefly review a few of the basics before moving on to the proofs of our theorems.

Definition 6 (Kähler manifolds). Let M be a complex manifold, with complex structure J and a strictly positive symplectic $(1,1)$-form ω on M (considered as a real manifold). We then call the triple (M, ω, J) a *Kähler manifold.*

Example 1 (Affine space). We identify $\mathbb{C}^M$ with $\mathbb{R}^{2M}$ and use coordinates $Z^i = X^i + \sqrt{-1}Y^i$. The *canonical* 2-form $\omega_Z = \sum_{i=1}^M \mathrm{d}X_i \wedge \mathrm{d}Y_i$ makes $\mathbb{C}^M$ into a symplectic manifold.

The natural complex structure J is just the multiplication by $\sqrt{-1}$. The triple $(\mathbb{C}^M, \omega_Z, J)$ is a Kähler manifold.

Example 2 (Projective space). Projective space $\mathbb{P}^{M-1}$ admits a *canonical* 2-form defined as follows. Let $Z = (Z^1, \ldots, Z^M) \in (\mathbb{C}^M)^*$, and let $[Z] = (Z^1 : \cdots : Z^M) \in \mathbb{P}^{M-1}$ be the corresponding point in $\mathbb{P}^{M-1}$. The tangent space $T_{[Z]}\mathbb{P}^{M-1}$ may be modeled by $Z^\perp \subset T_Z\mathbb{C}^M$. Then we can define a two-form on $\mathbb{P}^{M-1}$ by setting:

$$\omega_{[Z]}(u, v) = \|Z\|^{-2}\omega_Z(u, v),$$

where it is assumed that u and v are orthogonal to Z. The latter assumption tends to be quite inconvenient, and most people prefer to pull $\omega_{[Z]}$ back to $\mathbb{C}^M$ by the canonical projection $\pi : Z \mapsto [Z]$. It is standard to write the pull-back $\tau = \pi^*\omega_{[Z]}$ as

$$\tau_Z = -\tfrac{1}{2}\,\mathrm{d}J^* d\,\tfrac{1}{2}\,\log\,\|Z\|^2,$$

using the notation $\mathrm{d}\eta = \sum_i \partial\eta/p_i \wedge \mathrm{d}p_i + \partial\eta/q_i \wedge \mathrm{d}q_i$, and where J^* denotes the pull-back by J.

Projective space also inherits the complex structure from $\mathbb{C}^M$. Then $\omega_{[Z]}$ is a strictly positive $(1,1)$-form. The corresponding metric is called *Fubini-Study* metric in $\mathbb{C}^M$ or $\mathbb{C}^{M-1}$.

Remark 1. Some authors prefer to write $\sqrt{-1}\partial\bar{\partial}$ instead of $-\frac{1}{2}\,\mathrm{d}J^*d$. The following notation is assumed: $\partial\eta = \sum_i \partial\eta/Z_i \wedge \mathrm{d}Z_i$ and $\bar{\partial}\eta = \sum_i \partial\eta/\bar{Z}_i \wedge \mathrm{d}\bar{Z}_i$. Then they write τ_Z as

$$\tau_Z = \frac{\sqrt{-1}}{2}\left(\frac{\sum_i \mathrm{d}Z_i \wedge \mathrm{d}\bar{Z}_i}{\|Z\|^2} - \frac{\sum_i Z_i\,\mathrm{d}\bar{Z}_i \wedge \sum_j \bar{Z}_j\,\mathrm{d}Z_j}{\|Z\|^4}\right).$$

Example 3 (Toric Kähler manifolds from point sets). Let A be any $M \times n$ matrix with integer entries whose row vectors have n-dimensional convex hull and let C be any diagonal positive definite M times M matrix. Define the map $\hat{V}_A$ from $\mathbb{C}^n$ into $\mathbb{C}^M$ by

$$\hat{V}_A : z \mapsto C^{1/2}\begin{bmatrix} z^{A^1} \\ \vdots \\ z^{A^M} \end{bmatrix}.$$

We can also compose with the projection into projective space to obtain a slightly different map $V_A = \pi \circ \hat{V}_A : \mathbb{C}^n \to \mathbb{P}^{M-1}$ defined by $V_A : z \mapsto [\hat{V}_A(z)]$. When C is the identity, the Zariski closure of the image of V_A is called the *Veronese variety* and the map V_A is called the *Veronese embedding*. Note that V_A is not defined for certain values of z, like $z=0$. Those values comprise the *exceptional set* which is a subset of the coordinate hyper-planes.

There is then a natural symplectic structure on the closure of the image of V_A, given by the restriction of the Fubini-Study 2-form τ: We will see below (Lemma 1) that by our assumption on the convex hull of the rows of A, we have that DV_A is of rank n for $z \in (\mathbb{C}^*)^n$. Thus, we can pull-back this structure to $(\mathbb{C}^*)^n$ by $\Omega_A = V_A^*\tau$. Also, we can pull back the complex structure of $\mathbb{P}^{M-1}$, so that Ω_A becomes a strictly positive $(1,1)$-form. Therefore, the matrix A defines a Kähler manifold $((\mathbb{C}^*)^n, \Omega_A, J)$.

The reason we introduced C in the definition of $\hat{V}_A$ is as follows: if f denotes also the row-vector of the *scaled* coefficients of f, then $f(z) = \sum_a f_a (C_a)^{+1/2} z^a = f\hat{V}_A(z)$. This way, the 2-norm of the row vector f is also the norm of the polynomial f in $(F_A, \|\cdot\|_{C^{-1}}\|)$. A random normal polynomial with variance matrix C corresponds to a random normal row vector f with unit variance.

Example 4 (Toric manifolds in logarithmic coordinates). For any matrix A as in the previous example, we can pull-back the Kähler structure of $((\mathbb{C}^*)^n, \Omega_A, J)$ to obtain another Kähler manifold $(\mathscr{T}^n, \omega_A, J)$. (Actually, it is the same object in logarithmic coordinates, minus points at "infinity".) An equivalent definition is to pull back the Kähler structure of the Veronese variety by $\hat{v}_A \stackrel{\text{def}}{=} \hat{V}_A \circ \exp$.

Remark 2. The Fubini-Study metric on $\mathbb{C}^M$ was constructed by applying the operator $-\frac{1}{2}\,\mathrm{d}J^*d$ to a certain convex function (in our case, $\frac{1}{2}\log\|Z\|^2$). This is a general standard way to construct Kähler structures. In [14], it is explained how to associate a (non-unique) convex function to any convex body, thus producing an associated Kähler metric.

For the record, we state explicit formulæ for several of the invariants associated to the Kähler manifold $(\mathscr{T}^n, \omega_A, J)$. First of all, the function $g_A = g \circ \hat{v}_A$ is precisely:

Formula 2.1.1. The canonical Integral g_A (or *Kähler potential*) of the convex set associated to A

$$g_A(p) := \tfrac{1}{2}\log((\exp(A\cdot p))^{\mathrm{T}} C(\exp(A\cdot p))).$$

The terminology *integral* is borrowed from mechanics, and it refers to the invariance of g_A under a $[0, 2\Pi)^n$-action. Also, the gradient of g_A is called the *momentum map*. Recall that the Veronese embedding takes values in projective space. We will use the following notation: $v_A(p) = \hat{v}_A(p)/\|\hat{v}_A(p)\|$. This is independent of the representative of equivalence class $v_A(p)$. Now, let $v_A(p)^2$ mean coordinate-wise squaring and $v_A(p)^{2\mathrm{T}}$ be the transpose of $v_A(p)^2$. The gradient of g_A is then:

Formula 2.1.2. The Momentum Map associated to A

$$\nabla g_A = v_A(p)^{2\mathrm{T}} A.$$

Formula 2.1.3. Second derivative of g_A

$$D^2 g_A = 2Dv_A(p)^{\mathrm{T}} Dv_A(p).$$

We also have the following formulae:

Formula 2.1.4. The symplectic 2-form associated to A:

$$(\omega_A)_{(p,q)} = \frac{1}{2}\sum_{ij}(D^2 g_A)_{ij}\,\mathrm{d}p_i \wedge \mathrm{d}q_j.$$

Formula 2.1.5. Hermitian structure of $\mathscr{T}^n$ associated to A:

$$(\langle u, w\rangle_A)_{(p,q)} = u^H(\tfrac{1}{2}D^2 g_A)_p w.$$

In general, the function v_A goes from $\mathscr{T}^n$ into projective space. Therefore, its derivative is a mapping

$$(Dv_A)_{(p,q)} : T_{(p,q)}\mathscr{T}^n \to T_{v_{A(p+q\sqrt{-1})}}\mathbb{P}^{M-1} \simeq \hat{v}_A(p+q\sqrt{-1})^{\perp} \subset \mathbb{C}^M.$$

For convenience, we will write this derivative as a mapping into $\mathbb{C}^M$, with range $\hat{v}_A(p+q\sqrt{-1})^{\perp}$. Let P_v be the projection operator

$$P_v = I - \frac{1}{\|v\|^2} vv^H.$$

We then have the following formula.

Formula 2.1.6. Derivative of v_A

$$(Dv_A)_{(p,q)} = P_{\hat{v}_A(p+q\sqrt{-1})}\,\mathrm{Diag}\left(\frac{\hat{v}_A(p+q\sqrt{-1})}{\|\hat{v}_A(p+q\sqrt{-1})\|}\right)A.$$

Lemma 1. *Let A be a matrix with non-negative integer entries, such that* Conv(A) *has dimension n. Then $(Dv_A)_p$ (resp. $(Dv_A)_{p+\mathrm{i}q}$) is injective, for all $p\in\mathbb{R}^n$ (resp. for all $p+\mathrm{i}q\in\mathbb{C}^n$).*

Proof. We prove only the real case (the complex case is analogous). The conclusion of this Lemma can fail only if there are $p\in\mathbb{R}^n$ and $u\neq 0$ with $(Dv_A)_p u=0$. This means that

$$P_{v_A(p)}\,\mathrm{diag}(v_A)_p A u = 0.$$

This can only happen if $\mathrm{diag}(v_A)_p Au$ is in the space spanned by $(v_A)_p$, or, equivalently, Au is in the space spanned by $(1,1,\ldots,1)^{\mathrm{T}}$. This means that all the rows a of A satisfy $au=\lambda$ for some λ. Interpreting a row of A as a vertex of Conv(A), this means that Conv(A) is contained in the affine plane $\{a : au=\lambda\}$. □

An immediate consequence of Formula 2.1.6 is

Lemma 2. *Let $f\in\mathscr{F}_A$ and $(p,q)\in\mathscr{T}^n$ be such that $f\cdot\hat{v}_A(p+q\sqrt{-1})=0$. Then,*

$$f\cdot(Dv_A)_{(p,q)} = \frac{1}{\|\hat{v}_A(p,q)\|}\,f\cdot(D\hat{v}_A)_{(p,q)}.$$

In other words, when $(f\circ\exp)(p+q\sqrt{-1})$ vanishes, Dv_A and $D\hat{v}_A$ are the same up to scaling. Noting that the Hermitian metric can be written $(\langle u,w\rangle_A)_{(p,q)} = u^h Dv_A(p,q)^H Dv_A(p,q)w$, we also obtain the following formula.

Formula 2.1.7. Volume element of $(\mathscr{T}^n,\omega_A,J)$

$$\mathrm{d}\mathscr{T}^n_A = \det(\tfrac{1}{2}D^2 g_A(p))\mathrm{d}p_1\wedge\cdots\wedge\mathrm{d}p_n\wedge\mathrm{d}q_1\wedge\cdots\wedge\mathrm{d}q_n.$$

2.2. Toric actions and the momentum map

The *momentum map*, also called *moment map*, was introduced in its modern formulation by Smale [37] and Souriau [38]. The reader may consult one of the many textbooks in the subject (such as Abraham and Marsden [1] or McDuff and Salamon [26]) for a general exposition (see also the discussion at the end of [23]).

In this section we instead follow the point of view of Gromov [14]. The main results in this section are the two propositions below.

Proposition 1. *The momentum map ∇g_A maps $\mathscr{T}^n$ onto the interior of* Conv(A). *When ∇g_A is restricted to the real n-plane $[q=0]\subset\mathscr{T}^n$, this mapping is a bijection.*

This would appear to be a particular case of the Atiyah–Guillemin–Sternberg theorem [2,16]. However, technical difficulties prevent us from directly applying this result here. [3]

Proposition 2. *The momentum map* ∇g_A *is a volume-preserving map from the manifold* $(\mathscr{T}^n, \omega_A, J)$ *into* $\mathrm{Conv}(A)$, *up to a constant, in the following sense: if* U *is a measurable region of* $\mathrm{Conv}(A)$, *then*

$$\mathrm{Vol}((\nabla g_A)^{-1}(U)) = \pi^n \,\mathrm{Vol}(U).$$

Proof of Proposition 2. Consider the mapping

$$M: \quad \mathscr{T}^n \to \tfrac{1}{2}\,\mathrm{Conv}(A) \times \mathbb{T}^n$$
$$(p,q) \mapsto (\tfrac{1}{2}\nabla g_A(p), q).$$

Since we assume $\dim \mathrm{Conv}(A) = n$, we can apply Proposition 1 and conclude that M is a diffeomorphism.

The pull-back of the canonical symplectic structure in $\mathbb{R}^{2n}$ by M is precisely ω_A, because of Formulæ 2.1.3 and 2.1.4. Diffeomorphisms with that property are called *symplectomorphisms*. Since the volume form of a symplectic manifold depends only of the canonical 2-form, symplectomorphisms preserve volume. We compose with a scaling by $\frac{1}{2}$ in the first n variables, that divides $\mathrm{Vol}(U)$ by 2^n, and we are done. □

Before proving Proposition 1, we will need the following result about convexity which has been attributed to Legendre. (See also [14, Convexity Theorem 1.2] and a generalization in [4, Theorem 5.1].)

Legendre's Theorem. *If* f *is convex and of class* $\mathscr{C}^2$ *on* $\mathbb{R}^n$, *then the closure of the image* $\{\nabla f_r : r \in \mathbb{R}^n\}$ *in* $\mathbb{R}^n$ *is convex.*

By replacing f by g_A, we conclude that the image of the momentum map ∇g_A is convex.

Proof of Proposition 1. The momentum map ∇g_A maps $\mathscr{T}^n$ onto the interior of $\mathrm{Conv}(A)$. Indeed, let $a = A^\alpha$ be a row of A, associated to a vertex of $\mathrm{Conv}(A)$. Then there is a direction $v \in \mathbb{R}^n$ such that

$$a \cdot v = \max_{x \in \mathrm{Conv}(A)} x \cdot v$$

for some unique a.

We claim that $a \in \overline{\nabla g_A(\mathbb{R}^n)}$. Indeed, let $x(t) = v_A(tv)$, t a real parameter. If b is another row of A,

$$e^{a \cdot tv} = e^{ta \cdot v} \geqslant e^{tb \cdot v} = e^{b \cdot tv}$$

[3] The Atiyah–Guillemin–Sternberg Theorem applies to compact symplectic manifolds and the implied compactification of $\mathscr{T}^n$ may have singularities.

as $t \to \infty$. We can then write $\hat{v}_A(tv)^{2\mathrm{T}}$ as

$$\hat{v}_A(tv)^{2\mathrm{T}} = \begin{bmatrix} \vdots \\ \mathrm{e}^{ta\cdot v} \\ \vdots \end{bmatrix}^{\mathrm{T}} C\,\mathrm{Diag}\begin{bmatrix} \vdots \\ \mathrm{e}^{ta\cdot v} \\ \vdots \end{bmatrix}.$$

Since C is positive definite, $C_{\alpha\alpha} > 0$ and

$$\lim_{t\to\infty} v_A(tv)^{2\mathrm{T}} = \lim_{t\to\infty} \frac{\hat{v}_A(tv)^{2\mathrm{T}}}{\|\hat{v}_A(tv)\|^2} = \mathrm{e}_a^{\mathrm{T}} \frac{C_{\alpha\alpha}}{C_{\alpha\alpha}} = \mathrm{e}_a^{\mathrm{T}},$$

where e_a is the unit vector in $\mathbb{R}^M$ corresponding to the row a. It follows that $\lim_{t\to\infty} \nabla g_A(tv) = a$.

When we set $q = 0$, we have $\det D^2 g_A \neq 0$ on $\mathbb{R}^n$, so we have a local diffeomorphism at each point $p \in \mathbb{R}^n$. Assume that $(\nabla g_A)_p = (\nabla g_A)_{p'}$ for $p \neq p'$. Then, let $\gamma(t) = (1-t)p + tp'$. The function $t \mapsto (\nabla g_A)_{\gamma(t)}\gamma'(t)$ has the same value at 0 and at 1, hence by Rolle's Theorem its derivative must vanish at some $t^* \in (0,1)$.

In that case,

$$(D^2 g_A)_{\gamma(t^*)}(\gamma'(t^*), \gamma'(t^*)) = 0$$

and since $\gamma'(t^*) = p' - p \neq 0$, $\det D^2 g_A$ must vanish in some $p \in \mathbb{R}^n$. This contradicts Lemma 1. □

2.3. The condition matrix

Following [6], we look at the linearization of the implicit function $p + q\sqrt{-1} = G(f)$ for the equation $ev_{\mathscr{A}}(f, p + q\sqrt{-1}) = 0$.

Definition 7. The *condition matrix* of ev at $(f, p + q\sqrt{-1})$ is

$$DG = D_{\mathscr{T}^n}(ev)^{-1} D_{\mathscr{F}}(ev),$$

where $\mathscr{F} = \mathscr{F}_{A_1} \times \cdots \times \mathscr{F}_{A_n}$.

Above, $D_{\mathscr{T}^n}(ev)$ is a linear operator from an n-dimensional complex space into $\mathbb{C}^n$, while $D_{\mathscr{F}}(ev)$ goes from an $(M_1 + \cdots + M_n)$-dimensional complex space into $\mathbb{C}^n$.

Lemma 3. *If* $p + \mathrm{i}q \in \mathscr{T}^n$ *and* $f(\exp(p + \mathrm{i}q)) = \mathbf{O}$ *then*

$$\begin{aligned} &\det(DG\,DG^H)^{-1}\,\mathrm{d}p_1 \wedge \mathrm{d}q_1 \wedge \cdots \wedge \mathrm{d}p_n \wedge \mathrm{d}q_n \\ &\quad = (-1)^{n(n-1)/2} \wedge \sqrt{-1} f^i \cdot (Dv_{A_i})_{(p,q)}\,\mathrm{d}p \wedge \bar{f}^i \cdot (Dv_{A_i})_{(p,-q)}\,\mathrm{d}q. \end{aligned}$$

Note that although $f^i \cdot (Dv_{A_i})_{(p,q)}\,\mathrm{d}p$ is a complex-valued form, each wedge $f^i \cdot (Dv_{A_i})_{(p,q)}\mathrm{d}p \wedge \bar{f}^i \cdot (Dv_{A_i})_{(p,-q)}\,\mathrm{d}q$ is a real-valued 2-form.

Proof of Lemma 3. We compute:

$$D_{\mathscr{F}}(ev)|_{(p,q)} = \begin{bmatrix} \sum_{\alpha=1}^{M_1} \hat{v}_{A_1}^{\alpha}(p+q\sqrt{-1})\,\mathrm{d}f_{\alpha}^{1} \\ \vdots \\ \sum_{\alpha=1}^{M_n} \hat{v}_{A_n}^{\alpha}(p+q\sqrt{-1})\,\mathrm{d}f_{\alpha}^{n} \end{bmatrix},$$

and hence

$$D_{\mathscr{F}}(ev)D_{\mathscr{F}}(ev)^H = \operatorname{diag}\|\hat{v}_{A_i}\|^2.$$

Also,

$$D_{\mathscr{T}^n}(ev) = \begin{bmatrix} f^1 \cdot D\hat{v}_{A_1} \\ \vdots \\ f^n \cdot D\hat{v}_{A_n} \end{bmatrix}.$$

Therefore,

$$\det(DG_{(p,q)}DG_{(p,q)}^H)^{-1} = \left|\det \begin{bmatrix} f^1 \cdot \frac{1}{\|\hat{v}_{A_1}\|} D\hat{v}_{A_1} \\ \vdots \\ f^n \cdot \frac{1}{\|\hat{v}_{A_n}\|} D\hat{v}_{A_n} \end{bmatrix}\right|^2.$$

We can now use Lemma 2 to conclude the following:

Formula 2.3.1. Determinant of the condition matrix

$$\det(DG_{(p,q)}DG_{(p,q)}^H)^{-1} = \left|\det \begin{bmatrix} f^1 \cdot Dv_{A_1} \\ \vdots \\ f^n \cdot Dv_{A_n} \end{bmatrix}\right|^2.$$

We can now write the same formula as a determinant of a block matrix:

$$\det(DG_{(p,q)}DG_{(p,q)}^H)^{-1} = \det \begin{bmatrix} f^1 \cdot Dv_{A_1} & \\ \vdots & \\ f^n \cdot Dv_{A_n} & \\ & \bar{f}^1 \cdot D\bar{v}_{A_1} \\ & \vdots \\ & \bar{f}^n \cdot D\bar{v}_{A_n} \end{bmatrix}$$

and replace the determinant by a wedge. The factor $(-1)^{n(n-1)/2}$ comes from replacing $\mathrm{d}p_1 \wedge \cdots \wedge \mathrm{d}p_n \wedge \mathrm{d}q_1 \wedge \cdots \wedge \mathrm{d}q_n$ by $\mathrm{d}p_1 \wedge \mathrm{d}q_1 \wedge \cdots \wedge \mathrm{d}p_n \wedge \mathrm{d}q_n$. □

We are now ready to prove our main theorems.

3. The proofs of Theorems 1–6

We first prove that Theorem 1 follows from Theorem 4. Then we will prove our remaining main theorems in the following order: 2, 4, 5, 3, 6.

3.1. *Proof of Theorem 1*

The first assertion, modulo an exponential change of coordinates and using the multi-projective metric $d_{\mathbb{P}}(\cdot,\cdot)$, follows immediately from Theorem 4.

As for the rest of Theorem 1, Theorem 4 applied to the *linear* case then provides the following interpretation of $v^{\mathrm{Lin}}(n,\varepsilon)$:

$$v^{\mathrm{Lin}}(n,\varepsilon) = \mathrm{Prob}[d_{\mathbb{P}}(f,\Sigma_{(p,q)}) < \varepsilon],$$

where f is a complex random linear polynomial system, and (p,q) is such that $f(\exp(p+\mathrm{i}q))=0$. So we are on our way to proving the inequality

$$\mathrm{Prob}[d_{\mathbb{P}}(f,\Sigma_{(p,q)}) < \varepsilon] \leqslant n^3(n+1)\,\mathrm{Vol}(A)(\#A-1)(\#A-2)\varepsilon^4,$$

for *general* f, which clearly implies our desired bound.

To prove the latter inequality, recall that by the definition of the multi-projective distance $d_{\mathbb{P}}(\cdot,\cdot)$, we have the following equality:

$$d_{\mathbb{P}}(f,\Sigma_{(p,q)})^2 = \min_{\substack{g\in\Sigma_{(p,q)}\\ \lambda\in(\mathbb{C}^*)^n}} \sum_{i=1}^{n} \frac{\|f^i-\lambda_i g^i\|^2}{\|f^i\|^2}.$$

So let g be so that the above minimum is attained. Without loss of generality, we may scale the g^i so that $\lambda_1=\cdots=\lambda_n=1$. In that case,

$$d_{\mathbb{P}}(f,\Sigma_{(p,q)})^2 = \sum_{i=1}^{n} \frac{\|f^i-g^i\|^2}{\|f^i\|^2} \geqslant \frac{\sum_{i=1}^{n}\|f^i-g^i\|^2}{\sum_{j=1}^{n}\|f^j\|^2}.$$

We are then in the setting of [6, pp. 248–250], where we identify our linear f with a normally distributed $(n+1)\times n$ complex matrix. The right-hand side in the above inequality is then precisely the left-hand term in [6, Remark 2, p. 250]. Therefore, using the notation of [6, Proposition 4], $d_{\mathbb{P}}(f,\Sigma_{(p,q)})\geqslant d_{\mathrm{F}}(f,\Sigma_x)$. So it follows that

$$v^{\mathrm{Lin}}(n,\varepsilon) = \mathrm{Prob}[d_{\mathbb{P}}(f,\Sigma_{(p,q)}) < \varepsilon] \leqslant \mathrm{Prob}[d_{\mathrm{F}}(f,\Sigma_x) < \varepsilon]$$

and the last probability is bounded above by $n^3(n+1)(\#A-1)(\#A-2)\varepsilon^4$ via [6, Therorem 6, p. 254]. Theorem 1 now follows. □

3.2. Proof of Theorem 2

Using [6, Theorem 5, p. 243] (or Proposition 5, p. 31 below), we deduce that the average number of complex roots is

$$\mathrm{Avg} = \int_{(p,q)\in U} \int_{f\in \mathscr{F}_{(p,q)}} \left(\prod \frac{\mathrm{e}^{-\|f^i\|^2/2}}{(2\pi)^{M_i}} \right) \det(DG_{(p,q)} DG^H_{(p,q)})^{-1}.$$

By Lemma 3, we can replace the inner integral by a $2n$-form valued integral:

$$\mathrm{Avg} = (-1)^{n(n-1)/2} \int_{(p,q)\in U} \int_{f\in \mathscr{F}_{(p,q)}} \bigwedge_i \frac{\mathrm{e}^{-\|f^i\|^2/2}}{(2\pi)^{M_i}} f^i \cdot (Dv_{A_i})_{(p,q)}\, \mathrm{d}p \wedge \bar{f}^i \cdot (Dv_{A_i})_{(p,-q)}\, \mathrm{d}q.$$

Since the image of Dv_{A_i} is precisely $\mathscr{F}_{A_i,(p,q)} \subset \mathscr{F}_{A_i}$, one can add n extra variables corresponding to the directions $v_{A_i}(p+q\sqrt{-1})$ without changing the integral: we write $\mathscr{F}_{A_i} = \mathscr{F}_{A_i,(p,q)} \times \mathbb{C} v_{A_i}(p+q\sqrt{-1})$. Since $(f^i + t v_{A_i}(p+q\sqrt{-1}))Dv_{A_i}$ is equal to $f^i Dv_{A_i}$, the average number of roots is indeed:

$$\mathrm{Avg} = (-1)^{n(n-1)/2} \int_{(p,q)\in U} \int_{f\in \mathscr{F}} \bigwedge_i \frac{\mathrm{e}^{-\|f^i\|^2/2}}{(2\pi)^{M_i+1}} f^i \cdot (Dv_{A_i})_{(p,q)}\, \mathrm{d}p \wedge \bar{f}^i \cdot (Dv_{A_i})_{(p,-q)}\, \mathrm{d}q.$$

In the integral above, all the terms that are multiple of $f^i_\alpha \bar{f}^i_\beta$ for some $\alpha \neq \beta$ will cancel out. Therefore,

$$\mathrm{Avg} = (-1)^{n(n-1)/2} \int_{(p,q)\in U} \int_{f\in \mathscr{F}} \bigwedge_i \frac{\mathrm{e}^{-\|f^i\|^2/2}}{(2\pi)^{M_i+1}} \sum_\alpha |f^i_\alpha|^2 (Dv_{A_i})^\alpha_{(p,q)}\, \mathrm{d}p \wedge (Dv_{A_i})^\alpha_{(p,-q)}\, \mathrm{d}q.$$

Now, we apply the integral formula:

$$\int_{x\in\mathbb{C}^M} |x_1|^2 \frac{\mathrm{e}^{-\|x\|^2/2}}{(2\pi)^M} = \int_{x_1\in\mathbb{C}} |x_1|^2 \frac{\mathrm{e}^{-|x_1|^2/2}}{2\pi} = 2$$

to obtain:

$$\mathrm{Avg} = \frac{(-1)^{n(n-1)/2}}{\pi^n} \int_{(p,q)\in U} \bigwedge \sum_\alpha (Dv_{A_i})^\alpha_{(p,q)}\, \mathrm{d}p \wedge (Dv_{A_i})^\alpha_{(p,-q)}\, \mathrm{d}q.$$

According to formulæ 2.1.3 and 2.1.4, the integrand is just $2^{-n} \wedge \omega_{A_i}$, and thus

$$\mathrm{Avg} = \frac{(-1)^{n(n-1)/2}}{\pi^n} \int_U \bigwedge_i \omega_{A_i} = \frac{n!}{\pi^n} \int_U \mathrm{d}\mathscr{T}^n. \qquad \square$$

3.3. Proof of Theorem 4

Let $(p,q)\in\mathscr{T}^n$ and let $f\in\mathscr{F}_{(p,q)}$. Without loss of generality, we can assume that f is scaled so that for all i, $\|f^i\|=1$.

Let $\delta f\in\mathscr{F}_{(p,q)}$ be such that $f+\delta f$ is singular at (p,q), and assume that $\sum\|\delta f^i\|^2$ is minimal. Then, due to the scaling we choose,

$$d_{\mathbb{P}}(f,\Sigma_{(p,q)})=\sqrt{\sum\|\delta f^i\|^2}.$$

Since $f+\delta f$ is singular, there is a vector $u\neq 0$ such that

$$\begin{bmatrix}(f^1+\delta f^1)\cdot(D\hat{v}_{A_1})_{(p,q)}\\ \vdots\\ (f^n+\delta f^n)\cdot(D\hat{v}_{A_n})_{(p,q)}\end{bmatrix}u=0$$

and hence

$$\begin{bmatrix}(f^1+\delta f^1)\cdot(Dv_{A_1})_{(p,q)}\\ \vdots\\ (f^n+\delta f^n)\cdot(Dv_{A_n})_{(p,q)}\end{bmatrix}u=0.$$

This means that

$$\begin{aligned}&f^1\cdot Dv_{A_1}u=-\delta f^1\cdot Dv_{A_1}u\\ &\vdots\\ &f^n\cdot Dv_{A_n}u=-\delta f^n\cdot Dv_{A_n}u.\end{aligned}$$

Let $D(f)$ denote the matrix

$$D(f)\stackrel{\text{def}}{=}\begin{bmatrix}f^1\cdot(Dv_{A_1})_{(p,q)}\\ \vdots\\ f^n\cdot(Dv_{A_n})_{(p,q)}\end{bmatrix}.$$

Given $v=D(f)u$, we obtain:

$$\begin{aligned}&v_1=-\delta f^1\cdot Dv_{A_1}D(f)^{-1}v\\ &\vdots\\ &v_n=-\delta f^n\cdot Dv_{A_n}D(f)^{-1}v.\end{aligned}\tag{3.3.1}$$

We can then scale u and v, such that $\|v\|=1$.

Claim 1. *Under the assumptions above,* δf^i *is colinear to* $(Dv_{A_i}D(f)^{-1}v)^H$.

Proof. Assume that $\delta f^i=g+h$, with g colinear and h orthogonal to $(Dv_{A_i}D(f)^{-1}v)^H$. As the image of Dv_{A_i} is orthogonal to v_{A_i}, g is orthogonal to $v_{A_i}^H$, so $ev(g^i,(p,q))=0$ and hence $ev(h^i,(p,q))=0$. We can therefore replace δf^i by g without compromising equality (3.3.1). Since $\|\delta f\|$ was minimal, this implies $h=0$. □

We obtain now an explicit expression for δf^i in terms of v:

$$\delta f^i = -v_i \frac{(Dv_{A_i} D(f)^{-1} v)^H}{\|Dv_{A_i} D(f)^{-1} v\|^2}.$$

Therefore,

$$\|\delta f^i\| = \frac{|v_i|}{\|Dv_{A_i} D(f)^{-1} v\|} = \frac{|v_i|}{\|(D(f)^{-1} v)\|_{A_i}}.$$

So we have proved the following result:

Lemma 4. *Fix v so that $\|v\| = 1$ and let $\delta f \in \mathscr{F}_{(p,q)}$ be such that Eq. (3.3.1) holds and $\|\delta f\|$ is minimal. Then,*

$$\|\delta f^i\| = \frac{|v_i|}{\|D(f)^{-1} v\|_{A_i}}.$$

Lemma 4 provides an immediate lower bound for $\|\delta f\| = \sqrt{\sum \|\delta f^i\|^2}$: Since

$$\|\delta f^i\| \geqslant \frac{|v_i|}{\max_j \|D(f)^{-1} v\|_{A_j}},$$

we can use $\|v\| = 1$ to deduce that

$$\sqrt{\sum_i \|\delta f^i\|^2} \geqslant \frac{1}{\max_j \|D(f)^{-1} v\|_{A_j}} \geqslant \frac{1}{\max_j \|D(f)^{-1}\|_{A_j}}.$$

Also, for any v with $\|v\| = 1$, we can choose δf minimal so that Eq. (3.3.1) applies. Using Lemma 4, we obtain:

$$\|\delta f^i\| \leqslant \frac{|v_i|}{\min_j \|D(f)^{-1} v\|_{A_j}}.$$

Hence

$$\sqrt{\sum_i \|\delta f^i\|^2} \leqslant \frac{1}{\min_j \|D(f)^{-1} v\|_{A_j}}.$$

Since this is true for any v, and $\|\delta f\|$ is minimal for all v, we have

$$\sqrt{\sum_i \|\delta f^i\|^2} \leqslant \frac{1}{\max_{\|v\|=1} \min_j \|D(f)^{-1}\|_{A_j}}$$

and this proves Theorem 4. □

3.4. The idea behind the proof of Theorem 5

The proof of Theorem 5 is long. We first sketch the idea of the proof. Recall that $\mathscr{F}_{(p,q)}$ is the set of all $f \in \mathscr{F}$ such that $ev(f; p + q\sqrt{-1}) = 0$, and that $\Sigma_{(p,q)}$ is the restriction of the discriminant to the fiber $\mathscr{F}_{(p,q)}$:

$$\Sigma_{(p,q)} \stackrel{\text{def}}{=} \{f \in \mathscr{F}_{(p,q)} : D(f)_{(p,q)} \text{ does not have full rank}\}.$$

The space $\mathscr{F}$ is endowed with a Gaussian probability measure, with volume element

$$\frac{e^{-\|f\|^2/2}}{(2\pi)^{\sum M_i}} \, d\mathscr{F},$$

where $d\mathscr{F}$ is the usual volume form in $\mathscr{F} = (\mathscr{F}_{A_1}, \langle \cdot,\cdot \rangle_{A_1}) \times \cdots \times (\mathscr{F}_{A_n}, \langle \cdot,\cdot \rangle_{A_n})$ and $\|f\|^2 = \sum \|f^i\|^2_{A_i}$. For U a set in $\mathscr{T}^n$, we defined earlier (in the statement of Theorem 5) the quantity:

$$v^A(U,\varepsilon) \stackrel{\text{def}}{=} \text{Prob}[\boldsymbol{\mu}(f,U) > \varepsilon^{-1}] = \text{Prob}[\exists (p,q) \in U : d_{\mathbb{P}}(f, \Sigma_{(p,q)}) < \varepsilon].$$

The naïve idea for bounding $v^A(U,\varepsilon)$ is as follows: Let $V(\varepsilon) \stackrel{\text{def}}{=} \{(f,(p,q)) \in \mathscr{F} \times U : ev(f;(p,q)) = 0 \text{ and } d_{\mathbb{P}}(f,\Sigma_{(p,q)}) < \varepsilon\}$. We also define $\pi : V(\varepsilon) \to \mathscr{F}$ as the canonical projection mapping $\mathscr{F} \times U$ to $\mathscr{F}$, and set $\#_{V(\varepsilon)}(f) \stackrel{\text{def}}{=} \#\{(p,q) \in U : (f,(p,q)) \in V(\varepsilon)\}$. Then,

$$\begin{aligned} v^A(U,\varepsilon) &= \int_{f \in \mathscr{F}} \chi_{\pi(V(\varepsilon))}(f) \, \frac{e^{-\|f\|^2/2}}{(2\pi)^{\sum M_i}} \, d\mathscr{F} \\ &\leqslant \int_{f \in \mathscr{F}} \#_{V(\varepsilon)} \, \frac{e^{-\|f\|^2/2}}{(2\pi)^{\sum M_i}} \, d\mathscr{F} \end{aligned}$$

with equality in the linear case and when $\varepsilon > \sqrt{n}$.

Now we apply the coarea formula [6, Theorem 5, p. 243] to obtain:

$$v^A(U,\varepsilon) \leqslant \int_{(p,q) \in U \subset \mathscr{T}^n} \int_{\substack{f \in \mathscr{F}_{(p,q)} \\ d_{\mathbb{P}}(f,\Sigma_{(p,q)}) < \varepsilon}} \frac{1}{NJ(f;(p,q))} \, \frac{e^{-\|f\|^2/2}}{(2\pi)^{\sum M_i}} \, d\mathscr{F} \, dV_{\mathscr{T}^n},$$

where $dV_{\mathscr{T}^n}$ stands for Lebesgue measure in $\mathscr{T}^n$. Again, in the linear case, we have equality.

We already know from Lemma 3 that

$$1/NJ(f;(p,q)) = \bigwedge_{i=1}^{n} f^i \cdot (Dv_{A_i})_{(p,q)} \, dp \wedge \bar{f}^i \cdot (D\bar{v}_{A_i})_{(p,q)} \, dq.$$

We should focus now on the inner integral. In each coordinate space $\mathscr{F}_{A_i}$, we can introduce a new orthonormal system of coordinates (depending on (p,q)) by decomposing:

$$f^i = f^i_{\mathrm{I}} + f^i_{\mathrm{II}} + f^i_{\mathrm{III}},$$

where f^i_{I} is the component colinear to $v^H_{A_i}$, f^i_{II} is the projection of f^i to $(\text{range}\, Dv_{A_i})^H$, and f^i_{III} is orthogonal to f^i_{I} and f^i_{II}.

Of course, $f^i \in (\mathscr{F}_{A_i})_{(p,q)}$ if and only if $f^i_{\mathrm{I}} = 0$.
Also,

$$\bigwedge_{i=1}^{n} f^i \cdot (Dv_{A_i})_{(p,q)}\,\mathrm{d}p \wedge \bar{f}^i \cdot (D\bar{v}_{A_i})_{(p,q)}\,\mathrm{d}q = \bigwedge_{i=1}^{n} f^i_{\mathrm{II}} \cdot (Dv_{A_i})_{(p,q)}\,\mathrm{d}p \wedge \bar{f}^i_{\mathrm{II}} \cdot (D\bar{v}_{A_i})_{(p,q)}\,\mathrm{d}q.$$

It is an elementary fact that

$$d_{\mathbb{P}}(f^i_{\mathrm{II}} + f^i_{\mathrm{III}}, \Sigma_{(p,q)}) \leqslant d_{\mathbb{P}}(f^i_{\mathrm{II}}, \Sigma_{(p,q)}).$$

It follows that for $f \in \mathscr{F}_{(p,q)}$:

$$d_{\mathbb{P}}(f, \Sigma_{(p,q)}) \leqslant d_{\mathbb{P}}(f_{\mathrm{II}}, \Sigma_{(p,q)}),$$

with equality in the linear case. Hence, we obtain:

$$v^A(U,\varepsilon) \leqslant \int_{(p,q)\in U\subset \mathscr{T}^n} \int_{\substack{f\in\mathscr{F}_{(p,q)} \\ d_{\mathbb{P}}(f_{\mathrm{II}},\Sigma_{(p,q)})<\varepsilon}} \left(\bigwedge_{i=1}^{n} f^i_{\mathrm{II}} \cdot (Dv_{A_i})_{(p,q)}\,\mathrm{d}p \wedge \bar{f}^i_{\mathrm{II}} \cdot (D\bar{v}_{A_i})_{(p,q)}\,\mathrm{d}q \right) \cdot \frac{\mathrm{e}^{-\|f^i_{\mathrm{II}}+f^i_{\mathrm{III}}\|^2/2}}{(2\pi)^{\sum M_i}}\,\mathrm{d}\mathscr{F}\,\mathrm{d}V_{\mathscr{T}^n},$$

with equality in the linear case. We can integrate the $\sum(M_i - n - 1)$ variables f_{III} to obtain:

Proposition 3.

$$v^A(U,\varepsilon) \leqslant \int_{(p,q)\in U\subset \mathscr{T}^n} \int_{\substack{f_{\mathrm{II}}\in\mathbb{C}^{n^2} \\ d_{\mathbb{P}}(f_{\mathrm{II}},\Sigma_{(p,q)})<\varepsilon}} \left(\bigwedge_{i=1}^{n} f^i_{\mathrm{II}} \cdot (Dv_{A_i})_{(p,q)}\,\mathrm{d}p \wedge \bar{f}^i_{\mathrm{II}} \cdot (D\bar{v}_{A_i})_{(p,q)}\,\mathrm{d}q \right) \cdot \frac{\mathrm{e}^{-\|f^i_{\mathrm{II}}\|^2/2}}{(2\pi)^{n(n+1)}}\,\mathrm{d}V_{\mathscr{T}^n}$$

with equality in the linear case.

3.5. Proof of Theorem 5

The domain of integration in Proposition 3 makes integration extremely difficult. In order to estimate the inner integral, we will need to perform a change of coordinates.

Unfortunately, the Gaussian in Proposition 3 makes that change of coordinates extremely hard, and we will have to restate Proposition 3 in terms of integrals over a product of projective spaces.

The domain of integration will be $\mathbb{P}^{n-1} \times \cdots \times \mathbb{P}^{n-1}$. Translating an integral in terms of Gaussians to an integral in terms of projective spaces is not immediate, and we will use the following elementary fact about Gaussians:

Lemma 5. *Let $\varphi : \mathbb{C}^n \to \mathbb{R}$ be $\mathbb{C}^*$-invariant* (*in the sense of the usual scaling action*). *Then we can also interpret φ as a function from $\mathbb{P}^{n-1}$ into $\mathbb{R}$, and*:

$$\frac{1}{\mathrm{Vol}(\mathbb{P}^{n-1})} \int_{[x]\in\mathbb{P}^{n-1}} \varphi(x)\,\mathrm{d}[x] = \int_{x\in\mathbb{C}^n} \varphi(x)\,\frac{\mathrm{e}^{-\|x\|^2/2}}{(2\pi)^n}\,\mathrm{d}x,$$

where, *respectively*, *the natural volume forms on $\mathbb{P}^{n-1}$ and $\mathbb{C}^n$ are understood for each integral.*

Now the integrand in Proposition 3 is not $\mathbb{C}^*$-invariant. This is why we will need the following formula:

Lemma 6. *Under the hypotheses of Lemma* 5,

$$\frac{1}{\mathrm{Vol}(\mathbb{P}^{n-1})} \int_{[x]\in\mathbb{P}^{n-1}} \varphi(x)\,\mathrm{d}[x] = \frac{1}{2n} \int_{x\in\mathbb{C}^n} \|x\|^2 \varphi(x)\,\frac{\mathrm{e}^{-\|x\|^2/2}}{(2\pi)^n}\,\mathrm{d}x,$$

where, *respectively*, *the natural volume forms on $\mathbb{P}^{n-1}$ and $\mathbb{C}^n$ are understood for each integral.*

Proof.

$$\begin{aligned}
&\int_{x\in\mathbb{C}^n} \|x\|^2 \varphi(x)\,\frac{\mathrm{e}^{-\|x\|^2/2}}{(2\pi)^n}\,\mathrm{d}x \\
&\quad = \int_{\Theta\in S^{2n-1}} \int_{r=0}^{\infty} |r|^{2n+1} \varphi(\Theta)\,\frac{\mathrm{e}^{-|r|^2/2}}{(2\pi)^n}\,\mathrm{d}r\,\mathrm{d}\Theta \\
&\quad = \int_{\Theta\in S^{2n-1}} \left(-\left[|r|^{2n}\,\frac{\mathrm{e}^{-|r|^2/2}}{(2\pi)^n} \right]_0^{\infty} + 2n \int_{r=0}^{\infty} |r|^{2n-1}\,\frac{\mathrm{e}^{-|r|^2/2}}{(2\pi)^n}\,\mathrm{d}r \right) \varphi(\Theta)\,\mathrm{d}\Theta \\
&\quad = 2n \int_{x\in\mathbb{C}^n} \varphi(x)\,\frac{\mathrm{e}^{-\|x\|^2/2}}{(2\pi)^n}\,\mathrm{d}x. \qquad \square
\end{aligned}$$

We can now introduce the notation:

$$\mathrm{WEDGE}^A(f_{\mathrm{II}}) \stackrel{\mathrm{def}}{=} \bigwedge_{i=1}^{n} \frac{1}{\|f_{\mathrm{II}}^i\|^2}\, f_{\mathrm{II}}^i \cdot (Dv_{A_i})_{(p,q)}\,\mathrm{d}p \wedge \bar{f}_{\mathrm{II}}^i \cdot (D\bar{v}_{A_i})_{(p,q)}\,\mathrm{d}q.$$

This function is invariant under the $(\mathbb{C}^*)^n$-action $\lambda \star f_{\mathrm{II}} : f_{\mathrm{II}} \mapsto (\lambda_1 f_{\mathrm{II}}^1, \ldots, \lambda_n f_{\mathrm{II}}^n)$.

We adopt the following conventions: $\mathscr{F}_{\mathrm{II}} \subset \mathscr{F}$ is the space spanned by coordinates f_{II} and $\mathbb{P}(\mathscr{F}_{\mathrm{II}})$ is its quotient by $(\mathbb{C}^*)^n$.

We apply n times Lemma 6 and obtain:

Proposition 4. *Let* $\mathrm{VOL} \stackrel{\mathrm{def}}{=} \mathrm{Vol}(\mathbb{P}^{n-1})^n$. *Then,*

$$v^A(U,\varepsilon) \leqslant \frac{(2n)^n}{\mathrm{VOL}} \int_{(p,q)\in U\subset \mathcal{T}^n} \int_{\substack{f_{\mathrm{II}}\in\mathbb{P}(\mathcal{F}_{\mathrm{II}})\\ d_{\mathbb{P}}(f_{\mathrm{II}},\Sigma_{(p,q)})<\varepsilon}} \mathrm{WEDGE}^A(f_{\mathrm{II}})\,\mathrm{d}\mathbb{P}(\mathcal{F}_{\mathrm{II}})\,\mathrm{d}V_{\mathcal{T}^n}$$

with equality when $\varepsilon > \sqrt{n}$. *In the linear case,*

$$v^{\mathrm{Lin}}(U,\varepsilon) = \frac{(2n)^n}{\mathrm{VOL}} \int_{(p,q)\in U\subset \mathcal{T}^n} \int_{\substack{g_{\mathrm{II}}\in\mathbb{P}(\mathcal{F}_{\mathrm{II}}^{\mathrm{Lin}})\\ d_{\mathbb{P}}(g_{\mathrm{II}},\Sigma^{\mathrm{Lin}}_{(p,q)})<\varepsilon}} \mathrm{WEDGE}^{\mathrm{Lin}}(g_{\mathrm{II}})\,\mathrm{d}(P\mathcal{F}_{\mathrm{II}}^{\mathrm{Lin}})\,\mathrm{d}V_{\mathcal{T}^n}.$$

□

Now we introduce the following change of coordinates. Let $L\in GL(n)$ be such that the minimum in Definition 5, p. 6 is attained:

$$\varphi : \mathbb{P}^{n-1}\times\cdots\times\mathbb{P}^{n-1} \to \mathbb{P}^{n-1}\times\cdots\times\mathbb{P}^{n-1}$$

$$f_{\mathrm{II}} \mapsto g_{\mathrm{II}} \stackrel{\mathrm{def}}{=} \varphi(f_{\mathrm{II}}), \quad \text{such that } g^i_{\mathrm{II}} = f^i_{\mathrm{II}}\cdot Dv_{A_i}L.$$

Without loss of generality, we scale L such that $\det L = 1$. The following property follows from the definition of WEDGE:

$$\mathrm{WEDGE}^A(f_{\mathrm{II}}) = \mathrm{WEDGE}^{\mathrm{Lin}}(g_{\mathrm{II}}) \prod_{i=1}^n \frac{\|g^i_{\mathrm{II}}\|^2}{\|f^i_{\mathrm{II}}\|^2}. \tag{3.5.1}$$

Assume now that $d_{\mathbb{P}}(f_{\mathrm{II}},\Sigma_{(p,q)})<\varepsilon$. Then there is $\delta f\in\mathcal{F}_{\mathrm{II}}$, such that $f+\delta f\in\Sigma^{\mathrm{Lin}}_{(p,q)}$ and $\|\delta f\|\leqslant\varepsilon$ (assuming the scaling $\|f^i_{\mathrm{II}}\|=1$ for all i).

Setting $g_{\mathrm{II}}=\varphi(f_{\mathrm{II}})$ and $\delta g=\varphi(\delta g)$, we obtain that $g+\delta g\in\Sigma^{\mathrm{Lin}}_{(p,q)}$.

$$d_{\mathbb{P}}(g,\Sigma^{\mathrm{Lin}}_{(p,q)}) \leqslant \sqrt{\sum_{i=1}^n \frac{\|\delta g^i\|^2}{\|g^i_{\mathrm{II}}\|^2}}.$$

At each value of i,

$$\frac{\|\delta g^i\|}{\|g^i_{\mathrm{II}}\|} \leqslant \frac{\|\delta f^i\|}{\|f^i_{\mathrm{II}}\|}\,\kappa(D_{f^i_{\mathrm{II}}}\varphi^i),$$

where κ denotes Wilkinson's condition number of the linear operator $D_{f^i_{\mathrm{II}}}\varphi^i$. This is precisely $\kappa(Dv_{A_i}L)$. Thus,

$$d_{\mathbb{P}}(g,\Sigma^{\mathrm{Lin}}_{(p,q)}) \leqslant \varepsilon\max_i\ \kappa(Dv_{A_i}L) = \max_i\ \sqrt{\kappa(\omega_{A_i})}.$$

Thus, an ε-neighborhood of $\Sigma^A_{(p,q)}$ is mapped into a $\sqrt{\kappa_U}\varepsilon$ neighborhood of $\Sigma^{\mathrm{Lin}}_{(p,q)}$.

We use this property and Eq. (3.5.1) to bound:

$$\nu^A(U,\varepsilon) \leqslant \frac{(2n)^n}{\mathrm{VOL}} \int_{(p,q)\in U\subset \mathcal{T}^n} \int_{\substack{g_{\mathrm{II}}\in \mathbb{P}^{n-1}\times\cdots\times\mathbb{P}^{n-1} \\ d_{\mathbb{P}}(g_{\mathrm{II}},\Sigma^{\mathrm{Lin}}_{(p,q)})<\sqrt{\kappa_U}\varepsilon}} \mathrm{WEDGE}^{\mathrm{Lin}}(g_{\mathrm{II}})$$

$$\cdot \prod_{i=1}^{n} \frac{\|g^i_{\mathrm{II}}\|^2}{\|f^i_{\mathrm{II}}\|^2} |J_{g_{\mathrm{II}}}\varphi^{-1}|^2 \,\mathrm{d}(\mathbb{P}^{n-1}\times\cdots\times\mathbb{P}^{n-1})\,\mathrm{d}V_{\mathcal{T}^n} \tag{3.5.2}$$

where $J_{g_{\mathrm{II}}}\varphi^{-1}$ is the Jacobian of φ^{-1} at g_{II}.

Remark 3. Considering each Dv_{A_i} as a map from $\mathbb{C}^n$ into $\mathbb{C}^n$, the Jacobian is

$$J_{g_{\mathrm{II}}}\varphi^{-1} = \prod_{i=1}^{n} \frac{\|\varphi^{-1}(g_{\mathrm{II}})^i\|^n}{\|g^i_{\mathrm{II}}\|^n} (\det Dv^H_{A_i} Dv_{A_i})^{-1/2}.$$

We will not use this value in the sequel.

In order to simplify the expressions for the bound on $\nu^A(U,\varepsilon)$, it is convenient to introduce the following notation:

$$\mathrm{d}P \stackrel{\mathrm{def}}{=} \frac{(2n)^n}{\mathrm{VOL}} \mathrm{WEDGE}^{\mathrm{Lin}}(g_{\mathrm{II}}) \frac{\mathrm{d}(\mathbb{P}^{n-1}\times\cdots\times\mathbb{P}^{n-1})}{n!(\omega_{\mathrm{Lin}})^{\bigwedge n}},$$

$$H \stackrel{\mathrm{def}}{=} \prod_{i=1}^{n} \frac{\|g^i_{\mathrm{II}}\|^2}{\|f^i_{\mathrm{II}}\|^2} |J_g\varphi^{-1}|^2,$$

$$\chi_\delta \stackrel{\mathrm{def}}{=} \chi_{\{g: d_{\mathbb{P}}(g,\Sigma^{\mathrm{Lin}}_{(p,q)})<\delta\}}.$$

Now Eq. (3.5.2) becomes:

$$\nu^A(U,\varepsilon) \leqslant n! \int_{(p,q)\in U\subset\mathcal{T}^n} (\omega_{\mathrm{Lin}})^{\bigwedge n} \int_{g_{\mathrm{II}}\in\mathbb{P}^{n-1}\times\cdots\times\mathbb{P}^{n-1}} \mathrm{d}P\, H(g_{\mathrm{II}})\chi_{\sqrt{\kappa_U}\varepsilon}(g_{\mathrm{II}}). \tag{3.5.3}$$

Lemma 7. *Let (p,q) be fixed. Then $\mathbb{P}^{n-1}\times\cdots\times\mathbb{P}^{n-1}$ together with density function* $\mathrm{d}P$, *is a probability space.*

Proof. The expected number of roots in U for a linear system is

$$n! \int_{(p,q)\in U} \omega^{\bigwedge n}_{\mathrm{Lin}} \int_{g_{\mathrm{II}}\in\mathbb{P}^{n-1}\times\cdots\times\mathbb{P}^{n-1}} \mathrm{d}P$$

and is also equal to $n!\int_U \omega^{\bigwedge n}_{\mathrm{Lin}}$. This holds for all U, hence the volume forms are the same and

$$\int_{g_{\mathrm{II}}\in\mathbb{P}^{n-1}\times\cdots\times\mathbb{P}^{n-1}} \mathrm{d}P = 1. \qquad \square$$

This allows us to interpret the inner integral of Eq. (3.5.3) as the expected value of a product. This is less than the product of the expected values, and:

$$v^A(U,\varepsilon) \leqslant n! \int_{(p,q)\in U\subset\mathscr{T}^n} (\omega_{\mathrm{Lin}})^{\bigwedge n} \left(\int_{g_{\mathrm{II}}\in\mathbb{P}^{n-1}\times\cdots\times\mathbb{P}^{n-1}} \mathrm{d}P\, H(g_{\mathrm{II}}) \right) \cdot \left(\int_{g_{\mathrm{II}}\in\mathbb{P}^{n-1}\times\cdots\times\mathbb{P}^{n-1}} \mathrm{d}P \chi_{\sqrt{\kappa_U}\varepsilon}(g_{\mathrm{II}}) \right).$$

Because generic (square) systems of linear equations have exactly one root, we can also consider U as a probability space, with probability measure $(1/\mathrm{Vol}^{\mathrm{Lin}}(U))n!\omega_{\mathrm{Lin}}^{\bigwedge n}$. Therefore, we can bound:

$$v^A(U,\varepsilon) \leqslant \frac{1}{\mathrm{Vol}^{\mathrm{Lin}}(U)} \left(\int_{(p,q)\in U} n!(\omega_{\mathrm{Lin}})^{\bigwedge n} \int_{g_{\mathrm{II}}\in\mathbb{P}^{n-1}\times\cdots\times\mathbb{P}^{n-1}} \mathrm{d}P\, H(g_{\mathrm{II}}) \right) \cdot \left(\int_{(p,q)\in U} n!(\omega_{\mathrm{Lin}})^{\bigwedge n} \int_{g_{\mathrm{II}}\in\mathbb{P}^{n-1}\times\cdots\times\mathbb{P}^{n-1}} \mathrm{d}P\, \chi_{\sqrt{\kappa_U}\varepsilon}(g_{\mathrm{II}}) \right).$$

The first parenthetical expression is $\mathrm{Vol}^A(U)$, the volume of U with respect to the toric volume form associated to $A=(A_1,\ldots,A_n)$. The second parenthetical expression is $v^{\mathrm{Lin}}(\sqrt{\kappa_U}\varepsilon, U)$. This concludes the proof of Theorem 5. □

3.6. *Proof of Theorem 3*

As in the complex case (Theorem 2), the expected number of roots can be computed by applying the coarea formula:

$$AVG = \int_{p\in U} \int_{f\in\mathscr{F}_p^{\mathbb{R}}} \prod_{i=1}^{n} \frac{\mathrm{e}^{-\|f^i\|^2/2}}{\sqrt{2\pi}^{M_i}} \det(DG\, DG^H)^{-1/2}.$$

Now there are three big differences. The set U is in $\mathbb{R}^n$ instead of $\mathscr{T}^n$, the space $\mathscr{F}_p^{\mathbb{R}}$ contains only real polynomials (and therefore has half the dimension), and we are integrating the square root of $1/\det(DG\, DG^H)$.

Since we do not know in general how to integrate such a square root, we bound the inner integral as follows. We consider the real Hilbert space of functions integrable in $\mathscr{F}_p^{\mathbb{R}}$ endowed with Gaussian probability measure. The inner product in this space is:

$$\langle \varphi, \psi \rangle \stackrel{\mathrm{def}}{=} \int_{\mathscr{F}_p^{\mathbb{R}}} \varphi(f)\psi(f) \prod_{i=1}^{n} \frac{\mathrm{e}^{-\|f^i\|^2/2}}{\sqrt{2\pi}^{M_i-1}}\, \mathrm{d}V,$$

where $\mathrm{d}V$ is Lebesgue volume. If $\mathbf{1}$ denotes the constant function equal to 1, we interpret

$$AVG = \int_{p\in U} (2\pi)^{-n/2} \langle \det(DG\, DG^H)^{-1/2}, \mathbf{1} \rangle.$$

Hence Cauchy–Schwartz inequality implies:

$$AVG \leqslant \int_{p\in U} (2\pi)^{-n/2} \|\det(DG\, DG^H)^{-1/2}\| \|\mathbf{1}\|.$$

By construction, $\|\mathbf{1}\| = 1$, and we are left with:

$$AVG \leqslant \int_{p\in U} (2\pi)^{-n/2} \sqrt{\int_{\mathscr{F}_p^{\mathbb{R}}} \prod_{i=1}^{n} \frac{e^{-\|f^i\|^2/2}}{\sqrt{2\pi}^{M_i-1}} \det(DG\, DG^H)^{-1}}.$$

As in the complex case, we add extra n variables:

$$AVG \leqslant (2\pi)^{-n/2} \int_{p\in U} \sqrt{\int_{\mathscr{F}^{\mathbb{R}}} \prod_{i=1}^{n} \frac{e^{-\|f^i\|^2/2}}{\sqrt{2\pi}^{M_i}} \det(DG\, DG^H)^{-1}},$$

and we interpret $\det(DG\, DG^H)^{-1}$ in terms of a wedge. Since

$$\int_{x\in\mathbb{R}^M} |x_1|^2 \frac{e^{-\|x\|^2/2}}{\sqrt{2\pi}^M} = \int_{y\in\mathbb{R}} y^2 \frac{e^{-y^2/2}}{\sqrt{2\pi}} = \int_{y\in\mathbb{R}} \frac{e^{-y^2/2}}{\sqrt{2\pi}} = 1,$$

we obtain:

$$AVG \leqslant (2\pi)^{-n/2} \int_{p\in U} \sqrt{n!\, d\mathscr{T}^n} = (2\pi)^{-n/2} \int_{p\in U} \sqrt{n!\, d\mathscr{T}^n}.$$

Now we would like to use Cauchy–Schwartz again. This time, the inner product is defined as

$$\langle \varphi, \psi \rangle \stackrel{\text{def}}{=} \int_{p\in U} \varphi(p)\psi(p)\, dV.$$

Hence,

$$AVG \leqslant (2\pi)^{-n/2} \langle n!\, d\mathscr{T}^n, \mathbf{1}\rangle \leqslant (2\pi)^{-n/2} \|n!\, d\mathscr{T}^n\| \|\mathbf{1}\|.$$

This time, $\|\mathbf{1}\|^2 = \lambda(U)$, so we bound:

$$\begin{aligned} AVG &\leqslant (2\pi)^{-n/2} \sqrt{\lambda(U)} \sqrt{\int_U n!\, d\mathscr{T}^n} \\ &\leqslant (4\pi^2)^{-n/2} \sqrt{\lambda(U)} \sqrt{\int_{(p,q)\in\mathscr{T}^n,\, p\in U} n!\, d\mathscr{T}^n}. \qquad \square \end{aligned}$$

3.7. Proof of Theorem 6

Let $\varepsilon > 0$. As in the mixed case, we define:

$$v_{\mathbb{R}}(U,\varepsilon) \stackrel{\text{def}}{=} \text{Prob}_{f\in\mathscr{F}}[\boldsymbol{\mu}(f;U) > \varepsilon^{-1}]$$
$$= \text{Prob}_{f\in\mathscr{F}}[\exists p \in U : ev(f;p) = 0 \text{ and } d_{\mathbb{P}}(f,\Sigma_p) < \varepsilon]$$

where now $U \in \mathbb{R}^n$.

Let $V(\varepsilon) \stackrel{\text{def}}{=} \{(f,p) \in F_{\mathbb{R}} \times U : ev(f;p) = 0 \text{ and } d_{\mathbb{P}}(f,\Sigma_p) < \varepsilon\}$. We also define $\pi : V(\varepsilon) \to \mathbb{P}(\mathscr{F})$ to be the canonical projection mapping $F_{\mathbb{R}} \times U$ to $F_{\mathbb{R}}$ and set $\#_{V(\varepsilon)}(f) \stackrel{\text{def}}{=} \#\{p \in U : (f,p) \in V(\varepsilon)\}$. Then,

$$v_{\mathbb{R}}(U,\varepsilon) = \int_{f\in\mathscr{F}^{\mathbb{R}}} \frac{\mathrm{e}^{-\sum_i \|f^i\|^2/2}}{\sqrt{2\pi}^{\sum M_i}} \chi_{\pi(V(\varepsilon))}(f)\, \mathrm{d}\mathscr{F}^{\mathbb{R}}$$
$$\leqslant \int_{f\in\mathscr{F}^{\mathbb{R}}} \frac{\mathrm{e}^{-\sum_i \|f^i\|^2/2}}{\sqrt{2\pi}^{\sum M_i}} \#_{V(\varepsilon)}\, \mathrm{d}\mathscr{F}^{\mathbb{R}}$$
$$\leqslant \int_{p\in U\subset\mathbb{R}^n} \int_{\substack{f\in\mathscr{F}^{\mathbb{R}}_p \\ d_{\mathbb{P}}(f,\Sigma_p)<\varepsilon}} \frac{\mathrm{e}^{-\sum_i \|f^i\|^2/2}}{\sqrt{2\pi}^{\sum M_i}} \frac{1}{NJ(f;p)}\, \mathrm{d}\mathscr{F}^{\mathbb{R}}_p\, \mathrm{d}V_{\mathscr{T}^n}.$$

As before, we change coordinates in each fiber of $\mathscr{F}^{\mathbb{R}}_A$ by

$$f = f_{\mathrm{I}} + f_{\mathrm{II}} + f_{\mathrm{III}}$$

with f^i_{I} colinear to v_A^{T}, $(f^i_{\mathrm{II}})^{\mathrm{T}}$ in the range of Dv_A, and f^i_{III} orthogonal to f^i_{I} and f^i_{II}. This coordinate system is dependent on $p + q\sqrt{-1}$.

In the new coordinate system, formula 2.3.1 splits as follows:

$$\det(DG_{(p)}DG^H_{(p)})^{-1/2}\, \mathrm{d}V_{\mathscr{T}^n}$$
$$= \left|\det \begin{bmatrix} (f^1_{\mathrm{II}})_1 & \dots & (f^1_{\mathrm{II}})_n \\ \vdots & & \vdots \\ (f^n_{\mathrm{II}})_1 & \dots & (f^n_{\mathrm{II}})_n \end{bmatrix}\right| \left|\det \begin{bmatrix} (Dv_A{}^{\mathrm{II}})^1_1 & \dots & (Dv_A{}^{\mathrm{II}})^1_n \\ \vdots & & \vdots \\ (Dv_A{}^{\mathrm{II}})^n_1 & \dots & (Dv_A{}^{\mathrm{II}})^n_n \end{bmatrix}\right| \mathrm{d}V$$
$$= \left|\det \begin{bmatrix} (f^1_{\mathrm{II}})_1 & \dots & (f^1_{\mathrm{II}})_n \\ \vdots & & \vdots \\ (f^n_{\mathrm{II}})_1 & \dots & (f^n_{\mathrm{II}})_n \end{bmatrix}\right| \sqrt{\det Dv_A^H Dv_A}.$$

The integral $E(U)$ of $\sqrt{\det Dv_A\, Dv_A^H}$ is the expected number of real roots on U, therefore

$$v_{\mathbb{R}}(U,\varepsilon) \leqslant E(U) \int_{\substack{f_{\rm II}+f_{\rm III}\in\mathscr{F}_p^{\mathbb{R}} \\ d_{\mathbb{P}}(f_{\rm II}+f_{\rm III},\Sigma_p)<\varepsilon}} \frac{\mathrm{e}^{-\sum_i \|f_{\rm II}^i+f_{\rm III}^i\|^2/2}}{\sqrt{2\pi}^{\sum M_i}} \cdot \left|\det \begin{bmatrix} (f_{\rm II}^1)_1 & \dots & (f_{\rm II}^1)_n \\ \vdots & & \vdots \\ (f_{\rm II}^n)_1 & \dots & (f_{\rm II}^n)_n \end{bmatrix}\right| \mathrm{d}\mathscr{F}_p^{\mathbb{R}}.$$

In the new system of coordinates, Σ_p is defined by the equation:

$$\det \begin{bmatrix} (f_{\rm II}^1)_1 & \dots & (f_{\rm II}^1)_n \\ \vdots & & \vdots \\ (f_{\rm II}^n)_1 & \dots & (f_{\rm II}^n)_n \end{bmatrix} = 0.$$

Since $\|f_{\rm II}+f_{\rm III}\| \geqslant \|f_{\rm II}\|$,

$$d_{\mathbb{P}}(f_{\rm II}+f_{\rm III},\Sigma_p) < \varepsilon \Rightarrow d_{\mathbb{P}}(f_{\rm II},\Sigma_p) < \varepsilon.$$

This implies:

$$v_{\mathbb{R}}(U,\varepsilon) \leqslant E(U) \int_{\substack{f_{\rm II}+f_{\rm III}\in F_p^{\mathbb{R}} \\ d_{\mathbb{P}}(f_{\rm II},[\det=0])<\varepsilon}} \frac{\mathrm{e}^{-\sum_i \|f_{\rm II}^i+f_{\rm III}^i\|^2/2}}{\sqrt{2\pi}^{\sum M_i}} \cdot \left|\det \begin{bmatrix} (f_{\rm II}^1)_1 & \dots & (f_{\rm II}^1)_n \\ \vdots & & \vdots \\ (f_{\rm II}^n)_1 & \dots & (f_{\rm II}^n)_n \end{bmatrix}\right| \mathrm{d}\mathscr{F}_p^{\mathbb{R}}.$$

We can integrate the $(\sum M_i - n - 1)$ variables $f_{\rm III}$ to obtain:

$$v_{\mathbb{R}}(U,\varepsilon) = E(U) \int_{\substack{f_{\rm II}\in\mathbb{R}^{n^2} \\ d_{\mathbb{P}}(f_{\rm II},[\det=0])<\varepsilon}} \frac{\mathrm{e}^{-\sum_i \|f_{\rm II}^i\|^2/2}}{\sqrt{2\pi}^{n^2}} |\det f_{\rm II}|^2 \,\mathrm{d}\mathbb{R}^{n^2}.$$

This is $E(U)$ times the probability $v(n,\varepsilon)$ for the linear case. □

Acknowledgements

Steve Smale provided valuable inspiration for us to extend the theory of [32–36] to sparse polynomial systems. He also provided examples on how to eliminate the dependency upon unitary invariance in the dense case.

The paper by Gromov [14] was of foremost importance to this research. To the best of our knowledge, [14] is the only clear exposition available of mixed volume in terms of a wedge of differential forms. We thank Mike Shub for pointing out that reference and for many suggestions.

We would like to thank Jean-Pierre Dedieu for sharing his thoughts with us on Newton iteration in Riemannian and quotient manifolds.

Also, we would like to thank Felipe Acker, Felipe Cucker, Alicia Dickenstein, Ioannis Emiris, Askold Khovanskii, Eric Kostlan, T.Y. Li, Nelson Maculan, Martin Sombra and Jorge P. Zubelli for their suggestions and support.

This paper was written while G.M. was visiting the Liu Bie Ju Center for Mathematics at the City University of Hong Kong. He wishes to thank City U for its generous support.

Some of the material above was previously circulated in the technical report [22].

Appendix A. The coarea formula

Here we give a short proof of the coarea formula, in a version suitable to the setting of this paper. This means we take all manifolds and functions smooth and avoid measure theory as much as possible.

Proposition 5. (1) *Let X be a smooth Riemann manifold, of dimension M and volume form $|\mathrm{d}X|$.*

(2) *Let Y be a smooth Riemann manifold, of dimension n and volume form $|\mathrm{d}Y|$.*

(3) *Let U be an open set of X, and $F: U \to Y$ be a smooth map, such that DF_x is surjective for all x in U.*

(4) *Let $\varphi: X \to \mathbb{R}^+$ be a smooth function with compact support contained in U.*

Then for almost all $z \in F(U)$, $V_z \stackrel{\mathrm{def}}{=} F^{-1}(z)$ is a smooth Riemann manifold, and

$$\int_X \varphi(x) NJ(F;x) |\mathrm{d}X| = \int_{z \in Y} \int_{x \in V_z} \varphi(x) |\mathrm{d}V_z| |\mathrm{d}Y|$$

where $|\mathrm{d}V_z|$ is the volume element of V_z and $NJ(F,x) = \sqrt{\det DF_x^H DF_x}$ is the product of the singular values of DF_x.

By the implicit function theorem, whenever V_z is non-empty, it is a smooth $(N-n)$-dimensional Riemann submanifold of X. By the same reason, $V := \{(z,x) : x \in V_z\}$ is also a smooth manifold.

Let η be the following N-form restricted to V:

$$\eta = \mathrm{d}Y \wedge \mathrm{d}V_z.$$

This is *not* the volume form of V. The proof of Proposition 5 is divided into two steps:

Lemma 8.

$$\int_V \varphi(x) |\eta| = \int_X \varphi(x) NJ(F;x) |\mathrm{d}X|.$$

Lemma 9.

$$\int_V \varphi(x)|\eta| = \int_{z\in Y}\int_{x\in V_z} \varphi(x)|dV_z||dY|.$$

Proof of Lemma 8. We parametrize:

$$\psi : X \to V$$
$$x \mapsto (F(x), x).$$

Then,

$$\int_V \varphi(x)|\eta| = \int_X (\varphi\circ\psi)(x)|\psi^*\eta|.$$

We can choose an orthonormal basis $u_1,\ldots,u_M$ of T_xX such that $u_{n+1},\ldots,u_M \in \ker DF_x$. Then,

$$D\psi(u_i) = \begin{cases} (DF_xu_i, u_i) & i = 1,\ldots,n \\ (0, u_i) & i = n+1,\ldots,M. \end{cases}$$

Thus,

$$\begin{aligned} |\psi^*\eta(u_1,\ldots,u_M)| &= |\eta(D\psi\, u_1,\ldots,D\psi\, u_M)| \\ &= |dY(DF_xu_1,\ldots,DF_xu_n)|\,|dV_z(u_{n+1},\ldots,u_M)| \\ &= |\det DF_x|_{\ker DF_x^\perp}| \\ &= NJ(F,x) \end{aligned}$$

and hence

$$\int_V \varphi(x)|\eta| = \int_X \varphi(x)NJ(F;x)|dX|. \qquad \square$$

Proof of Lemma 9. We will prove this Lemma locally, and this implies the full Lemma through a standard argument (partitions of unity in a compact neighborhood of the support of φ).

Let x_0, z_0 be fixed. A small enough neighborhood of $(x_0, z_0) \subset V_{z_0}$ admits a fibration over V_{z_0} by planes orthogonal to $\ker DF_{x_0}$.

We parametrize:

$$\theta : Y \times V_{z_0} \to V$$

$$(z,x) \mapsto (z, \rho(x,z)),$$

where $\rho(x,z)$ is the solution of $F(\rho)=z$ in the fiber passing through (z_0,x). Remark that $\theta^*\, dY = dY$, and $\theta^*\, dV_z = \rho^* DV_z$. Therefore,

$$\theta^*(dY \wedge dV_z) = dY \wedge (\rho^* dV_z).$$

Also, if one fixes z, then ρ is a parametrization $V_{z_0} \to V_z$. We have:

$$\begin{aligned}\int_V \varphi(x)|\eta| &= \int_{Y\times V_{z_0}} \varphi(\rho(x,z))|\theta^*\eta| \\ &= \int_{z\in Y}\left(\int_{x\in V_{z_0}} \varphi(\rho(x,z))|\rho^*\,\mathrm{d}V_z|\right)|\mathrm{d}Y| \\ &= \int_{z\in Y}\left(\int_{x\in V_z} \varphi(x)|\mathrm{d}V_z|\right)|\mathrm{d}Y|. \qquad \square\end{aligned}$$

The proposition below is essentially Theorem 3, p. 240 of [6]. However, we do not require our manifolds to be compact. We assume all maps and manifolds are smooth, so that we can apply Proposition 5.

Proposition 6. (1) *Let X be a smooth M-dimensional manifold with volume element $|\mathrm{d}X|$.*

(2) *Let Y be a smooth n-dimensional manifold with volume element $|\mathrm{d}Y|$.*

(3) *Let V be a smooth M-dimensional submanifold of $X\times Y$, and let $\pi_1 : V \to X$ and $\pi_2 : V\to Y$ be the canonical projections from $X\times Y$ to its factors.*

(4) *Let Σ' be the set of critical points of π_1, we assume that Σ' has measure zero and that Σ' is a manifold.*

(5) *We assume that π_2 is regular (all points in $\pi_2(V)$ are regular values).*

(6) *For any open set $U\subset V$, for any $x\in X$, we write: $\#_U(x) \stackrel{\text{def}}{=} \#\{\pi_1^{-1}(x)\cap U\}$. We assume that $\int_{x\in X} \#_V(x)|\mathrm{d}X|$ is finite.*

Then, for any open set $U\subset V$,

$$\int_{x\in\pi_1(U)} \#_U(x)|\mathrm{d}X| = \int_{z\in Y}\int_{\substack{x\in V_z\\(x,z)\in U}} \frac{1}{\sqrt{\det DG_x\, DG_x^H}}|\mathrm{d}V_z||\mathrm{d}Y|,$$

where G is the implicit function for $(\hat{x}, G(\hat{x}))\in V$ in a neighborhood of $(x,z)\in V\setminus\Sigma'$. $\square$

Proof. Every $(x,z)\in U\setminus\Sigma'$ admits an open neighborhood such that π_1 restricted to that neighborhood is a diffeomorphism. This defines an open covering of $U\setminus\Sigma'$. Since $U\setminus\Sigma'$ is locally compact, we can take a countable sub-covering and define a partition of unity $(\varphi_\lambda)_{\lambda\in\Lambda}$ subordinated to that sub-covering.

Also, if we fix a value of z, then $(\varphi_\lambda)_{\lambda\in\Lambda}$ becomes a partition of unity for $\pi_1(\pi_1^{-1}(V_z)\cap U)$. Therefore,

$$\begin{aligned}\int_{x\in\pi_1(U)} \#_U(x)|\mathrm{d}X| &= \sum_{\lambda\in\Lambda}\int_{x,z\in \operatorname{Supp}\varphi_\lambda} \varphi_\lambda(x,z)|\mathrm{d}X| \\ &= \sum_{\lambda\in\Lambda}\int_{z\in Y}\int_{x,z\in\operatorname{Supp}\varphi_\lambda} \frac{\varphi_\lambda(x,z)}{NJ(G,x)}|\mathrm{d}X|\end{aligned}$$

$$= \int_{z \in Y} \sum_{\lambda \in \Lambda} \int_{x,z \in \operatorname{Supp} \varphi_\lambda} \frac{\varphi_\lambda(x,z)}{NJ(G,x)} \, |\mathrm{d}X|$$
$$= \int_{z \in Y} \int_{x \in V_z} \frac{1}{NJ(G,x)} \, |\mathrm{d}X|,$$

where the second equality uses Proposition 5 with $\varphi = \varphi_\lambda / NJ$. Since $NJ = \sqrt{\det DG_x \, DG_x^H}$, we are done. □

References

[1] R. Abraham, J.E. Marsden, Foundations of Mechanics, 2nd Edition, Benjamin/Cummings, Advanced Book Program, Reading, MA, 1978.
[2] M.F. Atiyah, Convexity and commuting Hamiltonians, Bull. London Math. Soc. 14 (1) (1982) 1–15.
[3] M.F. Atiyah, Angular momentum, convex polyhedra and algebraic geometry, Proc. Edinburgh Math. Soc. (2) 26 (2) (1983) 121–133.
[4] M. Avriel, Nonlinear Programming Analysis and Methods, Prentice-Hall, Englewood Cliffs, New Jersey, 1976.
[5] D.N. Bernstein, The number of roots of a system of equations, Functional Anal. Appl. 9 9 (3) (1975) (1976) 183–185.
[6] L. Blum, F. Cucker, M. Shub, S. Smale, Complexity and Real Computation, Springer, New York, 1998.
[7] Yu.D. Burago, V.A. Zalgaller, Geometric Inequalities, in: Grundlehren Der Mathematischen Wissenschaften, Vol. 285, Springer, Berlin, 1988.
[8] S.S. Chern, W.H. Chen, K.S. Lam, Lectures on Differential Geometry, World Scientific, River Edge, NJ, 1999.
[9] J.-P. Dedieu, Approximate Solutions of Numerical Problems, Condition Number Analysis and Condition Number Theorem, The Mathematics of Numerical Analysis (Park City, UT, 1995), American Mathematical Society, Providence, RI, 1996, pp. 263–283.
[10] T. Delzant, Hamiltoniens périodiques et images convexes de l'application moment, Bull. Soc. Math. France 116 (3) (1988) 315–339.
[11] J.W. Demmel, Applied Numerical Linear Algebra, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.
[12] A. Edelman, E. Kostlan, How many zeros of a random polynomial are real? Bull. Amer. Math. Soc. (N.S.) 32 (1) (1995) 1–37.
[13] G. Ewald, Combinatorial Convexity and Algebraic Geometry, Springer, New York, 1996.
[14] M. Gromov, Convex Sets and Kähler Manifolds, Advances in Differential Geometry and Topology, World Scientific Publishing, Teaneck, NJ, 1990, pp. 1–38.
[15] V. Guillemin, Moment Maps and Combinatorial Invariants of Hamiltonian T^n-Spaces, Progress in Mathematics, Vol. 122, Birkhuser Boston, Boston, MA, 1994.
[16] V. Guillemin, S. Sternberg, Convexity properties of the moment mapping, Invent. Math. 67 (3) (1982) 491–513.
[17] B. Huber, B. Sturmfels, A polyhedral method for solving sparse polynomial systems, Math. Comp. 64 (212) (1995) 1541–1555.
[18] B.Ja. Kazarnovskiĭ, On zeros of exponential sums, Soviet Math. Doklady 23 (2) (1981) 347–351.
[19] A.G. Khovanskiĭ, Fewnomials, in: Translations of Mathematical Monographs, Vol. 88, American Mathematical Society, Providence, Rhode Island, 1991.
[20] E. Kostlan, On the distribution of roots of random polynomials, from topology to computation; Proceedings of the Smalefest (Berkeley, CA, 1990), Springer, New York, 1993, pp. 419–431.
[21] Li, Tien-Yien, Li, Xing, Finding mixed cells in the mixed volume computation, Found. Comput. Math. 1 (2) (2001) 161–181.
[22] G. Malajovich, J.M. Rojas, Random sparse polynomial systems, Math ArXiv preprint NA/0012104, 2000.

[23] G. Malajovich, J.M. Rojas, Polynomial Systems and the Momentum map, in: F. Cucker, J.M. Rojas (Eds.), Foundations of Computational Mathematics, Proceedings of SMALEFEST 2000, World Scientific, Singapore, 2002, pp. 251–266.
[24] G. Malajovich, J. Zubelli, On the geometry of graeffe iteration, Journal of Complexity 17 (3) (2001) 541–573.
[25] G. Malajovich, J. Zubelli, Tangent graeffe iteration, Numerische Mathematik 89 (4) (2001) 749–782.
[26] D. McDuff, D. Salamon, Introduction to Symplectic Topology, 2nd Edition, Clarendon Press, Oxford University Press, New York, 1998.
[27] A. McLennan, The expected number of real roots of a multihomogeneous system of polynomial equations, Amer. J. Math. 124 (1) (2002) 49–73.
[28] G. Mikhalkin, Amoebas of algebraic varieties, preprint, Mathematics ArXiv AG/0108225, 2001.
[29] M. Passare, H. Rullgård, Amoebas, Monge-Ampère measures, and triangulations of the Newton polytope. Research Reports in Mathematics No. 10, Department of Mathematics, Stockholm University, 2000.
[30] J.M. Rojas, On the average number of real roots of certain random sparse polynomial systems, Lect. Appl. Math. 32 (1996) 689–699.
[31] J.R. Sangwine-Yager, Mixed volumes, Handbook of Convex Geometry, Vol. A, B, North-Holland, Amsterdam, 1993, pp. 43–71.
[32] M. Shub, S. Smale, Complexity of Bezout's Theorem. II. Volumes and Probabilities, Computational Algebraic Geometry (Nice, 1992), Birkhäuser Boston, Boston, MA, 1993, pp. 267–285.
[33] M. Shub, S. Smale, Complexity of Bézout's theorem. I, Geometric aspects, J. Amer. Math. Soc. 6 (2) (1993) 459–501.
[34] M. Shub, S. Smale, Complexity of Bezout's theorem. III. Condition number and packing, J. Complexity 9 (1) (1993) 4–14 (Festschrift for Joseph F. Traub, Part I).
[35] M. Shub, S. Smale, Complexity of Bezout's theorem. V. Polynomial time, Theoret. Comput. Sci. 133 (1) (1994) 141–164.
[36] M. Shub, S. Smale, Complexity of Bezout's theorem. IV. Probability of success; extensions, SIAM J. Numer. Anal. 33 (1) (1996) 128–148.
[37] S. Smale, Topology and mechanics, I, Invent. Math. 10 (1970) 305–331.
[38] J.-M. Souriau, Structure des systèmes dynamiques, Dunod, Paris, Maîtrises de mathématiques, 1970.
[39] A. Turing, On computable numbers, with an application to the Entscheidungsproblem, Proc. London Math. Soc. 2 (42) (1936) 230–265.
[40] J. Verschelde, P. Verlinden, R. Cools, Homotopies exploiting Newton polytopes for solving sparse polynomial systems, SIAM J. Numer. Anal. 31 (3) (1994) 915–930.
[41] O. Viro, Dequantization of Real Algebraic Geometry on Logarithmic Paper, European Congress of Mathematics, Vol. I, Barcelona, 2000; Progress in Mathematics, Vol. 201, Birkhuser, Basel, 2001, pp. 135–146.