Found Comput Math
DOI 10.1007/s10208-013-9155-y

FOUNDATIONS OF COMPUTATIONAL MATHEMATICS
The Journal of the Society for the Foundations of Computational Mathematics

Smale's Fundamental Theorem of Algebra Reconsidered

Diego Armentano · Michael Shub

Received: 9 May 2012 / Revised: 29 January 2013 / Accepted: 8 March 2013

Abstract In his 1981 Fundamental Theorem of Algebra paper Steve Smale initiated the complexity theory of finding a solution of polynomial equations of one complex variable by a variant of Newton's method. In this paper we reconsider his algorithm in the light of work done in the intervening years. Smale's upper bound estimate was infinite average cost. Ours is polynomial in the Bézout number and the dimension of the input. Hence it is polynomial for any range of dimensions where the Bézout number is polynomial in the input size. In particular it is not just for the case that Smale considered but for a range of dimensions as considered by Bürgisser–Cucker, where the max of the degrees is greater than or equal to $n^{1+\epsilon}$ for some fixed ϵ. It is possible that Smale's algorithm is polynomial cost in all dimensions and our main theorem raises some problems that might lead to a proof of such a theorem.

Keywords Polynomial system solving · Homotopy methods · Fundamental Theorem of Algebra · Smale's 17th problem

Mathematics Subject Classification (2010) 65H10 · 65H20 · 65H04 · 14Q20 · 12Y05 · 65Y20 · 68Q25 · 68W40

D. Armentano
Centro de Matemática, Universidad de la República, Iguá 4225, Montevideo 11400, Uruguay
e-mail: diego@cmat.edu.uy

M. Shub (✉)
CONICET, IMAS, Universidad de Buenos Aires, Ciudad Universitaria, Buenos Aires C1428EGA, Argentina
e-mail: shub.michael@gmail.com

Published online: 08 May 2013

1 Introduction and Main Result

1.1 Introduction and Preliminaries

In his 1981 paper [19] Steve Smale initiated the complexity theory of finding a solution of polynomial equations of one complex variable by a variant of Newton's method. More specifically he considered the affine space P_d of monic polynomials of degree d,

$$f(z) = \sum_{i=0}^{d} a_i z^i, \qquad a_d = 1 \quad \text{and} \quad a_i \in \mathbb{C} \quad (i = 0, \dots, d-1).$$

He identified P_d with $\mathbb{C}^d$, with coordinates $(a_0, \dots, a_{d-1}) \in \mathbb{C}^d$. In P_d he considered the polydisk

$$\mathcal{Q}_1 = \bigl\{ f \in \mathsf{P}_d : |a_i| < 1, i = 0, \dots, d-1 \bigr\}$$

to have finite volume and he obtained a probability space by normalizing the volume to 1. The algorithm he analyzed is given by the following. Let $0 < h \le 1$ and let $z_0 = 0$. Inductively define $z_n = T_h(z_{n-1})$ where T_h is the modified Newton's method for f given by $T_h(z) = z - h\frac{f(z)}{f'(z)}$.

His eponymous main theorem was:

Main Theorem *There is a universal polynomial $S(d, 1/\mu)$ and a function $h = h(d, \mu)$ such that for degree d and μ, $0 < \mu < 1$, the following is true with probability $1 - \mu$. Let $x_0 = 0$. Then $x_n = T_h(x_{n-1})$ is defined for all $n > 0$ and x_s is an approximate zero for f where $s = S(d, 1/\mu)$.*

In [19], that x_s is an approximate zero meant that there is an x^* such that $f(x^*) = 0$, $x_n \to x^*$ and $\frac{|f(x_{j+1})|}{|f(x_j)|} < \frac{1}{2}$, for $j \ge s$, where $x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$. That is, x_{k+1} is defined by the usual Newton's method for f. Smale mentions that the polynomial S may be taken to be $\frac{100(d+2)^9}{\mu^7}$. The notion of approximate zero was changed in later papers (see Blum et al. [8] for the later version or Sect. 1.2). The new version incorporates immediate quadratic convergence of Newton's method on an approximate zero. In the remainder of this paper an approximate zero refers to the new version.

Note that $\frac{1}{\mu^7}$ is not finitely integrable, so Smale's initial algorithm was not proven to be finite average time or cost when the upper bound is averaged over the polydisk $\mathcal{Q}_1$ (see Blum et al. [8, pp. 208, Proposition 2]).

A tremendous amount of work has been done in the last 30 years following on Smale's initial contribution, much too much to survey here. Let us mention a few of the main changes. In one variable a lot of work has been done concerning the choice of good starting point z_0 for Smale's algorithm other than zero. See Chaps. 8 and 9 of Blum et al. [8] and references in the commentary on Chap. 9. The latest work in this direction is Kim–Martens–Sutherland [12].

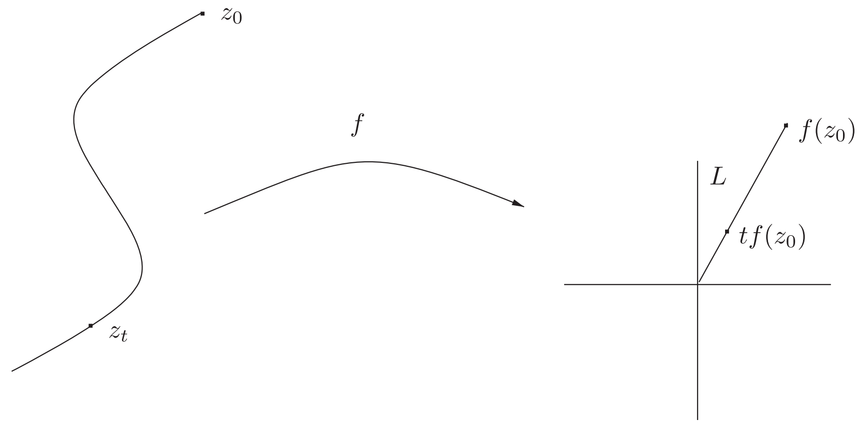


Fig. 1 The curve z_t, $0 \le t \le 1$, is the branch of $f^{-1}(L)$ containing z_0

Smale's algorithm may be given the following interpretation. For $z_0 \in \mathbb{C}$, consider $f_t = f - (1-t)f(z_0)$, for $0 \le t \le 1$. The polynomial f_t has the same degree as f, z_0 is a zero of f_0 and $f_1 = f$. So, we analytically continue z_0 to z_t a zero of f_t. For $t = 1$ we arrive at a zero of f. Newton's method is then used to produce a discrete numerical approximation to the path (f_t, z_t).

If we view f as a mapping from $\mathbb{C}$ to $\mathbb{C}$, then the curve z_t is the branch of the inverse image of the line segment $L = \{tf(z_0) : 0 \le t \le 1\}$, containing z_0. (See Fig. 1.)

Here are several of the changes made in the intervening years. Renegar [13] considered systems of n-complex polynomials in n-variables, without the restriction to be monic. Given a degree d, we let $\mathcal{P}_d$ stands for the vector space of degree d polynomials in n complex variables

$$\mathcal{P}_d = \left\{ f : f(z) = \sum_{\|\alpha\| \le d} a_\alpha z^\alpha \right\},$$

where $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ is a multi-index, $\|\alpha\| = \sum_{k=1}^d \alpha_k$, $z^\alpha = z_1^{\alpha_1} \cdots z_n^{\alpha_n}$, $a_\alpha \in \mathbb{C}$. We have suppressed the n for ease of notation. It should be understood from the context.

For $(d) = (d_1, \dots, d_n)$, let $\mathcal{P}_{(d)} = \mathcal{P}_{d_1} \times \cdots \times \mathcal{P}_{d_n}$ so $f = (f_1, \dots, f_n) \in \mathcal{P}_{(d)}$ is a system of n polynomial equations in n complex variables and f_i has degree d_i.

As Smale's, Renegar's results were not finite average cost or time. In a series of papers Shub and Smale [15–18], made some further changes and achieved enough results for Smale's 17th problem to emerge a reasonable if challenging research goal. Let us recall the 17th problem from Smale [20]:

Problem 17 Solving Polynomial Equations.

Can a zero of n complex polynomial equations in n unknowns be found approximately, on the average, in polynomial time with a uniform algorithm?

In place of $\mathcal{P}_{(d)}$ it is natural to consider $\mathcal{H}_{(d)} = \mathcal{H}_{d_1} \times \cdots \times \mathcal{H}_{d_n}$, where $\mathcal{H}_{d_i}$ is the vector space of homogeneous polynomials of degree d_i in $n+1$ complex variables.

The map

$$\mathrm{i}_{d_i} : \mathcal{P}_{d_i} \to \mathcal{H}_{d_i}, \qquad \mathrm{i}_{d_i}(f)(z_0, \dots, z_n) = z_0^{d_i} f\left(\frac{z_1}{z_0}, \dots, \frac{z_n}{z_0}\right),$$

is an isomorphism and $\mathrm{i} : \mathcal{P}_{(d)} \to \mathcal{H}_{(d)}$ for $\mathrm{i} = (\mathrm{i}_{d_1}, \dots, \mathrm{i}_{d_n})$ is an isomorphism.

For $f \in \mathcal{H}_{(d)}$ and $\lambda \in \mathbb{C}$,

$$f(\lambda\zeta) = \Delta\big(\lambda^{d_i}\big) f(\zeta),$$

where $\Delta(a_i)$ means the diagonal matrix whose ith diagonal entry is a_i. Thus the zeros of $f \in \mathcal{H}_{(d)}$ are now complex lines so may be considered as points in projective space $\mathbb{P}(\mathbb{C}^{n+1})$.

The affine chart

$$\mathrm{j} : \mathbb{C}^n \to \mathbb{P}\big(\mathbb{C}^{n+1}\big), \qquad \mathrm{j}(\zeta_1, \dots, \zeta_n) = \mathbb{C}(1 : \zeta_1 : \cdots : \zeta_n),$$

maps the zeros of $f \in \mathcal{P}_{(d)}$ to zeros of $\mathrm{i}(f) \in \mathcal{H}_{(d)}$. In addition $\mathrm{i}(f)$ may have zeros at infinity, i.e., zeros with $\zeta_0 = 0$.

From now on we consider $\mathcal{H}_{(d)}$ and $\mathbb{P}(\mathbb{C}^{n+1})$. On $\mathcal{H}_{d_i}$ we put a unitarily invariant Hermitian structure which we first encountered in the book [21] by Hermann Weyl and which is sometimes called Weyl, Bombieri–Weyl or Kostlan Hermitian structure depending on the applications considered.

For $\alpha = (\alpha_0, \dots, \alpha_n) \in \mathbb{N}^{n+1}$, $\|\alpha\| = d_i$, and the monomial $z^\alpha = z_0^{\alpha_0} \cdots z_n^{\alpha_n}$, the Weyl Hermitian structure makes $\langle z^\alpha, z^\beta \rangle = 0$, for $\alpha \neq \beta$ and

$$\big\langle z^\alpha, z^\alpha \big\rangle = \binom{d_i}{\alpha}^{-1} = \left(\frac{d_i!}{\alpha_0! \cdots \alpha_n!}\right)^{-1}.$$

On $\mathcal{H}_{(d)}$ we put the product structure

$$\langle f, g \rangle = \sum_{i=1}^{n} \langle f_i, g_i \rangle.$$

On $\mathbb{C}^{n+1}$ we put the usual Hermitian structure

$$\langle x, y \rangle = \sum_{k=0}^{n} x_k \, \overline{y_k}.$$

Given a complex vector space V with Hermitian structure and a vector $0 \neq v \in V$, we let $v^\perp$ be the Hermitian complement of v,

$$v^\perp = \big\{ w \in V : \langle v, w \rangle = 0 \big\}.$$

The subspace $v^\perp$ is a model for the tangent space, $T_v\mathbb{P}(V)$, of the projective space $\mathbb{P}(V)$ at the equivalence class of v (which we also denote by v).

The tangent space $T_v\mathbb{P}(V)$ inherits an Hermitian structure from $\langle\cdot,\cdot\rangle$ by the formula

$$\langle w_1, w_2\rangle_v = \frac{\langle w_1, w_2\rangle}{\langle v, v\rangle},$$

where $w_1, w_2 \in v^\perp$ represent the tangent vectors at $T_v\mathbb{P}(V)$.

This Hermitian structure which is well defined is called the Fubini–Study Hermitian structure.

The group of unitary transformations $\mathcal{U}(n+1)$ acts on $\mathcal{H}_{(d)}$ and $\mathbb{C}^{n+1}$ by $f \mapsto f \circ U^{-1}$ and $\zeta \mapsto U\zeta$ for $U \in \mathcal{U}(n+1)$.

This unitary action preserves the Hermitian structure on $\mathcal{H}_{(d)}$ and $\mathbb{C}^{n+1}$, see Blum et al. [8]. That is, for $U \in \mathcal{U}(n+1)$,

$$\begin{aligned} \left\langle f \circ U^{-1}, g \circ U^{-1}\right\rangle &= \langle f, g\rangle \quad \text{for } f, g \in \mathcal{H}_{(d)}; \\ \left\langle U\zeta, U\zeta'\right\rangle &= \left\langle \zeta, \zeta'\right\rangle \quad \text{for } \zeta, \zeta' \in \mathbb{C}^{n+1}. \end{aligned}$$

The zeros of λf and f for $0 \neq \lambda \in \mathbb{C}$ are the same, and we may consider the space $\mathbb{P}(\mathcal{H}_{(d)})$. Now the space of problem instances is compact and the space $\mathbb{P}(\mathbb{C}^{n+1})$ is compact as well. The set $\mathbb{P}(\mathcal{H}_{(d)})$ has a unitarily invariant Hermitian structure which gives rise to a volume form of finite volume $\frac{\pi^{N-1}}{\Gamma(N)}$, where $N = \dim \mathcal{H}_{(d)}$.

The average of a function $\phi : \mathbb{P}(\mathcal{H}_{(d)}) \to \mathbb{R}$ is

$$\mathbb{E}_{\mathbb{P}(\mathcal{H}_{(d)})}(\phi) = \frac{1}{\mathrm{vol}(\mathbb{P}(\mathcal{H}_{(d)}))} \int_{f \in \mathbb{P}(\mathcal{H}_{(d)})} \phi(f)\,\mathrm{d}f = \frac{\Gamma(N)}{\pi^{N-1}} \int_{f \in \mathbb{P}(\mathcal{H}_{(d)})} \phi(f)\,\mathrm{d}f.$$

If ϕ is induced by a homogeneous function $\phi : \mathcal{H}_{(d)} \to \mathbb{R}$ of degree zero, that is, $\phi(\lambda f) = \phi(f)$, $\lambda \in \mathbb{C} - \{0\}$, then we may also compute this average with respect to the Gaussian measure on $(\mathcal{H}_{(d)}, \langle\cdot,\cdot\rangle)$, that is,

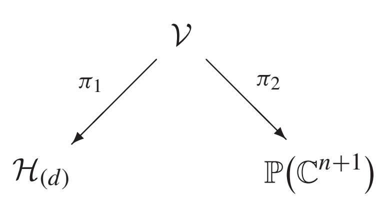
$$\mathbb{E}_{\mathcal{H}_{(d)}}(\phi) = \frac{1}{(2\pi)^N} \int_{\mathcal{H}_{(d)}} \phi(f) e^{-\|f\|^2/2}\,\mathrm{d}f. \tag{1}$$

It is this approach via the Gaussians above defined on $\mathcal{H}_{(d)}$ and the Fubini–Study Hermitian structure and volume form on $\mathbb{P}(\mathbb{C}^{n+1})$, which we take in this paper. The quantities we define on $\mathcal{H}_{(d)}$ are homogeneous of degree zero, thus are defined on $\mathbb{P}(\mathcal{H}_{(d)})$ and benefit from the compactness of this space and of $\mathbb{P}(\mathbb{C}^{n+1})$. While averages over systems of equations may be carried out in the vector space $\mathcal{H}_{(d)}$.

The solution variety

$$\mathcal{V} = \left\{(f, x) \in \left(\mathcal{H}_{(d)} - \{0\}\right) \times \mathbb{P}\left(\mathbb{C}^{n+1}\right) : f(x) = 0\right\}$$

is a central object of study. It is equipped with two projections



The solution variety $\mathcal{V}$ also has a projective version, namely,

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

1.2 Homotopy Methods

Homotopy methods for the solution of a system $f \in \mathcal{H}_{(d)}$ proceed as follows. Choose $(g,\zeta) \in \mathcal{V}$ a known pair. Connect g to f by a C^1 curve f_t in $\mathcal{H}_{(d)}$, $0 \le t \le 1$, such that $f_0 = g$, $f_1 = f$, and continue $\zeta_0 = \zeta$ to ζ_t such that $f_t(\zeta_t) = 0$, so that $f_1(\zeta_1) = 0$. The critical values of the projection of $\mathcal{V}$ on $\mathcal{H}_{(d)} - \{0\}$ are an algebraic subvariety, Σ, of $\mathcal{H}_{(d)} - \{0\}$ of complex codimension 1, called the discriminant variety. By the transversality theorem (see Abraham–Robbin [2]) a generic set of C^1 curves f_t do not intersect Σ. If a curve is in this generic set and $f_0(\zeta_0) = 0$, then by the implicit function theorem we may continue ζ_0 to ζ_t, $0 \le t \le 1$, such that $f_t(\zeta_t) = 0$. See Smale [19] for this type of argument. Indeed almost all "straight line" paths in $\mathcal{H}_{(d)}$ do not intersect Σ, again by a transversality argument, so if ζ_0 is a nondegenerate zero of g then for almost all f, ζ_0 may be continued to a zero of f along the curve $f_t = (1-t)g + tf$. We do not use this generality in this paper so we leave the above assertions as a sketch. In Proposition 1 we prove a precise version of the fact that the homotopies we consider in this paper may be almost always continued.

Now homotopy methods numerically approximate the path (f_t, ζ_t). One way to accomplish the approximation is via (projective) Newton's methods. Given an approximation x_t to ζ_t define

$$x_{t+\Delta t} = N_{f_{t+\Delta t}}(x_t),$$

where

$$N_f(x) = x - \left(Df(x)|_{x^\perp}\right)^{-1} f(x).$$

That x_t is an approximate zero of f_t associated with the zero ζ_t means that the sequence of Newton iterative $N_{f_t}^k(x_t)$ converges immediately quadratically to ζ_t.

The main result of Shub [14] is that Δt may be chosen so that $t_0 = 0$, $t_k = t_{k-1} + \Delta t_k$, x_{t_k} is an approximate zero of f_{t_k} with associated zero ζ_{t_k}, and $t_K = 1$ for

$$K = K(f,g,\zeta) \le C D^{3/2} \int_0^1 \mu(f_t,\zeta_t) \left\|(\dot{f}_t, \dot{\zeta}_t)\right\|_{(f_t,\zeta_t)} \mathrm{d}t. \tag{2}$$

Here C is a universal constant, $D = \max d_i$,

$$\mu(f,\zeta) = \|f\| \left\| \left(Df(\zeta)|_{\zeta^\perp}\right)^{-1} \Delta\left(\|\zeta\|^{d_i - 1}\sqrt{d_i}\right)\right\|$$

is the condition number of f at ζ, and

$$\left\|(\dot{f}_t, \dot{\zeta}_t)\right\|_{(f_t,\zeta_t)} = \left(\|\dot{f}_t\|_{f_t^2} + \|\dot{\zeta}_t\|_{\zeta_t^2}\right)^{1/2}$$

is the norm of the tangent vector to the projected curve in (f_t, ζ_t) in $\mathcal{V}_{\mathbb{P}} \subset \mathbb{P}(\mathcal{H}_{(d)}) \times \mathbb{P}(\mathbb{C}^{n+1})$. The choice of Δt_k is made explicit in Dedieu–Malajovich–Shub [10] and Beltrán [4].

In $\mathcal{V}_{\mathbb{P}}$, $\|\dot{\zeta}_t\|_{\zeta_t} \le \mu(f_t, \zeta_t)\|\dot{f}_t\|_{f_t}$, so the estimates (2) may be bounded from above by

$$K(f, g, \zeta) \le C D^{3/2} \int_0^1 \mu(f_t, \zeta_t)^2 \|\dot{f}_t\|_{f_t} \,\mathrm{d}t, \tag{3}$$

for a perhaps different universal constant C.

Finally in the case of straight line homotopy $\|\dot{f}_t\|_{f_t} = \frac{\sin(\theta)\|f_0\|\|f_1\|}{\|f_t\|^2}$, where θ is the angle between f_0 and f_1. So (3) may be rewritten as

$$K(f, g, \zeta) \le C D^{3/2} \sin(\theta)\|f_0\|\|f_1\| \int_0^1 \frac{\mu(f_t, \zeta_t)^2}{\|f_t\|^2} \,\mathrm{d}t, \tag{4}$$

see Bürgisser–Cucker [9], where (4) is a principal part of the analysis and where the increments Δ_{t_k}, which exhibit the right-hand side of (4) as an upper bound, are also made explicit.

Much attention has been devoted to studying the right hand of (4), for a good starting point (g, ζ).

In Beltrán–Pardo [5], an affirmative probabilistic solution to Smale's 17th problem is proven. They prove that a random point (g, ζ) is good in the sense that with random fixed starting point $(g, \zeta) = (f_0, \zeta_0)$ the average value of the right hand side of (4) is bounded by $O(nN)$. Moreover, Beltrán and Pardo show how to pick a random starting point starting from a random $n \times (n+1)$ matrix.

In [9], Bürgisser–Cucker exhibit a deterministic algorithm for Smale's 17th problem which is polynomial average cost, except for a narrow range of dimensions. More precisely:

> *There is a deterministic real number algorithm that on input $f \in \mathcal{H}_{(d)}$ computes an approximate zero of f in average time $N^{O(\log\log N)}$, where $N = \dim \mathcal{H}_{(d)}$ measures the size of the input f. Moreover, if we restrict data to polynomials satisfying*
>
> $$D \le n^{\frac{1}{1+\varepsilon}}, \quad \textit{or} \quad D \ge n^{1+\varepsilon},$$
>
> *for some fixed $\varepsilon > 0$, then the average time of the algorithm is polynomial in the input size N.*

So Smale's 17th problem in its deterministic form remains open for a narrow range of degrees and variables.

The case $D \le n^{\frac{1}{1+\varepsilon}}$ is dealt with by Bürgisser–Cucker by constructing a good starting point for a homotopy method while the case $D \ge n^{1+\varepsilon}$ is dealt with differently. Our Theorem 3 shows that we may use the homotopy method suggested by Smale's algorithm, described in the next section, in this range of dimensions and conclude a polynomial result as well.

1.3 Smale's Algorithm Reconsidered

Smale's 1981 algorithm depends on $f(0)$, so there is no fixed starting point for all systems. Given $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ we define for $f \in \mathcal{H}_{(d)}$ the straight line segment $f_t \in \mathcal{H}_{(d)}$,

$0 \le t \le 1$, by

$$f_t = f - (1-t)\Delta\left(\frac{\langle \cdot, \zeta\rangle^{d_i}}{\langle \zeta, \zeta\rangle^{d_i}}\right) f(\zeta).$$

So $f_0(\zeta) = 0$ and $f_1 = f$. Therefore we may apply homotopy methods to this line segment. (Here $\Delta(\frac{\langle \cdot,\zeta\rangle^{d_i}}{\langle \zeta,\zeta\rangle^{d_i}})$ must be understood as a matrix of functions and $f(\zeta)$ as a vector of constants which multiply the functions according to matrix vector multiplication.)

Note that if we restrict f to the affine chart $\zeta + \zeta^{\perp}$ then

$$f_t(z) = f(z) - (1-t) f(\zeta),$$

and if we take $\zeta = (1, 0, \dots, 0)$ and $n = 1$ we recover Smale's algorithm.

There is no reason to single out $\zeta = (1, 0, \dots, 0)$. Since the unitary group acts by isometries on $\mathbb{P}(\mathcal{H}_{(d)})$, $\mathbb{P}(\mathbb{C}^{n+1})$, $\mathcal{V}$ and $\mathcal{V}_{\mathbb{P}}$, and preserves μ and is transitive on $\mathbb{P}(\mathbb{C}^{n+1})$, all the points ζ yield algorithms with the same average cost.

Note that if we let

$$\mathcal{V}_\zeta = \left\{ f \in \mathcal{H}_{(d)} : f(\zeta) = 0 \right\},$$

then

$$f_0 = f - \Delta\left(\frac{\langle \cdot, \zeta\rangle^{d_i}}{\langle \zeta, \zeta\rangle^{d_i}}\right) f(\zeta)$$

is the orthogonal projection $\Pi_\zeta(f)$ of f on $\mathcal{V}_\zeta$. This follows from the reproducing kernel property of the Weyl Hermitian product on $\mathcal{H}_{d_i}$, namely,

$$\left\langle g, \langle \cdot, \zeta\rangle^{d_i} \right\rangle = g(\zeta), \tag{5}$$

for all $g \in \mathcal{H}_{d_i}$ $(i = 1, \dots, n)$. In particular $\|\langle \cdot, \zeta\rangle^{d_i}\| = \|\zeta\|^{d_i}$.

Then,

$$\left\| f - \Pi_\zeta(f) \right\| = \left\| \Delta\left(\|\zeta\|^{-d_i}\right) f(\zeta) \right\|,$$

while

$$\left\| \Pi_\zeta(f) \right\| = \left(\|f\|^2 - \left\| \Delta\left(\|\zeta\|^{-d_i}\right) f(\zeta) \right\|^2 \right)^{1/2}.$$

Let $\Phi : \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1}) \times [0, 1] \to \mathcal{V}$ be the map given by

$$\Phi(f, \zeta, t) = (f_t, \zeta_t), \tag{6}$$

where

$$f_t = (1-t)\Pi_\zeta(f) + tf,$$

that is,

$$f_t = f - (1-t)\Delta\left(\frac{\langle \cdot, \zeta\rangle^{d_i}}{\langle \zeta, \zeta\rangle^{d_i}}\right) f(\zeta),$$

and ζ_t is the homotopy continuation of ζ along the path f_t.

Proposition 1 *For almost every $f \in \mathcal{H}_{(d)}$, the set of $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ such that Φ is defined for all $t \in [0,1]$ has full measure. Moreover, for every $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$, the set of $f \in \mathcal{H}_{(d)}$ such that Φ is defined for all $t \in [0,1]$ has full measure.*

Remark In fact, the proof also shows that the complement of the set (f, ζ) such that Φ is defined for all $t \in [0,1]$ is a real algebraic set. The proof of Proposition 1 is in Sect. 2.

The norm of $\dot{f}_t$ is given now by the formula

$$\|\dot{f}_t\|_{f_t} = \frac{\|f_0\|\|f_1\|\sin(\theta)}{\|f_t\|^2} = \frac{\|\Pi_\zeta(f)\|\|f - \Pi_\zeta(f)\|}{\|f_t\|^2}$$
$$= \frac{(\|f\|^2 - \|\Delta(\|\zeta\|^{-d_i})f(\zeta)\|^2)^{1/2}\|\Delta(\|\zeta\|^{-d_i})f(\zeta)\|}{\|f_t\|^2}.$$

Let $\mathcal{T}(f,\zeta) = K(f, \Pi_\zeta(f), \zeta)$ and $\mathcal{T}_\zeta(f) = \mathcal{T}(f,\zeta)$. Then, the average cost of this algorithm satisfies

Proposition 2

$$\mathbb{E}_{\mathcal{H}_{(d)}}(\mathcal{T}_\zeta) = \mathbb{E}_{\mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1})}(\mathcal{T}) \le (\mathrm{I}),$$

where

$$(\mathrm{I}) = \frac{CD^{3/2}}{(2\pi)^N \operatorname{vol}(\mathbb{P}(\mathbb{C}^{n+1}))} \int_{f \in \mathcal{H}_{(d)}} \int_{\zeta \in \mathbb{P}(\mathbb{C}^{n+1})} \int_{t \in [0,1]} \frac{\mu(f_t, \zeta_t)^2}{\|f_t\|^2}$$
$$\times \left(\|f\|^2 - \left\|\Delta\left(\|\zeta\|^{-d_i}\right) f(\zeta)\right\|^2\right)^{1/2} \left\|\Delta\left(\|\zeta\|^{-d_i}\right) f(\zeta)\right\| e^{-\|f\|^2/2}\, df\, d\zeta\, dt.$$

As we have mentioned above, it is easily seen, by unitary invariance of all the quantities involved, that the average $\mathbb{E}_{\mathcal{H}_{(d)}}(\mathcal{T}_\zeta)$ on $\mathcal{H}_{(d)}$ is independent of ζ and equal to $\mathbb{E}_{\mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1})}(\mathcal{T})$. This argument proves the first equality of this proposition. The inequality follows immediately from the definition of $\mathcal{T}(f,\zeta)$.

What is gained by letting ζ vary and dividing by $\operatorname{vol}(\mathbb{P}(\mathbb{C}^{n+1}))$ is that a new way to see the integral leads to new theorems and interesting questions.

Suppose η is a non-degenerate zero of $h \in \mathcal{H}_{(d)}$. We define the basin of η, $B(h,\eta)$, as those $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ such that the zero ζ of $h - \Delta(\frac{\langle \cdot, \zeta \rangle^{d_i}}{\langle \zeta, \zeta \rangle^{d_i}})h(\zeta)$ continues to η for the homotopy h_t. From the proof of Proposition 1 we observe that the basins are open sets.

Let (I) be the expression defined in Proposition 2. Then, the main result of this paper is:

Theorem 1 (Main Theorem)

$$(\mathrm{I}) = \frac{CD^{3/2}}{(2\pi)^N} \int_{h \in \mathcal{H}_{(d)}} \left[\sum_{\eta / h(\eta) = 0} \frac{\mu^2(h,\eta)}{\|h\|^2} \Theta(h,\eta) \right] e^{-\|h\|^2/2}\, dh,$$

where

$$\Theta(h,\eta) = \frac{1}{\mathrm{vol}(\mathbb{P}(\mathbb{C}^{n+1}))} \int_{\zeta \in B(h,\eta)} \theta_h(\zeta)\,\mathrm{d}\zeta,$$

$$\theta_h(\zeta) = \left(\|h\|^2 - \left\|\Delta\left(\|\zeta\|^{-d_i}\right)h(\zeta)\right\|^2\right)^{1/2} \times \left\|\Delta\left(\|\zeta\|^{-d_i}\right)h(\zeta)\right\| \mathcal{I}_n\left(\left\|\Delta\left(\|\zeta\|^{-d_i}\right)h(\zeta)\right\|^2/2\right),$$

and $\mathcal{I}_n(\alpha) = \int_0^1 e^{(1-t^{-2})\alpha} t^{-2n-1}\,\mathrm{d}t$.

From Proposition 1 we find that the function Θ, defined in the statement of Theorem 1, is defined for almost every pair $(h,\eta) \in \mathcal{V}$.

Summing $\Theta(h,\eta)$ over the roots of h we let $\hat{\Theta}(h) = \sum_{\eta/h(\eta)=0} \Theta(h,\eta)$, and for almost all h we have

$$\hat{\Theta}(h) = \frac{1}{\mathrm{vol}(\mathbb{P}(\mathbb{C}^{n+1}))} \int_{\zeta \in \mathbb{P}(\mathbb{C}^{n+1})} \theta_h(\zeta)\,\mathrm{d}\zeta. \tag{7}$$

That is, $\hat{\Theta}(h) = \|\theta_h\|_{L^1}$.

More generally, for $p > 1$, consider the L_p-norm of θ_h

$$\|\theta_h\|_{L^p}^p = \frac{1}{\mathrm{vol}(\mathbb{P}(\mathbb{C}^{n+1}))} \int_{\zeta \in \mathbb{P}(\mathbb{C}^{n+1})} \theta_h(\zeta)^p\,\mathrm{d}\zeta. \tag{8}$$

The next theorem shows that the average of $\|\theta_h\|_{L^p}^p$ over $\mathcal{H}_{(d)}$, for all $p \geq 1$, is polynomial in N.

Theorem 2

$$\mathbb{E}_{\mathcal{H}_{(d)}}\left(\|\theta_h\|_{L^p}^p\right) \leq \frac{2^p}{p} \frac{\Gamma(N-n+p/2)}{\Gamma(N-n)} \frac{\Gamma(n+p/2)}{\Gamma(n)}.$$

(*The equality holds for* $p = 1$.)

Theorem 3

$$(\mathrm{I}) \leq 18 C D^{3/2} \mathcal{D} n^{3/2} N^{3/2}.$$

That is, (I) is polynomial in the Bézout number and the input size, N, and polynomial in the input size alone for any range of dimensions where the Bézout number $\mathcal{D}$ is polynomial in N.

Since our method of proof of Theorem 3 relies on Theorem 2, where the basins are not taking into account, it is possible that Smale's algorithm is polynomial cost in all dimensions.

Understanding the basins better might lead to a proof of such a theorem. The integral

$$\frac{1}{(2\pi)^N} \int_{h \in \mathcal{H}_{(d)}} \sum_{\eta/h(\eta)=0} \frac{\mu^2(h,\eta)}{\|h\|^2} e^{-\|h\|^2/2}\,\mathrm{d}h \leq \frac{e(n+1)}{2} \mathcal{D},$$

where $\mathcal{D} = d_1 \cdots d_n$ is the Bézout number (see Bürgisser–Cucker [9]). So the question is how does the factor $\Theta(h, \eta)$ affect the integral.[1]

From Theorem 2, the expected value of $\hat{\Theta}(h) = \|\theta_h\|_{L^1}$ is controlled, then, if the integral on the $\mathcal{D}$ basins are reasonably balanced, the factor of $\mathcal{D}$ in Theorem 3 and the integral above may cancel.

Remark The proof of Theorem 1 involves complicated formulas which exhibit enormous cancellations. We do not have a good explanation for these cancellations.

At the end of the paper we present some numerical experiments with $n = 1$ and $d = 7$ which were done by Carlos Beltrán on the Altamira super computer at the Universidad de Cantabria (partially supported by MTM2010-16051 Spanish Ministry of Science and Innovation MICINN). We thank Carlos and the Universidad de Cantabria. We also thanks Gregorio Malajovich for many useful discussions and Santiago Laplagne for having done some more experiments. It would be interesting to see more experimental data. The proof of Theorem 1 is in Sect. 3, and the proofs of Theorems 2 and 3 are in Sect. 4.

2 Proof of Proposition 1

For the proof of Proposition 1 we need a technical lemma.

Lemma 1 *Let E be a vector bundle over B, F be finite dimensional vector space, and consider the trivial vector bundle $F \times B$. Let $\varphi : F \times B \to E$ be a smooth bundle map, covering the identity in B, which is a fiberwise surjective linear map. Then, φ is a surjective submersion.*

The proof is left to the reader.

Recall that $\Phi : \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1}) \times [0, 1] \to \mathcal{V}$ is the map given by

$$\Phi(f, \zeta, t) = (f_t, \zeta_t),$$

where

$$f_t = f - (1 - t)\Delta\left(\frac{\langle \cdot, \zeta\rangle^{d_i}}{\langle \zeta, \zeta\rangle^{d_i}}\right) f(\zeta),$$

and ζ_t is the homotopy continuation of ζ along the path f_t.

[1] In an earlier version of this paper we asked:

(d) Evaluate or estimate

$$\int_{\zeta \in \mathbb{P}(\mathbb{C}^{n+1})} \frac{1}{\|\Delta(\|\zeta\|^{-d_i})h(\zeta)\|^{2n-1}} e^{\|\Delta(\|\zeta\|^{-d_i})h(\zeta)\|^2/2}\, d\zeta.$$

It is easy to see, as in the proof of Theorem 2, that the expected value of this integral over $\mathcal{H}_{(d)}$ is infinite. In Fernandez–Pardo [11] the authors consider the more meaningful average over the unit sphere of $\mathcal{H}_{(d)}$ and get a precise formula for it. Our initial goal in asking question (d) was to get an upper-bound estimate of the integral we now evaluate in Theorem 2.

This map is defined at (f, ζ, t) provided that $\operatorname{rank}(Df_s(\zeta_s)|_{\zeta_s^{\perp}}) = n$, for all $s \in [0, t]$.

Let $\overline{K}$ be the vector bundle over $\mathbb{C}^{n+1} - \{0\}$ with fiber $\overline{K}_z = L(z^{\perp}, \mathbb{C}^n)$, $z \in \mathbb{C}^{n+1} - \{0\}$, where $L(z^{\perp}, \mathbb{C}^n)$ is the space of linear transformations from $z^{\perp}$ to $\mathbb{C}^n$. For $k = 0, \ldots, n$, let $\overline{K_k}$ be the sub-bundle of rank k linear transformations. Note that $\overline{K_k}$ has $(n-k)^2$ complex codimension (cf. [3]). These sub-bundles define a stratification of the bundle $\overline{K}$.

Lemma 2 *Let $\Omega^{(0)}$ be the set of pairs $(f, \zeta) \in \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1})$ such that $\operatorname{rank}(Df_0(\zeta)|_{\zeta^{\perp}}) < n$, where $(f_0, \zeta) = \Phi(f, \zeta, 0)$. Then $\Omega^{(0)}$ is a stratified set of smooth manifolds of complex codimension $(n-k)^2$, for $k = 0, \ldots, n-1$.*

Proof Let $\hat{\Omega}^{(0)}$ be the inverse image of $\Omega^{(0)}$ under the canonical projection $\mathcal{H}_{(d)} \times (\mathbb{C}^{n+1} - \{0\}) \to \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1})$.

Let $\varphi : \mathcal{H}_{(d)} \times (\mathbb{C}^{n+1} - \{0\}) \to \overline{K}$ be the map $\varphi(f, \zeta) = Df(\zeta)|_{\zeta^{\perp}}$. For each $k = 0, \ldots, n-1$, let $\hat{\Omega}_k^{(0)} = \varphi^{-1}(\overline{K_k})$. Since $Df_0(\zeta)|_{\zeta^{\perp}} = Df(\zeta)|_{\zeta^{\perp}}$, then $\hat{\Omega}^{(0)} = \bigcup_{k=0}^{n-1} \hat{\Omega}_k^{(0)}$.

Claim: φ is transversal to $\overline{K_k}$ for $k = 0, \ldots, n-1$:

Note that $\varphi(f, \cdot) : \mathbb{C}^{n+1} - \{0\} \to \overline{K}$ is a section of the vector bundle $\overline{K}$ for each $f \in \mathcal{H}_{(d)}$. Moreover, for each $\zeta \in \mathbb{C}^{n+1} - \{0\}$, the linear map $\varphi(\cdot, \zeta) : \mathcal{H}_{(d)} \to \overline{K}_{\zeta}$ is a surjective linear map. This fact follows by construction: given $L \in \overline{K}_{\zeta} = L(\zeta^{\perp}, \mathbb{C}^n)$, let $\tilde{L} \in L(\mathbb{C}^{n+1}, \mathbb{C}^n)$ be any linear extension of L to $\mathbb{C}^{n+1}$. Then, the system $f = \Delta(\frac{\langle \cdot, \zeta \rangle^{d_i - 1}}{\langle \zeta, \zeta \rangle^{d_i - 1}}) \tilde{L}(\cdot) \in \mathcal{H}_{(d)}$ satisfy $Df(\zeta)|_{\zeta^{\perp}} = L$. Then, the claim follows from Lemma 1.

Since φ is transversal, we conclude that the inverse image of a stratification is a stratification of the same codimension (cf. [3]). That is, $\hat{\Omega}^{(0)}$ is a stratification of smooth submanifolds of complex codimension $(n-k)^2$, for $k = 0, \ldots, n-1$. (The leaves of the strata are not analytic submanifolds since their definition relies on complex conjugation but they are real analytic submanifolds whose tangent space is modelled by a complex vector space at each point.)

Moreover, since each strata $\hat{\Omega}_k^{(0)}$ contains the fiber of the canonical projection $\mathcal{H}_{(d)} \times (\mathbb{C}^{n+1} - \{0\}) \to \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1})$, then, its image, $\Omega_k^{(0)}$, is a smooth manifold of codimension $(n-k)^2$, and the lemma follows. □

One can define the homotopy continuation of the pair $(f, \zeta) \in \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1})$ for all $t \in [0, 1]$ whenever $(f, \zeta) \notin \Omega^{(0)}$ and lies outside the subset of pairs such that there exist $(w, t) \in \mathbb{P}(\mathbb{C}^{n+1}) \times (0, 1]$ satisfying the following equations:

$$f(w) = (1-t)\Delta\left(\frac{\langle w, \zeta \rangle^{d_i}}{\langle \zeta, \zeta \rangle^{d_i}}\right) f(\zeta), \quad \text{and} \quad \operatorname{rank}\left(Df_t(w)|_{w^{\perp}}\right) < n.$$

Note that, since f_t is homogeneous, $\operatorname{rank}(Df_t(w)|_{w^{\perp}})$ is well defined on $w \in \mathbb{P}(\mathbb{C}^{n+1})$.

Differentiating f_t we get

$$Df_t(w) = Df(w) - (1-t)\Delta\left(\frac{d_i\langle w,\zeta\rangle^{d_i-1}\langle\cdot,\zeta\rangle}{\langle\zeta,\zeta\rangle^{d_i}}\right)f(\zeta).$$

Therefore, taking $s = 1-t$, we conclude that one can define the homotopy continuation of the pair $(f,\zeta) \in \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1})$ for all $t \in [0,1]$ whenever $(f,\zeta) \notin \Omega^{(0)}$ and lies outside the subset of pairs such that there exist $(w,s) \in \mathbb{P}(\mathbb{C}^{n+1}) \times [0,1)$ satisfying, for some $k = 0,\dots,n-1$, the following equations:

$$\Delta\big(\langle\zeta,\zeta\rangle^{d_i}\big)f(w) - s\Delta\big(\langle w,\zeta\rangle^{d_i}\big)f(\zeta) = 0, \tag{9}$$

$$\operatorname{rank}\big(\big[\Delta\big(\langle\zeta,\zeta\rangle^{d_i}\big)Df(w) - s\Delta\big(d_i\langle w,\zeta\rangle^{d_i-1}\langle\cdot,\zeta\rangle\big)f(\zeta)\big]\big|_{w^\perp}\big) = k. \tag{10}$$

Let $\Sigma' \subset \mathcal{V}$ be the set of critical points of the projection $\pi_1 : \mathcal{V} \to \mathcal{H}_{(d)}$. Recall that $\Sigma = \pi_1(\Sigma') \subset \mathcal{H}_{(d)}$ is the discriminant variety. If $(f,w) \in \Sigma'$ then w is a degenerate root of f, that is, $\operatorname{rank}(Df(w)|_{w^\perp}) < n$ (cf. Blum et al. [8]).

Note that if $f \in \Sigma$ then every $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ satisfies equations (9) and (10) for $s = 0$ and $w \in \mathbb{P}(\mathbb{C}^{n+1})$ a degenerate root of f. Hence, it is natural to remove the discriminant variety Σ and the case $s = 0$ from this discussion.

Lemma 3 *Let $\Lambda \subset (\mathcal{H}_{(d)} - \Sigma) \times \mathbb{P}(\mathbb{C}^{n+1}) \times \mathbb{P}(\mathbb{C}^{n+1}) \times (0,1)$ be the set of tuples (f,ζ,w,s) such that Eqs.* (9) *and* (10) *holds for some $k = 0,\dots,n-1$. Then, Λ is stratified set of smooth manifolds of real codimension $2(n+(n-k)^2)$ for $k = 0,\dots,n-1$.*

Proof Similar to the preceding proof, for each $k = 0,\dots,n-1$, we consider the set $\hat{\Lambda}_k \subset (\mathcal{H}_{(d)} - \Sigma) \times (\mathbb{C}^{n+1} - \{0\}) \times (\mathbb{C}^{n+1} - \{0\}) \times (0,1)$ of tuples (f,ζ,w,s) such that equations (9) and (10) holds.

Let $(f,\zeta,w,s) \in \hat{\Lambda}_k$ for some $k \in \{0,\dots,n-1\}$. Since $f \notin \Sigma$ then $\langle w,\zeta\rangle \neq 0$. Therefore from (9), equation (10) takes the form

$$\operatorname{rank}\big(\big(\langle w,\zeta\rangle Df(w) - \Delta(d_i)f(w)\langle\cdot,\zeta\rangle\big)\big|_{w^\perp}\big) = k,$$

for $k = 0,\dots,n-1$.

Let

$$F = (F_1,F_2) : (\mathcal{H}_{(d)} - \Sigma) \times \big(\mathbb{C}^{n+1} - \{0\}\big) \times \big(\mathbb{C}^{n+1} - \{0\}\big) \times (0,1) \to \mathbb{C}^n \times \overline{K}$$

be the map defined by

$$F_1(f,\zeta,w,s) = \Delta\big(\langle\zeta,\zeta\rangle^{d_i}\big)f(w) - s\Delta\big(\langle w,\zeta\rangle^{d_i}\big)f(\zeta) \in \mathbb{C}^n,$$

$$F_2(f,\zeta,w,s) = \big(w, \big(\langle w,\zeta\rangle Df(w) - \Delta(d_i)f(w)\langle\cdot,\zeta\rangle\big)|_{w^\perp}\big) \in \overline{K}.$$

Note that $\hat{\Lambda}_k = F^{-1}(\{0\} \times \overline{K_k})$.

Claim: F is transversal to $\{0\} \times \overline{K_k}$:

In fact, what we prove is that DF is surjective at any point (f,ζ,w,s) which F maps into $\{0\} \times \overline{K_k}$, for any $k = 0,\dots,n-1$, that is, any point in $\hat{\Lambda}_k$.

Recall that $\mathcal{V}_\zeta = \{f \in \mathcal{H}_{(d)} : f(\zeta) = 0\}$. Consider the orthogonal decomposition $\mathcal{H}_{(d)} = \mathcal{V}_\zeta \oplus C_\zeta$, where $C_\zeta = \mathcal{V}_\zeta{}^\perp$.

Let $(f, \zeta, w, s) \in \hat{\Lambda}_k$. We first prove that $DF_1(f, \zeta, w, s)|_{C_\zeta} : C_\zeta \to \mathbb{C}^n$ is surjective.

Note that the linear map $\xi : \mathbb{C}^n \to C_\zeta$ given by $\xi(a) = \Delta(\frac{\langle \cdot, \zeta\rangle^{d_i}}{\langle \zeta, \zeta\rangle^{d_i}})a$, is an isomorphism, where $\xi^{-1} : C_\zeta \to \mathbb{C}^n$ is given by $\xi^{-1}(f) = f(\zeta)$. Then, under this identification, the restriction to C_ζ of the derivative of F_1 is the linear map given by

$$DF_1(f, \zeta, w, s)|_{C_\zeta} = (1 - s)\Delta\big(\langle w, \zeta\rangle^{d_i}\big),$$

for all tuples (f, ζ, w, s). Moreover, since $(f, \zeta, w, s) \in \hat{\Lambda}_k$, then $\langle w, \zeta\rangle \neq 0$ and $s \neq 1$, hence $DF_1(f, \zeta, w, s)|_{C_\zeta}$ is onto.

Now we prove that $DF_2(f, \zeta, w, s)|_{\mathcal{V}_\zeta \times T_w \mathbb{P}(\mathbb{C}^{n+1})}$ is surjective onto the tangent space $T_{F_2(f,\zeta,w,s)}\overline{K}$, at every $(f, \zeta, w, s) \in \hat{\Lambda}_k$.

Note that the map $F_2(f, \zeta, \cdot, s) : \mathbb{C}^{n+1} - \{0\} \to \overline{K}$ is a section of the vector bundle $\overline{K}$. Therefore, from Lemma 1, it is enough to prove that $F_2(\cdot, \zeta, w, s) : \mathcal{H}_{(d)} \to \overline{K}_w$ is a surjective linear map.

Fix a tuple $(f, \zeta, w, s) \in \hat{\Lambda}_k$, for some $k = 0, \ldots, n - 1$. The unitary group $\mathcal{U}(n + 1)$ acts by isometries on $(\mathcal{H}_{(d)} - \Sigma) \times (\mathbb{C}^{n+1} - \{0\}) \times (\mathbb{C}^{n+1} - \{0\}) \times (0, 1)$ by $U \cdot (f, \zeta, w, s) = (f \circ U^{-1}, U(\zeta), U(w), s)$, and leave $\hat{\Lambda}_k$ invariant. Therefore we may assume that $w = e_0$. Write $f_i(z) = \sum_{\|\alpha\|=d_i} a_\alpha^{(i)} z^\alpha$ $(i = 1, \ldots, n)$. Then, the linear map $F_2(\cdot, \zeta, e_0, s) : \mathcal{H}_{(d)} \to \overline{K}_{e_0}$ is given by

$$F_2(f, \zeta, e_0, s) = \big(\big(\overline{\zeta_0} a^{(i)}_{(d_i-1,v_j)} - d_i a^{(i)}_{(d_i,0,\ldots,0)}\overline{\zeta_j}\big)\big)_{i,j=1,\ldots,n},$$

where v_j is the n-vector with the j-entry equal to 1 and the other entries equal to 0.

In particular, since $\zeta_0 \neq 0$, the restriction $F_2(\cdot, \zeta, e_0, s) : \mathcal{V}_\zeta \to \overline{K}_{e_0}$ is surjective, concluding the claim.

Then, since F is transversal to $\{0\} \times \overline{K_k}$, we conclude that $\hat{\Lambda}_k = F^{-1}(\{0\} \times \overline{K_k})$ is a submanifold of real codimension $2(n + (n - k)^2)$, for $k = 0, \ldots, n - 1$.

To end the proof, we note that $\hat{\Lambda}_k$ contains the fiber of the canonical projection $(\mathcal{H}_{(d)} - \Sigma) \times (\mathbb{C}^{n+1} - \{0\}) \times (\mathbb{C}^{n+1} - \{0\}) \times (0, 1) \to (\mathcal{H}_{(d)} - \Sigma) \times \mathbb{P}(\mathbb{C}^{n+1}) \times \mathbb{P}(\mathbb{C}^{n+1}) \times (0, 1)$. □

Let $\Pi : \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1}) \times \mathbb{P}(\mathbb{C}^{n+1}) \times (0, 1) \to \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1})$ be the canonical projection

$$\Pi(f, \zeta, w, s) = (f, \zeta).$$

Then, from Lemmas 2 and 3 the set of pairs $(f, \zeta) \in \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1})$ such that the homotopy is not defined for all $t \in [0, 1]$ is contained by the union

$$\Omega^{(0)} \cup \Pi(\Lambda) \cup \Sigma \times \mathbb{P}\big(\mathbb{C}^{n+1}\big) \subset \mathcal{H}_{(d)} \times \mathbb{P}\big(\mathbb{C}^{n+1}\big).$$

Remark We could conclude the proof by Fubini's Theorem. But we give a different argument. See the remark at the end.

Proof of Proposition 1 For $k = 0, \ldots, n-1$, let $\Omega_k^{(0)} \subset \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1})$ be the subset given in the proof of Lemma 2, and let $\hat{\pi}_1 : \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1}) \to \mathcal{H}_{(d)}$ be the projection in the first coordinate. From Sard's Lemma we see that almost every $f \in \mathcal{H}_{(d)}$ is a regular value of the restriction $\hat{\pi}_1|_{\Omega_k^{(0)}} : \Omega_k^{(0)} \to \mathcal{H}_{(d)}$, for each $k = 0, \ldots, n-1$. Therefore, from Lemma 2, we conclude that for almost every $f \in \mathcal{H}_{(d)}$ the subset

$$\hat{\pi}_1|_{\Omega_k^{(0)}}{}^{-1}(f) = \hat{\pi}_1^{-1}(f) \cap \Omega_k^{(0)} \subset \mathbb{P}\big(\mathbb{C}^{n+1}\big)$$

is an empty set or a smooth submanifold of complex dimension $n - (n-k)^2$, for $k = 0, \ldots, n-1$. Hence, for almost every $f \in \mathcal{H}_{(d)}$, the set of $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ such that Φ is not defined at $t = 0$ has measure zero.

Similar to the preceding argument, for each $k = 0, \ldots, n-1$, let $\Lambda_k \subset (\mathcal{H}_{(d)} - \Sigma) \times \mathbb{P}(\mathbb{C}^{n+1}) \times \mathbb{P}(\mathbb{C}^{n+1}) \times \mathbb{P}(\mathbb{C}^{n+1}) \times (0,1)$ be the set of tuples (f, ζ, w, s) such that Eqs. (9) and (10) hold, and let $\hat{\Pi}_f : \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1}) \times \mathbb{P}(\mathbb{C}^{n+1}) \times \mathbb{P}(\mathbb{C}^{n+1}) \times (0,1) \to \mathcal{H}_{(d)}$ be the projection in the first coordinate. Then by Sard's Lemma, almost every $f \in \mathcal{H}_{(d)}$ is a regular value of the restriction $\hat{\Pi}_f|_{\Lambda_k} : \Lambda_k \to \mathcal{H}_{(d)}$. Therefore, from Lemma 3, we conclude that for almost every $f \in \mathcal{H}_{(d)}$ the subset

$$\hat{\Pi}_f|_{\Lambda_k}^{-1}(f) = \hat{\Pi}_f^{-1}(f) \cap \Lambda_k \subset \mathbb{P}\big(\mathbb{C}^{n+1}\big) \times \mathbb{P}\big(\mathbb{C}^{n+1}\big) \times (0,1)$$

is an empty set or a smooth submanifold of real dimension $2n + 1 - 2(n-k)^2$, for $k = 0, \ldots, n-1$. Then, projecting in the ζ-space we see that for almost every $f \in \mathcal{H}_{(d)}$, the set of $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ such that Φ is not defined at $t \in (0,1)$ is a finite union of measure zero sets. Moreover, since $\Sigma \subset \mathcal{H}_{(d)}$ has measure zero, the proof of the first statement of the proposition follows.

The second statement of Proposition 1 follows directly from proofs of the claims of Lemma 2, and Lemma 3, and the subsequent analysis of dimensions. □

Remark The proof of Proposition 1 follows immediately from Fubini's Theorem. But we say more because this discussion may be useful for the discussion of the basins. This proposition proves that the boundary of the basins are contained in this stratified set, the structure of which should be persistent by the isotopy theorem (cf. [3]) on the connected components of the complement of the critical values of the projection. We do not know if there is more than one component.

3 Proof of Theorem 1

Let us first state the notation in the forthcoming computations. Most of the maps are defined between Hermitian spaces, however, they are real differentiable. Therefore, unless we mention the contrary, all derivatives are real derivatives. Moreover, if a map is defined on $\mathbb{P}(\mathbb{C}^{n+1})$ then is natural to restrict its derivative at ζ to the complex tangent space $T_\zeta\mathbb{P}(\mathbb{C}^{n+1})$. If $L : E \to F$ is a linear map between finite dimensional Hermitian vector spaces, then its determinant, $\det(L)$, is the determinant of the linear map $L : E \to \mathrm{Im}(L)$, computed with respect to the associated canonical real structures, namely, the real part of the Hermitian product of E and the real part of the inherited Hermitian product on $\mathrm{Im}(L) \subset F$. The adjoint operator $L^* : F \to E$ is also computed with respect to the associated canonical real structures.

In general, if E is a set, Id_E means the identity map defined on that set.

Since the set of triples $(f, \zeta, t) \in \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1}) \times [0, 1]$ such that $t = 0$ or $t = 1$ has measure zero, we may assume in the rest of this section that $t \in (0, 1)$.

Recall that $\Phi : \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1}) \times [0, 1] \to \mathcal{V}$ is the map given by

$$\Phi(f, \zeta, t) = (f_t, \zeta_t),$$

where

$$f_t = f - (1-t)\Delta\left(\frac{\langle \cdot, \zeta\rangle^{d_i}}{\langle \zeta, \zeta\rangle^{d_i}}\right) f(\zeta),$$

and ζ_t is the homotopy continuation of ζ along the path f_t.

For each $t \in (0, 1)$, let $\Phi_t : \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1}) \to \mathcal{V}$ be the restriction $\Phi_t(\cdot, \cdot) = \Phi(\cdot, \cdot, t)$.

Recall that for each non-degenerate root η of h, $B(h, \eta)$ is the non-empty open set of those $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ such that the zero ζ of $\Pi_\zeta(h)$ continues to η for the homotopy $h_t = (1-t)\Pi_\zeta(h) + th$.

Given $h \in \mathcal{H}_{(d)}$ and $t \in (0, 1)$, let $\hat{H}_t : \mathbb{P}(\mathbb{C}^{n+1}) \to \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1})$,

$$\hat{H}_t(\zeta) = \big(\hat{h}_t(\zeta), \zeta\big), \quad \text{and} \quad \hat{h}_t(\zeta) = h + \left(\frac{1-t}{t}\right)\Delta\left(\frac{\langle \cdot, \zeta\rangle^{d_i}}{\langle \zeta, \zeta\rangle^{d_i}}\right) h(\zeta), \tag{11}$$

for all $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$. (We have suppressed the h for ease of notation.)

Lemma 4 *Let $t \in (0, 1)$, and let $(h, \eta) \in \mathcal{V}$ be a regular value of Φ_t. Then, the fiber $\Phi_t(h, \eta)^{-1}$ is given by*

$$\Phi_t^{-1}(h, \eta) = \hat{H}_t\big(B(h, \eta)\big).$$

Proof For $0 < t < 1$, we have $(f, \zeta) \in \Phi_t^{-1}(h, \eta)$, provided that

(i) $h = f_t = tf + (1-t)\Pi_\zeta(f)$;
(ii) the homotopy continuation of ζ on the path $\{sh + (1-s)\Pi_\zeta(f)\}_{s \in [0,1]}$ is η.

Since $\Pi_\zeta(h) = \Pi_\zeta(f)$ we conclude that

$$f = \frac{1}{t}\big(h - (1-t)\Pi_\zeta(h)\big) = h + \left(\frac{1-t}{t}\right)\big(h - \Pi_\zeta(h)\big),$$

and $\zeta \in B(h, \eta)$. □

Proposition 3 *Let $(f, \zeta) \in \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1})$ such that Φ_t is defined and let $(h, \eta) = \Phi_t(f, \zeta)$. Then the normal Jacobian of Φ_t is given by*

$$NJ_{\Phi_t}(f, \zeta) = t^{2n} \frac{Jac_{\hat{H}_t}(\zeta)}{NJ_{\pi_1}(h, \eta)},$$

where $Jac_{\hat{H}_t}(\zeta) = |\det(D\hat{H}_t(\zeta))|$ is the Jacobian of the map $\hat{H}_t$ defined in (11).

The proof of this proposition is divided in several lemmas and is left to the end of this section.

Let us first recall *the co-area formula*. Let $\pi : X \to Y$ be a smooth surjective map between Riemannian manifolds X and Y. If almost every $y \in Y$ is a regular value of π, and $\varphi : X \to \mathbb{R}$ is integrable, then

$$\int_{x\in X} \varphi(x)\,\mathrm{d}X = \int_{y\in Y}\int_{x\in\pi^{-1}(y)} \frac{\varphi(x)}{NJ_\pi(x)}\,\mathrm{d}\pi^{-1}(y)\,\mathrm{d}Y.$$

In particular, the co-area formula for the projection $\pi_1 : \mathcal{V} \to \mathcal{H}_{(d)}$ and a function $\varphi : \mathcal{V} \to \mathbb{R}$ yields

$$\int_{(h,\eta)\in\mathcal{V}} \varphi(h,\eta)\,\mathrm{d}\mathcal{V} = \int_{h\in\mathcal{H}_{(d)}} \left(\sum_{\eta/\,h(\eta)=0} \frac{\varphi(h,\eta)}{NJ_{\pi_1}(h,\eta)}\right)\mathrm{d}h.$$

Proof of Theorem 1 Recall from Proposition 2 that (I) is defined by

$$\text{(I)} = \frac{CD^{3/2}}{(2\pi)^N \operatorname{vol}(\mathbb{P}(\mathbb{C}^{n+1}))} \int_{f\in\mathcal{H}_{(d)}}\int_{\zeta\in\mathbb{P}(\mathbb{C}^{n+1})}\int_{t\in[0,1]} \frac{\mu(f_t,\zeta_t)^2}{\|f_t\|^2}$$
$$\times \left\|\Pi_\zeta(f)\right\| \left\|\Delta\left(\|\zeta\|^{-d_i}\right) f(\zeta)\right\| e^{-\|f\|^2/2}\,\mathrm{d}t\,\mathrm{d}\zeta\,\mathrm{d}f.$$

Then, for $0 < t < 1$, by the co-area formula for the map $\Phi_t : \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1}) \to \mathcal{V}$, and Proposition 3 we obtain

$$\text{(I)} = \frac{CD^{3/2}}{(2\pi)^N \operatorname{vol}(\mathbb{P}(\mathbb{C}^{n+1}))} \int_0^1 t^{-2n} \int_{(h,\eta)\in\mathcal{V}} \frac{\mu(h,\eta)^2}{\|h\|^2} NJ_{\pi_1}(h,\eta)$$
$$\times \int_{(f,\zeta)\in\Phi_t^{-1}(h,\eta)} \frac{\|\Pi_\zeta(f)\|\|\Delta(\|\zeta\|^{-d_i}) f(\zeta)\|}{Jac_{\hat{H}_t}(\zeta)} e^{-\|f\|^2/2}\,\mathrm{d}\Phi_t^{-1}(h,\eta)\,\mathrm{d}\mathcal{V}\,\mathrm{d}t.$$

If $\Phi_t(f,\zeta) = (h,\eta)$, then $f(\zeta) = h(\zeta)/t$ and $\Pi_\zeta(f) = \Pi_\zeta(h)$. From Lemma 4 we find that, for all $t \in (0,1)$, $\hat{H}_t|_{B(h,\eta)} : B(h,\eta) \to \Phi_t^{-1}(h,\eta)$ given by $\zeta \mapsto (\hat{h}_t(\zeta),\zeta)$, is a parameterization of the fiber $\Phi_t^{-1}(h,\eta)$. Moreover, since $\zeta = \hat{H}_t^{-1}(f,\zeta)$ whenever $\hat{H}_t(\zeta) = (f,\zeta)$, applying the change of variable formula we conclude that

$$\text{(I)} = \frac{CD^{3/2}}{(2\pi)^N \operatorname{vol}(\mathbb{P}(\mathbb{C}^{n+1}))} \int_0^1 t^{-2n-1} \int_{(h,\eta)\in\mathcal{V}} \frac{\mu(h,\eta)^2}{\|h\|^2} NJ_{\pi_1}(h,\eta)$$
$$\times \int_{\zeta\in B(h,\eta)} \left\|\Pi_\zeta(h)\right\| \left\|\Delta\left(\|\zeta\|^{-d_i}\right) h(\zeta)\right\| e^{-\|\hat{h}_t(\zeta)\|^2/2}\,\mathrm{d}\zeta\,\mathrm{d}\mathcal{V}\,\mathrm{d}t. \qquad (12)$$

From the definition of $\hat{h}_t(\zeta)$ in (11) and the reproducing kernel property of the Weyl Hermitian product (5), we obtain

$$\|\hat{h}_t(\zeta)\|^2 = \|h\|^2 + 2\left(\frac{1-t}{t}\right)\mathrm{Re}\big\langle h, \Delta\big(\langle \zeta,\zeta\rangle^{-d_i}\langle\cdot,\zeta\rangle^{d_i}\big)h(\zeta)\big\rangle$$
$$+\left(\frac{1-t}{t}\right)^2\big\|\Delta\big(\langle\zeta,\zeta\rangle^{-d_i}\langle\cdot,\zeta\rangle^{d_i}\big)h(\zeta)\big\|^2,$$

and then

$$\|\hat{h}_t(\zeta)\|^2 = \|h\|^2 - \left(1-\frac{1}{t^2}\right)\big\|\Delta\big(\|\zeta\|^{-d_i}\big)h(\zeta)\big\|^2. \tag{13}$$

Therefore, from (12) and (13) we obtain

$$(\mathrm{I}) = \frac{CD^{3/2}}{(2\pi)^N}\int_{(h,\eta)\in\mathcal{V}}\frac{\mu(h,\eta)^2}{\|h\|^2}\Theta(h,\eta)NJ_{\pi_1}(h,\eta)e^{-\|h\|^2/2}\,\mathrm{d}\mathcal{V}, \tag{14}$$

where

$$\Theta(h,\eta) = \frac{1}{\mathrm{vol}(\mathbb{P}(\mathbb{C}^{n+1}))}\int_{\zeta\in B(h,\eta)}\big(\|h\|^2 - \big\|\Delta\big(\|\zeta\|^{-d_i}\big)h(\zeta)\big\|^2\big)^{1/2}$$
$$\times\big\|\Delta\big(\|\zeta\|^{-d_i}\big)h(\zeta)\big\|\mathcal{I}_n\big(\big\|\Delta\big(\|\zeta\|^{-d_i}\big)h(\zeta)\big\|^2/2\big)\,\mathrm{d}\zeta,$$

and $\mathcal{I}_n(\alpha) = \int_0^1 e^{(1-t^{-2})\alpha}t^{-2n-1}\,\mathrm{d}t$.

Now, the proof of Theorem 1 follows applying the co-area formula for the projection $\pi_1 : \mathcal{V}\to\mathcal{H}_{(d)}$. □

3.1 Proof of Proposition 3

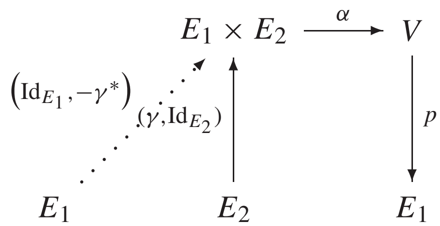
The map $\hat{h}_t : \mathbb{P}(\mathbb{C}^{n+1}) \to \mathcal{H}_{(d)}$ given in (11) is differentiable, and therefore $\hat{H}_t$ is also differentiable.

Lemma 5 *Let $(f,\zeta)\in\mathcal{H}_{(d)}\times\mathbb{P}(\mathbb{C}^{n+1})$ such that Φ_t is defined and let $(h,\eta) = \Phi_t(f,\zeta)$. Then*

$$NJ_{\Phi_t}(f,\zeta) = \frac{|\det[D(\pi_1\circ\Phi_t)(\hat{h}_t(\zeta),\zeta)\cdot(\mathrm{Id}_{\mathcal{H}_{(d)}}, -(D\hat{h}_t(\zeta)|_{\zeta^\perp})^*)]|}{|\det(\mathrm{Id}_{\zeta^\perp} + (D\hat{h}_t(\zeta)|_{\zeta^\perp})^* D\hat{h}_t(\zeta))|_{\zeta^\perp}|^{1/2}NJ_{\pi_1}(h,\eta)},$$

where $(\mathrm{Id}_{\mathcal{H}_{(d)}}, -(D\hat{h}_t(\zeta)|_{\zeta^\perp})^*) : \mathcal{H}_{(d)}\to\mathcal{H}_{(d)}\times T_\zeta\mathbb{P}(\mathbb{C}^{n+1})$ is the linear map $\dot{f}\mapsto(\dot{f}, -(D\hat{h}_t(\zeta)|_{\zeta^\perp})^*\dot{f})$.

Proof In general, let E_1, E_2, and V be finite dimensional vector spaces with inner product. Assume that $\dim(V) = \dim(E_1)$, and let $p : V\to E_1$ be an isomorphism. Let $\gamma : E_2\to E_1$ and $\alpha : E_1\times E_2\to V$ be linear operators. Consider the following diagram:



where $(\gamma, \mathrm{Id}_{E_2}) : E_2 \to E_1 \times E_2$.

Note that the image of the operator $(\mathrm{Id}_{E_1}, -\gamma^*) : E_1 \to E_1 \times E_2$ is the orthogonal complement of $(\gamma, \mathrm{Id})(E_2)$ in $E_1 \times E_2$; therefore we get

$$\begin{aligned}\left|\det(\alpha|_{((\gamma,\mathrm{Id}_{E_2})(E_2))^\perp})\right| &= \frac{|\det(p\cdot\alpha\cdot(\mathrm{Id}_{E_1}, -\gamma^*))|}{|\det(\mathrm{Id}_{E_1} + \gamma\gamma^*)|^{1/2}|\det(p)|}\\ &= \frac{|\det(p\cdot\alpha\cdot(\mathrm{Id}_{E_1}, -\gamma^*))|}{|\det(\mathrm{Id}_{E_2} + \gamma^*\gamma)|^{1/2}|\det(p)|},\end{aligned}$$

where the last equality follows by the Sylvester Theorem: if A and B are matrices of size $n \times m$ and $m \times n$, respectively, then

$$\det(\mathrm{Id}_m + BA) = \det(\mathrm{Id}_n + AB). \tag{15}$$

Now the proof follows taking $E_1 = \mathcal{H}_{(d)}$, $E_2 = T_\zeta\mathbb{P}(\mathbb{C}^{n+1})$, $V = T_{(h,\eta)}\mathcal{V}$, with the associated real inner products, $\gamma = D\hat{h}_t(\zeta)|_{\zeta^\perp}$, $\alpha = D\Phi_t(f,\zeta)$ and $p = D\pi_1(h,\eta)$. □

The derivative of $\hat{h}_t$ at $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ in the direction $\dot{\zeta} \in T_\zeta\mathbb{P}(\mathbb{C}^{n+1})$ is given by

$$D\hat{h}_t(\zeta)\dot{\zeta} = \left(\frac{1-t}{t}\right)\big(K_\zeta(\dot{\zeta}) + L_\zeta(\dot{\zeta})\big),$$

where $K_\zeta, L_\zeta : T_\zeta\mathbb{P}(\mathbb{C}^{n+1}) \to \mathcal{H}_{(d)}$ are given by

$$K_\zeta(\dot{\zeta}) = \Delta\left(\frac{\langle\cdot,\zeta\rangle^{d_i}}{\langle\zeta,\zeta\rangle^{d_i}}\right)Dh(\zeta)\dot{\zeta}; \tag{16}$$

$$L_\zeta(\dot{\zeta}) = \Delta\left(\frac{d_i\langle\cdot,\zeta\rangle^{d_i-1}\langle\cdot,\dot{\zeta}\rangle}{\langle\zeta,\zeta\rangle^{d_i}}\right)h(\zeta), \tag{17}$$

for all $\dot{\zeta} \in T_\zeta\mathbb{P}(\mathbb{C}^{n+1})$.

Lemma 6 *The adjoint operators $K_\zeta{}^*, L_\zeta{}^* : \mathcal{H}_{(d)} \to T_\zeta\mathbb{P}(\mathbb{C}^{n+1})$ are given by*

$$K_\zeta{}^*(\dot{f}) = \big(Dh(\zeta)|_{\zeta^\perp}\big)^*\Delta\big(\langle\zeta,\zeta\rangle^{-d_i+1}\big)\dot{f}(\zeta), \tag{18}$$

and

$$L_\zeta{}^*(\dot{f}) = \big(D\dot{f}(\zeta)|_{\zeta^\perp}\big)^*\Delta\big(\langle\zeta,\zeta\rangle^{-d_i+1}\big)h(\zeta), \tag{19}$$

for any $\dot{f} \in \mathcal{H}_{(d)}$.

Proof By the definition of adjoint, the definition of K_ζ and the reproducing kernel property of the Weyl Hermitian product (5), we get

$$\begin{aligned}\operatorname{Re}\big\langle {K_\zeta}^*(\dot f), \dot\zeta\big\rangle &= \|\zeta\|^2 \operatorname{Re}\big\langle \dot f, \Delta\big(\langle\zeta,\zeta\rangle^{-d_i}\langle\cdot,\zeta\rangle^{d_i}\big)Dh(\zeta)\dot\zeta\big\rangle \\ &= \operatorname{Re}\big\langle \dot f(\zeta), \Delta\big(\langle\zeta,\zeta\rangle^{-d_i+1}\big)Dh(\zeta)\dot\zeta\big\rangle \\ &= \operatorname{Re}\big\langle \big(Dh(\zeta)|_{\zeta^\perp}\big)^*\Delta\big(\langle\zeta,\zeta\rangle^{-d_i+1}\big)\dot f(\zeta), \dot\zeta\big\rangle.\end{aligned}$$

Moreover, differentiating equation (5) with respect to ζ, we obtain for ${L_\zeta}^*$

$$\begin{aligned}\operatorname{Re}\big\langle {L_\zeta}^*(\dot f), \dot\zeta\big\rangle &= \|\zeta\|^2 \operatorname{Re}\big\langle \dot f, \Delta\big(\langle\zeta,\zeta\rangle^{-d_i} d_i\langle\cdot,\zeta\rangle^{d_i-1}\langle\cdot,\dot\zeta\rangle\big)h(\zeta)\big\rangle \\ &= \operatorname{Re}\big\langle D\dot f(\zeta)\dot\zeta, \Delta\big(\langle\zeta,\zeta\rangle^{-d_i+1}\big)h(\zeta)\big\rangle \\ &= \operatorname{Re}\big\langle \big(D\dot f(\zeta)|_{\zeta^\perp}\big)^*\Delta\big(\langle\zeta,\zeta\rangle^{-d_i+1}\big)h(\zeta), \dot\zeta\big\rangle. \qquad \square\end{aligned}$$

Lemma 7 *One has*

$$\begin{aligned}&\big|\det\big(\mathrm{Id}_{\zeta^\perp} + \big(D\hat h_t(\zeta)|_{\zeta^\perp}\big)^* D\hat h_t(\zeta)|_{\zeta^\perp}\big)\big| \\ &\quad = \left(1 + \left(\frac{1-t}{t}\right)^2 \big\|\Delta\big(\sqrt{d_i}\|\zeta\|^{-d_i}\big)h(\zeta)\big\|^2\right)^{2n} \\ &\qquad \times \left|\det\left(\mathrm{Id}_{\zeta^\perp} + \frac{(\frac{1-t}{t})^2(Dh(\zeta)|_{\zeta^\perp})^*\Delta(\|\zeta\|^{-d_i+1})^2 Dh(\zeta)_{\zeta^\perp}}{1+(\frac{1-t}{t})^2\|\Delta(\sqrt{d_i}\|\zeta\|^{-d_i})h(\zeta)\|^2}\right)\right|.\end{aligned}$$

Proof By direct computation we get

$$\begin{gathered}{K_\zeta}^*K_\zeta = \big(Dh(\zeta)|_{\zeta^\perp}\big)^*\Delta\big(\langle\zeta,\zeta\rangle^{-d_i+1}\big)Dh(\zeta)|_{\zeta^\perp}; \\ {K_\zeta}^*L_\zeta = {L_\zeta}^*K_\zeta = 0.\end{gathered}$$

Note that, if $\dot f = L_\zeta(\dot\zeta)$ for some $\dot\zeta \in T_\zeta\mathbb{P}(\mathbb{C}^{n+1})$, then, for all $\theta \in \mathbb{C}^n$, we get

$$\big(D\dot f(\zeta)|_{\zeta^\perp}\big)^*\theta = \big(\operatorname{Re}\big\langle\theta, \Delta\big(d_i\|\zeta\|^{-2}\big)h(\zeta)\big\rangle\big)\dot\zeta.$$

Hence,

$${L_\zeta}^*L_\zeta = \big\|\Delta\big(\sqrt{d_i}\|\zeta\|^{-d_i}\big)h(\zeta)\big\|^2\,\mathrm{Id}_{\zeta^\perp}.$$

Therefore we get

$$\begin{aligned}\big(D\hat h_t(\zeta)|_{\zeta^\perp}\big)^* D\hat h_t(\zeta)\big|_{\zeta^\perp} &= \left(\frac{1-t}{t}\right)^2\big({K_\zeta}^*K_\zeta + {L_\zeta}^*L_\zeta\big) \\ &= \left(\frac{1-t}{t}\right)^2\big(\big(Dh(\zeta)|_{\zeta^\perp}\big)^*\Delta\big(\|\zeta\|^{-2d_i+2}\big)Dh(\zeta)|_{\zeta^\perp} \\ &\qquad + \big\|\Delta\big(\sqrt{d_i}\|\zeta\|^{-d_i}\big)h(\zeta)\big\|^2\,\mathrm{Id}_{\zeta^\perp}\big).\end{aligned}$$

The proof follows. $\square$

Lemma 8 *One has*

$$\begin{aligned}&\big|\det\big[D(\pi_1\circ\Phi_t)\big(\hat{h}_t(\zeta),\zeta\big)\cdot\big(\mathrm{Id}_{\mathcal{H}_{(d)}},-\big(D\hat{h}_t(\zeta)|_{\zeta^\perp}\big)^*\big)\big]\big|\\&\quad=\big|\det\big(\mathrm{Id}_{\zeta^\perp}+\big(D\hat{h}_t(\zeta)|_{\zeta^\perp}\big)^* D\hat{h}_t(\zeta)|_{\zeta^\perp}\big)\big|t^{2n}.\end{aligned}$$

Proof First we find an expression for the term inside the determinant. For short, let

$$\psi=D(\pi_1\circ\Phi_t)\big(\hat{h}_t(\zeta),\zeta\big)\cdot\big(\mathrm{Id}_{\mathcal{H}_{(d)}},-\big(D\hat{h}_t(\zeta)|_{\zeta^\perp}\big)^*\big).$$

One gets

$$\left[\frac{\partial}{\partial f}(\pi_1\circ\Phi_t)(f,\zeta)\right](\dot{f})=\dot{f}-(1-t)\Delta\left(\frac{\langle\cdot,\zeta\rangle^{d_i}}{\langle\zeta,\zeta\rangle^{d_i}}\right)\dot{f}(\zeta), \tag{20}$$

and

$$\begin{aligned}&\left[\frac{\partial}{\partial \zeta}(\pi_1\circ\Phi_t)(f,\zeta)\right](\dot{\zeta})\\&\quad=-(1-t)\left[\Delta\left(\frac{\langle\cdot,\zeta\rangle^{d_i}}{\langle\zeta,\zeta\rangle^{d_i}}\right)Df(\zeta)\dot{\zeta}+\Delta\left(\frac{d_i\langle\cdot,\zeta\rangle^{d_i-1}\langle\cdot,\dot{\zeta}\rangle}{\langle\zeta,\zeta\rangle^{d_i}}\right)f(\zeta)\right].\end{aligned} \tag{21}$$

Since $\hat{h}_t(\zeta)(\zeta)=h(\zeta)/t$, and $D[\hat{h}_t(\zeta)](\zeta)|_{\zeta^\perp}=Dh(\zeta)|_{\zeta^\perp}$, from (20) and (21) we get

$$\begin{aligned}\psi(\dot{f})&=\dot{f}-(1-t)\Delta\left(\frac{\langle\cdot,\zeta\rangle^{d_i}}{\langle\zeta,\zeta\rangle^{d_i}}\right)\dot{f}(\zeta)\\&\quad+(1-t)\left[\Delta\left(\frac{\langle\cdot,\zeta\rangle^{d_i}}{\langle\zeta,\zeta\rangle^{d_i}}\right)Dh(\zeta)|_{\zeta^\perp}\big(D\hat{h}_t(\zeta)|_{\zeta^\perp}\big)^*\dot{f}\right.\\&\quad\left.+\Delta\left(\frac{d_i\langle\cdot,\zeta\rangle^{d_i-1}\langle\cdot,(D\hat{h}_t(\zeta)|_{\zeta^\perp})^*\dot{f}\rangle}{\langle\zeta,\zeta\rangle^{d_i}}\right)\frac{h(\zeta)}{t}\right],\end{aligned}$$

for all $\dot{f}\in\mathcal{H}_{(d)}$. That is, with the notation K_ζ and L_ζ given in (16) and (17), we get

$$\begin{aligned}\psi(\dot{f})&=\dot{f}-(1-t)\left[\Delta\left(\frac{\langle\cdot,\zeta\rangle^{d_i}}{\langle\zeta,\zeta\rangle^{d_i}}\right)\dot{f}(\zeta)-\left(\frac{1-t}{t}\right)K_\zeta\big({K_\zeta}^*+{L_\zeta}^*\big)\dot{f}\right]\\&\quad+\left(\frac{1-t}{t}\right)^2L_\zeta\big({K_\zeta}^*+{L_\zeta}^*\big)\dot{f}\end{aligned} \tag{22}$$

for all $\dot{f}\in\mathcal{H}_{(d)}$.

Note that $\psi=\mathrm{Id}_{\mathcal{H}_{(d)}}-\mathcal{L}$, for a certain operator $\mathcal{L}$. Therefore $\det(\psi)=\det((\mathrm{Id}_{\mathcal{H}_{(d)}}-\mathcal{L})|_{\mathrm{Im}\mathcal{L}})$, where last determinant must be understood as the determinant of the linear operator $(\mathrm{Id}_{\mathcal{H}_{(d)}}-\mathcal{L})|_{\mathrm{Im}\,\mathcal{L}}:\mathrm{Im}\,\mathcal{L}\to\mathrm{Im}\,\mathcal{L}$.

The image of $\mathcal{L}$ is decomposed into two orthogonal subspaces, namely:

$$C_\zeta := \left\{ \Delta\left(\frac{\langle \cdot, \zeta\rangle^{d_i}}{\langle \zeta, \zeta\rangle^{d_i}}\right) a : a = (a_1, \ldots, a_n)^T \in \mathbb{C}^n \right\};$$

$$R_\zeta := \left\{ L_\zeta(w) : w \in T_\zeta \mathbb{P}\left(\mathbb{C}^{n+1}\right)\right\}.$$

Note that $\operatorname{Im} K_\zeta = C_\zeta \subset \ker {L_\zeta}^*$ and $\operatorname{Im} L_\zeta = R_\zeta \subset \ker {K_\zeta}^*$.

Consider the linear map

$$\tau : \mathbb{C}^n \to C_\zeta, \qquad \tau(b) = \Delta\left(\frac{\langle \cdot, \zeta\rangle^{d_i}}{\langle \zeta, \zeta\rangle^{d_i}}\right) \Delta\left(\|\zeta\|^{d_i}\right) b, \quad b \in \mathbb{C}^n.$$

Note that $\tau^{-1}(\Delta(\frac{\langle \cdot,\zeta\rangle^{d_i}}{\langle \zeta,\zeta\rangle^{d_i}})a) = \Delta(\|\zeta\|^{-d_i})a$. Since

$$\left\| \Delta\left(\frac{\langle \cdot, \zeta\rangle^{d_i}}{\langle \zeta, \zeta\rangle^{d_i}}\right) a \right\| = \left\| \Delta\left(\|\zeta\|^{-d_i}\right) a \right\|,$$

we conclude that τ is a linear isometry between $\mathbb{C}^n$ and C_ζ.

Let

$$\eta : T_\zeta \mathbb{P}\left(\mathbb{C}^{n+1}\right) \to R_\zeta, \qquad \eta(\cdot) = \frac{\|\zeta\|}{\|\Delta(\sqrt{d_i}\|\zeta\|^{-d_i})h(\zeta)\|} L_\zeta(\cdot).$$

Since

$$\left\| L_\zeta(w) \right\| = \left\| \Delta\left(\sqrt{d_i}\|\zeta\|^{-d_i}\right) h(\zeta) \right\| \frac{\|w\|}{\|\zeta\|},$$

for all $w \in T_\zeta \mathbb{P}(\mathbb{C}^{n+1})$, we find that η is a linear isometry between $T_\zeta \mathbb{P}(\mathbb{C}^{n+1})$ and R_ζ.

Let $\Pi_{C_\zeta} \psi$ and $\Pi_{R_\zeta} \psi$ be the orthogonal projections on C_ζ and R_ζ, respectively.

Then $|\det(\psi)|$ is equal to the absolute value of the determinant of

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

where $A = \tau^{-1} \circ \Pi_{C_\zeta} \psi|_{C_\zeta} \circ \tau$, $B = \tau^{-1} \circ \Pi_{C_\zeta} \psi|_{R_\zeta} \circ \eta$, $C = \eta^{-1} \circ \Pi_{R_\zeta} \psi|_{C_\zeta} \circ \tau$ and $D = \eta^{-1} \circ \Pi_{R_\zeta} \psi|_{R_\zeta} \circ \eta$.

Straightforward computation shows that

$$A = t \,\mathrm{Id}_{\mathbb{C}^n} + \frac{(1-t)^2}{t} \Delta\left(\|\zeta\|^{-d_i+1}\right) Dh(\zeta)|_{\zeta^\perp} \left(Dh(\zeta)|_{\zeta^\perp}\right)^* \Delta\left(\|\zeta\|^{-d_i+1}\right);$$

$$B = \frac{(1-t)^2}{t} \left\| \Delta\left(\sqrt{d_i}\|\zeta\|^{-d_i}\right) h(\zeta) \right\| \Delta\left(\|\zeta\|^{-d_i+1}\right) Dh(\zeta)|_{\zeta^\perp};$$

$$C = \left(\frac{1-t}{t}\right)^2 \left\| \Delta\left(\sqrt{d_i}\|\zeta\|^{-d_i}\right) h(\zeta) \right\| \left(Dh(\zeta)|_{\zeta^\perp}\right)^* \Delta\left(\|\zeta\|^{-d_i+1}\right);$$

$$D = \left(1 + \left(\frac{1-t}{t}\right)^2 \left\|\Delta\left(\sqrt{d_i}\|\zeta\|^{-d_i}\right)h(\zeta)\right\|^2\right) \mathrm{Id}_{\zeta^\perp}.$$

Since D is invertible, we may write

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A - BD^{-1}C & B \\ 0 & D \end{pmatrix} \cdot \begin{pmatrix} I & 0 \\ D^{-1}C & I \end{pmatrix},$$

hence $\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det D \det(A - BD^{-1}C)$.

Thus,

$$\begin{aligned} &\left|\det(\psi)\right| \\ &= t^{2n}\left(1 + \left(\frac{1-t}{t}\right)^2 \left\|\Delta\left(\sqrt{d_i}\|\zeta\|^{-d_i}\right)h(\zeta)\right\|^2\right)^{2n} \\ &\quad \times \left|\det\left(\mathrm{Id}_{\mathbb{C}^n} + \frac{(\frac{1-t}{t})^2 \Delta(\|\zeta\|^{-d_i+1}) Dh(\zeta)|_{\zeta^\perp} (Dh(\zeta)|_{\zeta^\perp})^* \Delta(\|\zeta\|^{-d_i+1})}{1 + (\frac{1-t}{t})^2 \|\Delta(\sqrt{d_i}\|\zeta\|^{-d_i})h(\zeta)\|^2}\right)\right|^2. \end{aligned}$$

Observe that

$$\begin{aligned} &\left(Dh(\zeta)|_{\zeta^\perp}\right)^* \Delta\left(\|\zeta\|^{-d_i+1}\right)^2 Dh(\zeta)|_{\zeta^\perp} \\ &\quad = \left(\Delta\left(\|\zeta\|^{-d_i+1}\right) Dh(\zeta)|_{\zeta^\perp}\right)^* \left(\Delta\left(\|\zeta\|^{-d_i+1}\right) Dh(\zeta)|_{\zeta^\perp}\right). \end{aligned}$$

Then, the proof follows from Lemma 7 and the Sylvester Theorem (15). □

Proof of Proposition 3 The Jacobian of $\hat{H}_t : \mathbb{P}(\mathbb{C}^{n+1}) \to \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1})$ at ζ is given by

$$\left|\det\left(\mathrm{Id}_{\zeta^\perp} + \left(D\hat{h}_t(\zeta)|_{\zeta^\perp}\right)^* D\hat{h}_t(\zeta)|_{\zeta^\perp}\right)\right|^{1/2}.$$

Then, the proof follows from Lemmas 5 and 8. □

4 Proof of Theorem 2 and Theorem 3

4.1 Proof of Theorem 2

Recall that

$$\theta_h(\zeta) = \left\|\Pi_\zeta(h)\right\| \left\|h - \Pi_\zeta(h)\right\| \mathcal{I}_n\left(\left\|h - \Pi_\zeta(h)\right\|^2/2\right),$$

where $\mathcal{I}_n(\alpha) = \int_0^1 e^{(1-t^{-2})\alpha} t^{-2n-1}\, \mathrm{d}t$.

In order to prove Theorem 2 we need some extra notation. Denote by S^{2n+1} the unit sphere in $\mathbb{C}^{n+1} = \mathbb{R}^{2(n+1)}$. For all $k \in \mathbb{N}$, let $\sigma_k : \mathbb{R}^k \to \mathbb{R}$ be the norm function $\sigma_k(x) = \|x\|$. Given a measurable function $\varphi : \mathbb{R}^k \to \mathbb{R}$, let us denote by $\mathbb{E}_{\mathbb{R}^k}(\varphi)$ the expected value with respect to the standard Gaussian measure on $\mathbb{R}^k$, i.e.,

$$\mathbb{E}_{\mathbb{R}^k}(\varphi) = \frac{1}{(2\pi)^{k/2}} \int_{\mathbb{R}^k} \varphi(x) e^{-\|x\|^2/2}\, \mathrm{d}x.$$

Lemma 9 *Let W be a complex vector space of dimension m with an Hermitian product $\langle\cdot,\cdot\rangle_W$. Denote $\|w\| = \sqrt{\langle w, w\rangle_W}$. Then, for every integrable function $\psi : \mathbb{R} \to \mathbb{R}$ one has*

$$\frac{1}{(2\pi)^m}\int_{w\in W} \psi\big(\|w\|\big)e^{-\|w\|^2/2}\,\mathrm{d}w = \mathbb{E}_{\mathbb{R}^{2m}}(\psi\circ\sigma_{2m}).$$

Proof If $w_1,\ldots,w_m$ is an orthonormal basis of $(W,\langle\cdot,\cdot\rangle_W)$, then, $w_1,\ldots,w_m$, $iw_1,\ldots,iw_m$ is an orthonormal basis of the associated $2m$-dimensional real vector space, namely $W_{\mathbb{R}}$, with the real inner product $\mathrm{Re}\langle\cdot,\cdot\rangle_W$. Define the linear map $A : W_{\mathbb{R}} \to \mathbb{R}^{2m}$, by $A(w_k) = e_k$, and $A(iw_k) = e_{m+k}$, where e_k is the kth element of the standard orthonormal basis of $\mathbb{R}^{2m}$. A straightforward computation shows that A is a (real) linear isometry. Then the lemma follows by the change of variable formula. □

Lemma 10 *Let $p \geq 1$. Then*

$$\mathbb{E}_{\mathcal{H}_{(d)}}\big(\|\theta_h\|^p_{L^p}\big) = \mathbb{E}_{\mathbb{R}^{2(N-n)}}\big(\sigma^p_{2(N-n)}\big)\mathbb{E}_{\mathbb{R}^{2n}}\big(\sigma^p_{2n}\mathcal{I}_n{}^p\big(\sigma^2_{2n}/2\big)\big).$$

Proof Given the canonical projection $S^{2n+1} \to \mathbb{P}(\mathbb{C}^{n+1})$, by the co-area formula we get

$$\|\theta_h\|^p_{L^p} = \frac{1}{2\pi}\frac{1}{\mathrm{vol}(\mathbb{P}(\mathbb{C}^{n+1}))}\int_{w\in S^{2n+1}}\theta_h(w)^p\,\mathrm{d}w. \tag{23}$$

Recall that $\mathcal{V}_w = \{f \in \mathcal{H}_{(d)} : f(w) = 0\}$ and C_w is the Hermitian complement of $\mathcal{V}_w$, and that $\Pi_w : \mathcal{H}_{(d)} \to \mathcal{V}_w$ and $\Pi|_{C_W}$ are the orthogonal projections onto $\mathcal{V}_w$ and C_w, respectively. Then we may write $\mathcal{H}_{(d)} = \mathcal{V}_w \oplus C_w$. Denote by $\psi(\alpha) = \alpha\mathcal{I}_n(\alpha^2/2)$. Then $\theta_h(w) = \|\Pi_w(h)\|\psi(\|\Pi|_{C_w}(h)\|)$. Since $\|h\|^2 = \|\Pi_w(h)\|^2 + \|\Pi_{C_w}(h)\|^2$, by Fubini's Theorem we get

$$\int_{h\in\mathcal{H}_{(d)}}\theta_h(w)^p\frac{e^{-\|h\|^2/2}}{(2\pi)^N}\,\mathrm{d}h = \int_{f\in\mathcal{V}_w}\|f\|^p\frac{e^{-\|f\|^2/2}}{(2\pi)^{N-n}}\,\mathrm{d}f\int_{g\in C_w}\psi\big(\|g\|\big)^p\frac{e^{-\|g\|^2/2}}{(2\pi)^n}\,\mathrm{d}g.$$

Since $\mathcal{V}_w$ and C_w are complex vector spaces with an Hermitian product of dimensions $N-n$ and n, respectively, the proof follows interchanging in (23) the integral sign with the sign of expectation, Lemma 9 and the fact that $\mathrm{vol}(S^{2n+1}) = 2\pi\,\mathrm{vol}(\mathbb{P}(\mathbb{C}^{n+1}))$. □

Lemma 11 *Let $p \geq 1$. With the above definitions, one has*

(i) $\mathbb{E}_{\mathbb{R}^{2(N-n)}}(\sigma^p_{2(N-n)}) = 2^{p/2}\frac{\Gamma(N-n+p/2)}{\Gamma(N-n)}$.

(ii) $\mathbb{E}_{\mathbb{R}^{2n}}(\sigma^p_{2n}\mathcal{I}_n{}^p(\sigma^2_{2n}/2)) \leq \frac{2^{p/2}}{p}\frac{\Gamma(n+p/2)}{\Gamma(n)}$, *where the equality holds for $p = 1$.*

Proof (i): By definition, and integrating by polar coordinates, we have

$$\begin{aligned}\mathbb{E}_{\mathbb{R}^{2(N-n)}}\big(\sigma^p_{2(N-n)}\big) &= \frac{1}{(2\pi)^{N-n}}\int_{\mathbb{R}^{2(N-n)}}\|x\|^p e^{-\|x\|^2/2}\,\mathrm{d}x\\ &= \frac{\mathrm{vol}(S^{2(N-n)-1})}{(2\pi)^{N-n}}\int_0^{+\infty}\rho^{2(N-n)-1}\rho^p e^{-\rho^2/2}\,\mathrm{d}\rho.\end{aligned}$$

Then performing the change of variable $u = \rho^2/2$ we obtain

$$\begin{aligned}\mathbb{E}_{\mathbb{R}^{2(N-n)}}(\sigma_{2(N-n)}) &= \frac{\mathrm{vol}(S^{2(N-n)-1})2^{N-n+p/2-1}}{(2\pi)^{N-n}} \int_0^{+\infty} u^{N-n+p/2-1} e^{-u}\,\mathrm{d}u \\ &= \frac{\mathrm{vol}(S^{2(N-n)-1})2^{N-n+p/2-1}}{(2\pi)^{N-n}} \Gamma(N-n+p/2).\end{aligned}$$

The assertion follows from the fact that $\mathrm{vol}(S^k) = 2\frac{\pi^{(k+1)/2}}{\Gamma((k+1)/2)}$.

(ii) By definition of $\mathcal{I}_n$, we have

$$\mathbb{E}_{\mathbb{R}^{2n}}\left(\sigma_{2n}^p {\mathcal{I}_n}^p\left(\sigma_{2n}^2/2\right)\right) = \int_0^1 t^{-2n-1} \mathbb{E}_{\mathbb{R}^{2n}}\left(\sigma_{2n}^p e^{p(1-\frac{1}{t^2})\sigma_{2n}^2/2}\right) \mathrm{d}t. \tag{24}$$

Then, for fixed $t \in (0, 1)$,

$$\begin{aligned}\mathbb{E}_{\mathbb{R}^{2n}}\left(\sigma_{2n}^p e^{p(1-\frac{1}{t^2})\sigma_{2n}^2/2}\right) &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^{2n}} \|x\|^p e^{p(1-\frac{1}{t^2})\|x\|^2/2} e^{-\|x\|^2/2}\,\mathrm{d}x \\ &= \frac{\mathrm{vol}(S^{2n-1})}{(2\pi)^n} \int_0^{+\infty} \rho^{2n-1} \rho^p e^{-(\frac{p}{t^2}-(p-1))\rho^2/2}\,\mathrm{d}\rho.\end{aligned}$$

Now, performing the change of variable $u = (\frac{p}{t^2} - (p-1))\rho^2/2$, we get

$$\begin{aligned}\mathbb{E}_{\mathbb{R}^{2n}}\left(\sigma_{2n}^p e^{p(1-\frac{1}{t^2})\sigma_{2n}^2/2}\right) &= \frac{\mathrm{vol}(S^{2n-1})2^{n+p/2-1}}{(2\pi)^n(\frac{p}{t^2}-(p-1))^{n+p/2}} \Gamma(n+p/2) \\ &= \frac{t^{2n+p}}{(p-t^2(p-1))^{n+p/2}} 2^{p/2} \frac{\Gamma(n+p/2)}{\Gamma(n)}.\end{aligned}$$

Since $t \in (0, 1)$, and $p \geq 1$, we have the bound $p - t^2(p-1) \geq 1$, and hence

$$\mathbb{E}_{\mathbb{R}^{2n}}\left(\sigma_{2n}^p e^{p(1-\frac{1}{t^2})\sigma_{2n}^2/2}\right) \leq t^{2n+p} 2^{p/2} \frac{\Gamma(n+p/2)}{\Gamma(n)}$$

(where the equality holds for $p = 1$). Then the proof follows from (24). □

Proof of Theorem 2 The proof follows from Lemmas 10 and 11. □

4.2 Proof of Theorem 3

Proposition 4 *Let $p, q \geq 1$ such that $1/p + 1/q = 1$. If $1 < q < 2$, then we have*

$$\begin{aligned}(\mathrm{I}) \leq C D^{3/2} \mathfrak{D} &\left[\frac{2^p}{p} \frac{\Gamma(N-n+p/2)}{\Gamma(N-n)} \frac{\Gamma(n+p/2)}{\Gamma(n)}\right]^{1/p} \\ &\times \left[\frac{\Gamma(N+1)}{\Gamma(N+1-q)} \frac{\Gamma(n^2+n-q)}{\Gamma(n^2+n)} \frac{2^{2q+2}}{4-2q} n^{3q}\right]^{1/q}.\end{aligned}$$

Proof Consider on $\mathcal{V}$ the following density measure:

$$\mathrm{d}\rho_{\mathcal{V}} = (2\pi)^{-N}\mathcal{D}^{-1}N J_{\pi_1}(h,\eta)e^{-\|h\|^2/2}\,\mathrm{d}\mathcal{V}.$$

By the co-area formula we see that this measure is a probability measure on $\mathcal{V}$.

From (14), the definition of $\hat{\Theta}$ in (7), and Fubini's Theorem, we obtain

$$\begin{aligned}
(\mathrm{I}) &= C D^{3/2}\mathcal{D}\int_{\mathcal{V}} \frac{\mu(h,\eta)^2}{\|h\|^2}\Theta(h,\eta)\,\mathrm{d}\rho_{\mathcal{V}} \\
&\le C D^{3/2}\mathcal{D}\int_{\mathcal{V}} \frac{\mu(h,\eta)^2}{\|h\|^2}\hat{\Theta}(h)\,\mathrm{d}\rho_{\mathcal{V}} \\
&= \frac{C D^{3/2}\mathcal{D}}{\mathrm{vol}(\mathbb{P}(\mathbb{C}^{n+1}))}\int_{\mathbb{P}(\mathbb{C}^{n+1})}\left(\int_{\mathcal{V}} \frac{\mu(h,\eta)^2}{\|h\|^2}\theta_h(\zeta)\,\mathrm{d}\rho_{\mathcal{V}}\right)\mathrm{d}\zeta.
\end{aligned}$$

The function $\theta_h(\zeta)$, as a function defined on $\mathcal{V}$, is constant and equal to $\|\Pi_\zeta(h)\|\|h - \Pi_\zeta(h)\|\mathcal{I}_n(\|h-\Pi_\zeta(h)\|^2/2)$ on the fiber of the projection $\pi_1 : \mathcal{V} \to \mathcal{H}_{(d)}$.

For $p, q > 0$ such that $1/p + 1/q = 1$, Hölder inequality on $(\mathcal{V}, \mathrm{d}\rho_{\mathcal{V}})$ yields

$$(\mathrm{I}) \le \frac{C D^{3/2}\mathcal{D}}{\mathrm{vol}(\mathbb{P}(\mathbb{C}^{n+1}))}\int_{\mathbb{P}(\mathbb{C}^{n+1})}\left(\int_{\mathcal{V}} \frac{\mu(h,\eta)^{2q}}{\|h\|^{2q}}\,\mathrm{d}\rho_{\mathcal{V}}\right)^{1/q}\left(\int_{\mathcal{V}}\theta_h(\zeta)^p\,\mathrm{d}\rho_{\mathcal{V}}\right)^{1/p}\mathrm{d}\zeta.$$

Then, applying Fubini's Theorem and taking the canonical projection $\pi_1 : \mathcal{V} \to \mathcal{H}_{(d)}$ we get by the co-area formula

$$\begin{aligned}
(\mathrm{I}) &\le \frac{C D^{3/2}\mathcal{D}}{\mathcal{D}^{1/q+1/p}}\left(\mathbb{E}_{\mathcal{H}_{(d)}}\left(\sum_{\eta/h(\eta)=0}\frac{\mu^{2q}(h,\eta)}{\|h\|^{2q}}\right)\right)^{1/q} \\
&\quad\times \frac{1}{\mathrm{vol}(\mathbb{P}(\mathbb{C}^{n+1}))}\int_{\mathbb{P}(\mathbb{C}^{n+1})}\left(\mathcal{D}\mathbb{E}_{\mathcal{H}_{(d)}}\left(\theta_h(\zeta)^p\right)\right)^{1/p}\mathrm{d}\zeta,
\end{aligned}$$

where, abusing notation, we denote

$$\mathbb{E}_{\mathcal{H}_{(d)}}\left(\sum_{\eta/h(\eta)=0}\frac{\mu^{2q}(h,\eta)}{\|h\|^{2q}}\right) = \frac{1}{(2\pi)^N}\int_{h\in\mathcal{H}_{(d)}}\left[\sum_{\eta/h(\eta)=0}\frac{\mu^{2q}(h,\eta)}{\|h\|^{2q}}\right]e^{-\|h\|^2/2}\,\mathrm{d}h.$$

From the proof of Lemma 10 we see that $\mathbb{E}_{\mathcal{H}_{(d)}}(\theta_h(\zeta)^p)$ is independent of $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$. Then

$$\mathbb{E}_{\mathcal{H}_{(d)}}\left(\theta_h(\zeta)^p\right) = \mathbb{E}_{\mathcal{H}_{(d)}}\left(\|\theta_h\|_{L^p}^p\right),$$

and hence

$$(\mathrm{I}) \le C D^{3/2}\mathcal{D}^{1/p}\left(\mathbb{E}_{\mathcal{H}_{(d)}}\left(\sum_{\eta/h(\eta)=0}\frac{\mu^{2q}(h,\eta)}{\|h\|^{2q}}\right)\right)^{1/q}\mathbb{E}_{\mathcal{H}_{(d)}}\left(\|\theta_h\|_{L^p}^p\right)^{1/p}. \tag{25}$$

In Beltrán–Shub [7] (see also Beltrán–Pardo [6]) it is proved that, for $0<\alpha<4$,

$$\mathbb{E}_{\mathcal{H}_{(d)}}\left(\sum_{\eta/h(\eta)=0}\frac{\mu^{\alpha}(h,\eta)}{\|h\|^{\alpha}}\right) \le \mathcal{D}\frac{\Gamma(N+1)}{\Gamma(N+1-\alpha/2)}\frac{\Gamma(n^2+n-\alpha/2)}{\Gamma(n^2+n)}\frac{2^{\alpha+2}}{4-\alpha}n^{3\alpha/2}. \tag{26}$$

Now the proof follows from (25), (26) and Theorem 2. □

Proof of Theorem 3 The Gamma function $\Gamma(x)$, for $x>0$, is logarithmic convex (see Artin [1]). Then it is an easy exercise to check that $\Gamma(x+1/2)\le\sqrt{x}\Gamma(x)$, for all $x>0$.

Let $p=3$ and $q=3/2$. Then applying this inequality to Proposition 4 yields

$$(\mathrm{I})\le C'D^{3/2}\mathcal{D}Nn\sqrt{N-n+1/2}\sqrt{n+1/2},$$

where $C'=C2^{13/3}3^{-1/3}$. Then, the result of the theorem follows from the trivial bound $N-n+1/2\le N$ and $n+1/2\le(3/2)n$. □

5 Numerical Experiments

In this section we present some numerical experiments for $n=1$ and $d=7$ that were performed by Carlos Beltrán on the Altamira supercomputer at the Universidad de Cantabria.

Recall from Theorem 1 that

$$\Theta(h,\eta)=\frac{1}{\mathrm{vol}(\mathbb{P}(\mathbb{C}^{n+1}))}\int_{\zeta\in B(h,\eta)}\left(\|h\|^2-\left\|\Delta\left(\|\zeta\|^{-d_i}\right)h(\zeta)\right\|^2\right)^{1/2} \times\left\|\Delta\left(\|\zeta\|^{-d_i}\right)h(\zeta)\right\|\mathcal{I}_n\left(\left\|\Delta\left(\|\zeta\|^{-d_i}\right)h(\zeta)\right\|^2/2\right)\mathrm{d}\zeta,$$

where $\mathcal{I}_n(\alpha)=\int_0^1 e^{(1-t^{-2})\alpha}t^{-2n-1}\,\mathrm{d}t$.

Table 1 concerns a degree 7 polynomial h, chosen at random with the Bombieri–Weyl distribution. The condition numbers $\mu(h,\eta)$, $\Theta(h,\eta)$ and $\mathrm{vol}(B(h,\eta))$, at each root η of h are computed.

Table 1 Degree 7 random polynomial

Roots in $\mathbb{C}$	$\mu(h,\cdot)$	$\Theta(h,\cdot)$	$\mathrm{vol}(B(h,\cdot))$
$3.260883-i1.658800$	1.712852	0.4733570	0.140509π
$-2.357860-i1.329208$	1.738380	0.5502839	0.138576π
$-0.210068+i1.868947$	1.608231	0.5049662	0.144054π
$0.227994-i0.782004$	1.909433	0.4914771	0.125685π
$-0.044701+i0.384342$	3.231554	1.003594	0.147277π
$-0.308283+i0.049618$	3.183603	0.8892611	0.152433π
$0.213950-i0.068700$	2.948318	0.8426484	0.151466π

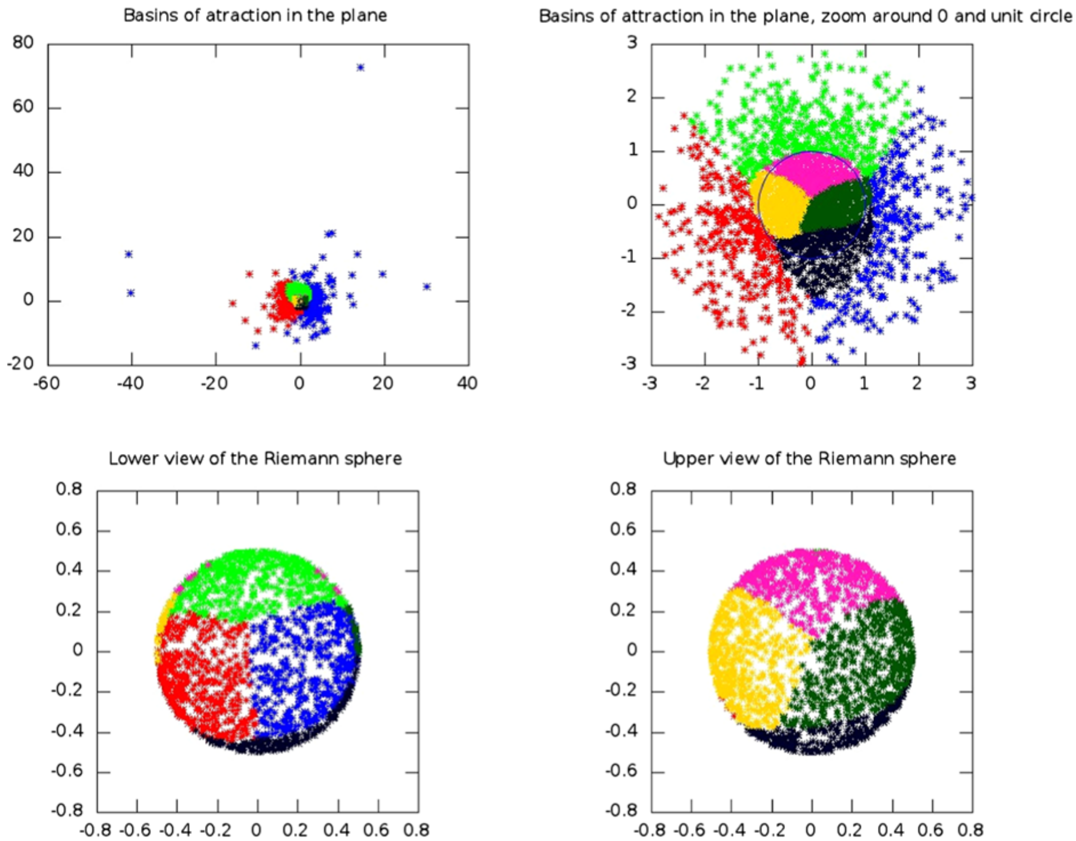


Fig. 2 The basins $B(h,\eta)$ in $\mathbb{C}$ and in the Riemann sphere of the degree 7 random polynomial (GNU Octave) (Color figure online)

The data of the chosen random polynomial is given by

$$a_7 = -0.152840 - i0.757630,$$

$$a_6 = 1.283080 + i0.357670,$$

$$a_5 = 2.000560 + i3.302700,$$

$$a_4 = 13.004500 + i0.203300,$$

$$a_3 = -1.138140 + i7.094290,$$

$$a_2 = 3.110090 + i2.618830,$$

$$a_1 = 0.282940 + -i0.276260,$$

$$a_0 = -0.316220 + i0.036590.$$

One gets $\|h\| = 2.9631$.

In Fig. 2 we have plotted, using GNU Octave, the basins $B(h,\eta)$ at each root η of the chosen random polynomial h are plotted, in $\mathbb{C}$ and in the Riemann sphere.

In Table 2 the same quantities are computed for the polynomial given by $a_0 = -1$, $a_1 = a_2 = \cdots = a_6 = 0$, $a_7 = 1$. In this case the roots are the seventh roots of unity, and it is not difficult to see that the actual values of $\mu(h,\eta)$, $\Theta(h,\eta)$, and $\mathrm{vol}(B(h,\eta))$

Table 2 $h(z_0, z_1) = z_1^7 - z_0^7$

Roots in $\mathbb{C}$	$\mu(h, \cdot)$	$\Theta(h, \cdot)$	$\mathrm{vol}(B(h, \cdot))$
$-0.900969 + i0.433884$	3.023716	0.7035899	0.128982π
$-0.900969 - i0.433884$	3.023716	0.8354068	0.153846π
$-0.222521 + i0.974928$	3.023716	0.7405610	0.135198π
$-0.222521 - i0.974928$	3.023716	0.7549753	0.141414π
$1.000000 + i0.000000$	3.023716	0.9128278	0.156954π
$0.623490 + i0.781831$	3.023716	0.6800328	0.135198π
$0.623490 - i0.781831$	3.023716	0.8122845	0.148407π

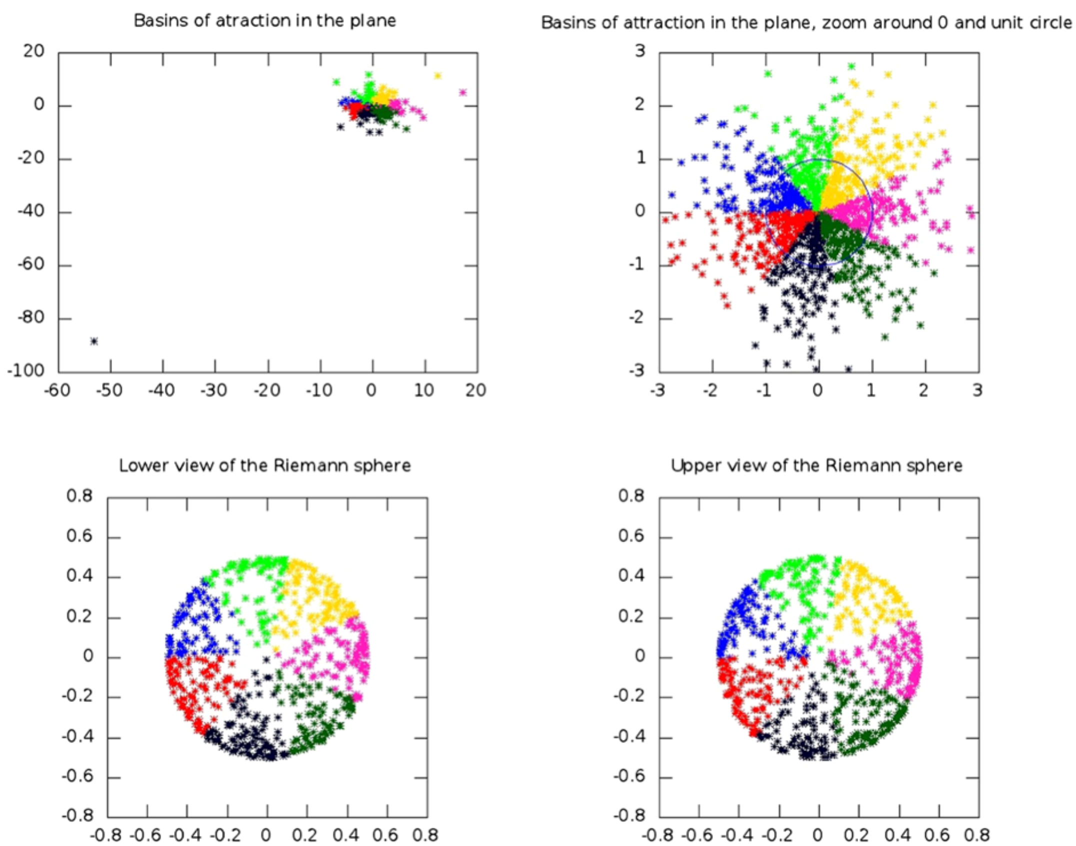


Fig. 3 The basins $B(h, \eta)$ in $\mathbb{C}$ and in the Riemann sphere for $h(z_0, z_1) = z_1^7 - z_0^7$ (GNU Octave) (Color figure online)

are constant at the roots of h by symmetry; cf. Fig. 3. This example illustrates the extent of the accuracy of the computations.

In this case we get $\|h\| = \sqrt{2}$.

The errors for the root of unity case in the third column are of the order of 25 %. But 25 % does not seem enough to explain the variation in the computed quantities in the third column of the random example where the ratio of the max to min is greater than 2. So it is likely that they are not all equal. On the other hand, the ratios of the volumes of the basins in the fourth columns of the random and roots of unity examples do seem of the same order of magnitude. So perhaps for $n = 1$ they are all equal? Also, the graphics of the basins are very encouraging in the random case. There appear to be seven connected regions with a root in each. So there is some hope that this is true in general. That is, there may generically be a root in each connected component of the basins. This would be very interesting and would be a very good start for understanding the integrals. It would be good to have some more experiments and even better some theorems.

Acknowledgements Michael Shub was supported by CONICET PIP 0801 2010-2012 and ANPCyT PICT 2010-00681. Diego Armentano was partially supported by Grupo de Investigación “Sistemas Dinámicos” (Comisión Sectorial de Investigación Científica (CSIC), Nro. 618). This paper was partially funded by Mathamsud grant “Complexity”.

References

1. E. Artin, *The Gamma Function* (Holt, Rinehart and Winston, New York, 1964).
2. R. Abraham, J. Robbin, *Transversal Mappings and Flows* (Benjamin, New York, 1967).
3. V.I. Arnold, S.M. Gusein-Zade, A.N. Varchenko, *Singularities of Differentiable Maps*, vol. 1 (Birkhäuser, Basel, 1985).
4. C. Beltrán, A continuation method to solve polynomial systems and its complexity, *Numer. Math.* **117**(1), 89–113 (2011).
5. C. Beltrán, L.M. Pardo, Smale’s 17th problem: average polynomial time to compute affine and projective solutions, *J. Am. Math. Soc.* **22**(2), 363–385 (2009).
6. C. Beltrán, L.M. Pardo, Fast linear homotopy to find approximate zeros of polynomial systems, *Found. Comput. Math.* **11**(1), 95–129 (2011).
7. C. Beltrán, M. Shub, A note on the finite variance of the averaging function for polynomial system solving, *Found. Comput. Math.* **10**(1), 115–125 (2010).
8. L. Blum, F. Cucker, M. Shub, S. Smale, *Complexity and Real Computation* (Springer, New York, 1998).
9. P. Bürgisser, F. Cucker, On a problem posed by Steve Smale, *Ann. Math.* **174**(3), 1785–1836 (2011).
10. J.-P. Dedieu, G. Malajovich, M. Shub, Adaptative step size selection for homotopy methods to solve polynomial equations, *IMAJNA* **33**(1), 1–29 (2013).
11. M. Fernandez, L.M. Pardo, An arithmetic Poisson formula for the multi-variate resultant, *J. Complexity* (2013). doi:10.1016/j.jco.2013.04.005.
12. M.-H. Kim, M. Martens, S. Sutherland, Bounds for the cost of root finding (2011). arXiv:0903.3674.
13. J. Renegar, On the efficiency of Newton’s method in approximating all zeros of a system of complex polynomials, *Math. Oper. Res.* **12**(1), 121–148 (1987).
14. M. Shub, Complexity of Bezout’s theorem. VI. Geodesics in the condition (number) metric, *Found. Comput. Math.* **9**(2), 171–178 (2009).
15. M. Shub, S. Smale, Complexity of Bézout’s theorem. I. Geometric aspects, *J. Am. Math. Soc.* **6**(2), 459–501 (1993).
16. M. Shub, S. Smale, Complexity of Bezout’s theorem. II. Volumes and probabilities, in *Computational Algebraic Geometry*. Progr. Math., vol. 109, Nice, 1992 (Birkhäuser, Boston, 1993), pp. 267–285.
17. M. Shub, S. Smale, Complexity of Bezout’s theorem. III. Condition number and packing, *J. Complex.* **9**(1), 4–14 (1993).
18. M. Shub, S. Smale, Complexity of Bezout’s theorem. IV. Probability of success; extensions, *SIAM J. Numer. Anal.* **33**(1), 128–148 (1996).
19. S. Smale, The fundamental theorem of algebra and complexity theory, *Bull., New Ser., Am. Math. Soc.* **4**(1), 1–36 (1981).
20. S. Smale, *Mathematical Problems for the Next Century, Mathematics: Frontiers and Perspectives* (Amer. Math. Soc., Providence, 2000), pp. 271–294.
21. H. Weyl, *The Classical Groups. Their Invariants and Representations* (Princeton University Press, Princeton, 1939).