ON THE GEOMETRY AND TOPOLOGY OF THE SOLUTION VARIETY FOR POLYNOMIAL SYSTEM SOLVING

CARLOS BELTRÁN AND MICHAEL SHUB

Abstract. We study the geometry and topology of the rank stratification for polynomial system solving, i. e. the set of pairs (system, solution) such that the derivative of the system at the solution has a given rank. Our approach is to study the gradient flow of the Frobenius condition number defined on each stratum.

1. Introduction

For $\mathbb{K} = \mathbb{C}$ or $\mathbb{K} = \mathbb{R}$ and positive integers n, d let $\mathcal{H}_d$ denote the vector space of homogeneous polynomials of degree d with coefficients in $\mathbb{K}$ and unknowns $X_0, \ldots, X_n$. So if $f \in \mathcal{H}_d$, $f : \mathbb{K}^{n+1} \to \mathbb{K}$. For a multi-index of non-negative integers $(\alpha) = (\alpha_0, \ldots, \alpha_n)$ with $|(\alpha)| = \sum_0^n \alpha_i = d$ let $x^{(\alpha)} = \prod_0^n x^{\alpha_i}$ be the monomial with multi-index (α). Give $\mathcal{H}_d$ the Hermitian product or inner product which makes the basis of monomials of $\mathcal{H}_d$ orthogonal and makes $||x^{(\alpha)}||^2 = \binom{d}{(\alpha)}^{-1}$ where $\binom{d}{(\alpha)}$ is the multinomial coefficient. This Hermitian (inner) product is frequently called the Bombieri-Weyl product. For a list of positive degrees $(d) = (d_1, \ldots, d_m) \in \mathbb{N}^m$, let $\mathcal{H}_{(d)} = \prod_1^m \mathcal{H}_{d_i}$. So $\mathcal{H}_{(d)}$ is the set of all systems $f = (f_1, \ldots, f_m)$ of homogeneous polynomials of respective degrees $\deg(f_i) = d_i, 1 \le i \le m$, and unknowns $X_0, \ldots, X_n$. Considered as a function, $f : \mathbb{K}^{n+1} \longrightarrow \mathbb{K}^m$. We consider $\mathcal{H}_{(d)}$ endowed with the product Hermitian (inner) product, and the corresponding norm (denoted $\| \cdot \|$).

The *solution variety* $V \subseteq \mathbb{P}(\mathcal{H}_{(d)}) \times \mathbb{P}(\mathbb{K}^{n+1})$ is defined as the set of pairs (f, ζ) such that $f(\zeta) = 0$. Observe that V is a smooth manifold (cf. [BCSS98, p. 193]), endowed with a natural Riemannian structure (and corresponding volume form) inherited from the Bombieri-Weyl product in $\mathcal{H}_{(d)}$ and the naturally induced Riemannian structure in $\mathbb{P}(\mathbb{K}^{n+1})$, which we will refer to as the Fubini-Study metric, both in the real and complex cases. We refer to this Riemannian structure in V and the metric it defines as the Fubini-Study metric.

Let $1 \le k \le \min\{m, n\}$ and consider the stratum

$$W = W^k = W^k_{(d)} = \{(f, \zeta) \in \mathbb{P}(\mathcal{H}_{(d)}) \times \mathbb{P}(\mathbb{K}^{n+1}) : f(\zeta) = 0 \text{ and } rank(Df(\zeta)) = k\},$$

and in the special case of rank 0,

$$\mathcal{W} = \{(f, \zeta) \in \mathbb{P}(\mathcal{H}_{(d)}) \times \mathbb{P}(\mathbb{K}^{n+1}) : f(\zeta) = 0 \text{ and } Df(\zeta) = 0\}.$$

C. Beltrán was partially supported by MTM2004-01167 and by a postdoctoral grant from the Spanish Government. C. Beltrán and M. Shub were partially supported by an NSERC Discovery Grant.

In this paper we will study the geometry and topology of $W^k_{(d)}$ via the gradient flow of the Frobenius condition number $\widetilde{\mu}(f,\zeta)$ defined on this variety. We recall and introduce a few definitions before we state our principal results.

The Frobenius condition number in W is defined as follows:

$$\widetilde{\mu}(f,\zeta) = \|f\| \, \|Df(\zeta)^\dagger \; Diag(\|\zeta\|^{d_i-1} d_i^{1/2})\|_F, \quad \forall (f,\zeta) \in W,$$

where $\|\cdot\|_F$ is Frobenius norm (i.e. $Trace(L^*L)^{1/2}$ where L^* is the adjoint of L) and † is Moore-Penrose pseudoinverse. The Moore-Penrose inverse $L^\dagger : \mathbb{E} \to \mathbb{F}$ of a linear operator $L : \mathbb{E} \to \mathbb{F}$ of finite dimensional Hilbert spaces is defined as the the composition

$$L^\dagger = (L \mid_{Ker(L)^\perp})^{-1} \circ \pi_{Image(L)}, \tag{1.1}$$

where $\pi_{Image(L)}$ is the orthogonal projection on image L and $Ker(L)^\perp$ is the orthogonal complement of the nullspace of L.

From Proposition 6 below, the mapping $A \mapsto A^\dagger$ restricted to the set $\mathcal{R}^k$ of rank k matrices is smooth and hence $\widetilde{\mu} : W \longrightarrow \mathbb{R}$ is also smooth, for every choice of m, n, k.

Our main interest is the case $m = n = k$ and $\mathbb{K} = \mathbb{C}$. This is the case of n homogeneous equations in $n+1$ unknowns which we have studied in a series of papers on The Complexity of Bezout's Theorem. We recall a result in this case. In [Shu07] we have bounded the number of steps of projective Newton's method sufficient to follow a homotopy $\Gamma_t = (f_t, \zeta_t)$ in W, by the length of the path Γ_t in the condition metric, using the condition number of [SS93, Shu07]. Recall that if we are given a Riemann structure on a manifold M, the length of a piecewise $\mathcal{C}^1$ curve Γ_t in M is then given by,

$$Length(\Gamma_t) = \int \|\dot{\Gamma}_t\| \, dt,$$

and a metric d on M is defined by $d(x,y) = inf Length(\gamma)$ over piecewise $\mathcal{C}^1$ differentiable paths γ in W joining x to y. Now we define the normalized condition number $\mu_{\text{norm}}(f,\zeta)$ for $(f,\zeta) \in W$ by,

$$\mu_{\text{norm}}(f,\zeta) = \|f\| \|(Df(\zeta) \mid_{\zeta^\perp})^{-1} Diag(\|\zeta\|^{d_i-1} d_i^{1/2})\|_{op},$$

or ∞ if $\det(Df(\zeta) \mid_{\zeta^\perp}) = 0$. Here, $\|\cdot\|_{op}$ denotes the operator norm of a linear map. The *Condition Riemann Structure* $\|(\dot{f}_t, \dot{\zeta}_t)\|_\kappa$, on W is defined by $\|(\dot{f}_t, \dot{\zeta}_t)\|_\kappa = \mu(f_t, \zeta_t)\|(\dot{f}_t, \dot{\zeta}_t)\|$. We denote by $D = \max\{d_i : 1 \leq i \leq m\}$ the maximum of the degrees. The Main Theorem of [Shu07] is:

Theorem 1. *Let $m = n = k$. There is a constant $C > 0$, such that: if $\Gamma_t = (f_t, \zeta_t)$ $t_0 \leq t \leq t_1$ is a C^1 path in W with the Condition Riemann Structure , then*

$$CD^{3/2} Length_\kappa(\Gamma_t)$$

steps of projective Newton method are sufficient to continue an approximate zero x_0 of f_{t_0} with associated zero ζ_0 to an approximate zero x_1 of f_{t_1} with associated zero ζ_1.

The projective Newton method is Newton's method adapted to homogeneous problems in projective space. An approximate zero x with associated zero ζ is a point for which the projective Newton method converges quadratically to the zero ζ, see [Shu07].

This theorem makes the geometry of W in the condition Riemann structure and the distance function a central object of study with potential application to the understanding of algorithms for solving square systems of polynomial equations. There is a difficulty in that the condition number μ_{norm} is not a smooth mapping. So we have introduced the Frobenius condition number which is smooth. The corresponding smooth *Frobenius Condition Riemann Structure* on W is defined by $\|(\dot{f}_t,\dot{\zeta}_t)\|_F = \widetilde{\mu}(f_t,\zeta_t)\|(\dot{f}_t,\dot{\zeta}_t)\|$. Note that $\mu_{\text{norm}}(f,\zeta) \le \widetilde{\mu}(f,\zeta) \le \sqrt{n}\mu_{\text{norm}}(f,\zeta)$. Thus the Main Theorem of [Shu07] can now be rephrased with at most an extra factor of $\sqrt{n}$ in the estimate as:

Theorem 2. *Let $m = n = k$. There is a constant $C > 0$, such that: if $\Gamma_t = (f_t,\zeta_t)$ $t_0 \le t \le t_1$ is a C^1 path in W with the Frobenius Condition Riemann Structure , then*

$$CD^{3/2} Length_F(\Gamma_t)$$

steps of projective Newton method are sufficient to continue an approximate zero x_0 of f_{t_0} with associated zero ζ_0 to an approximate zero x_1 of f_{t_1} with associated zero ζ_1.

Besides the Main Theorem of [Shu07] all the known results about μ_{norm} can be stated, up to some factors of $\sqrt{n}$, in terms of $\widetilde{\mu}$. For example, the analysis of good initial pairs for homotopy methods in [BP06, BP07] can be done using $\widetilde{\mu}$ instead of μ_{norm}.

So now we wish to study the geometry of W with the Frobenius Condition Riemann Structure and the corresponding distance function. We will do this by studying the gradient of $\widetilde{\mu}$ on W. It turns out that our analysis extends to all m, n, k so we have made our definitions and state our theorems with this generality and now we return to considering general W. The topology of W is told to us by the gradient of $\widetilde{\mu}$. It is remarkably simple! The varieties V and W would seem to be so natural in algebraic geometry that we would not be surprised if some or even most of our results are known. But we are not aware of any references.

We start with the following basic result, whose proof may be found in Section 2.

Proposition 1. *Let $\mathbb{K} = \mathbb{C}$ ($\mathbb{R}$). For any choice of $m, n, k, (d)$, the set W is a complex (real) submanifold of V of complex (real) codimension $(m-k)(n-k)$. Moreover, if any of the degrees d_i is greater than 1, then $\mathcal{W}$ is a complex (real) submanifold of V of complex (real) codimension mn.*

Let $\mathcal{B} = \mathcal{B}^k \subseteq W$ be the set of pairs (f,ζ) such that there exist unitary (orthogonal) matrices U, V of sizes $m, n+1$ respectively with

$$Diag(d_i^{-1/2}\|\zeta\|^{1-d_i})Df(\zeta) = \frac{\|f\|}{\sqrt{k}}\, U \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix} V^*,$$

For $(a_1,\dots a_m) \in \mathbb{K}^m$ we denote by $Diag(a_i)$ the $m \times m$ diagonal matrix whose ith diagonal entry is a_i. Here $Df(\zeta)$ is considered as a matrix in the standard bases and $\begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$ is an $m \times (n+1)$ matrix whose whose upper left $k \times k$ corner is the $k \times k$ identity matrix. Note that this property does not depend on the chosen affine representatives of (f,ζ). The reader may check that a pair (f,ζ) belongs to $\mathcal{B}$ if and only if there are representatives $\|f\| = \|\zeta\| = 1$ and unitary matrices U, V such that

$$f(z) = \frac{1}{\sqrt{k}} Diag\left(d_i^{1/2}\langle z,\zeta\rangle^{d_i-1}\right) U \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix} V^* z.$$

and $U\begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix} V^*(\zeta)=0$.

Let $\mathcal{G}_n$ denote $\mathcal{U}_n$, the group of $n\times n$ unitary matrices, in the case $\mathbb{K}=\mathbb{C}$ and $\mathcal{O}_n$, the group of $n\times n$ orthogonal matrices, in the case $\mathbb{K}=\mathbb{R}$. There is a natural action of the group $\mathcal{G}_{n+1}$ on V which leaves each W invariant, so we consider the action restricted to W:

$$(U,(f,\zeta)) \quad \mapsto \quad (f\circ U^*, U\zeta) \tag{1.2}$$

When all the $d_i's$ are equal the group $\mathcal{G}_m\times\mathcal{G}_{n+1}$ also acts on $\mathcal{B}$ by

$$\rho((W_1,W_2))\left(\frac{1}{\sqrt{k}}Diag\left(d_i^{1/2}\langle z,\zeta\rangle^{d_i-1}\right)U\begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}V^*z,\zeta\right)=$$

$$\left(\frac{1}{\sqrt{k}}Diag\left(d_i^{1/2}\langle z,W_2\zeta\rangle^{d_i-1}\right)W_1U\begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}V^*W_2^*z,W_2(\zeta)\right).$$

Proposition 2. (1) *$\mathcal{B}$ is the quotient of a homogeneous space of $\mathcal{G}_m\times\mathcal{G}_{n+1}$ by a further free $\mathbb{S}^1\times\mathbb{S}^1$ action in the complex case or a further free $\mathbb{Z}_2\times\mathbb{Z}_2$ action in the real case.*
(2) *$\mathcal{B}$ has a structure of fiber bundle over $\mathbb{P}(\mathbb{K}^{n+1})$ with fiber a homogeneous space of $\mathcal{G}_m\times\mathcal{G}_n$.*
(3) *If $\mathbb{K}=\mathbb{C}$, $\mathcal{B}$ has real dimension $2mk+2nk+2n-3k^2-1$.*
(4) *If $\mathbb{K}=\mathbb{R}$, $\mathcal{B}$ has real dimension $mk+nk+n-\frac{3}{2}k^2-\frac{1}{2}k$.*
(5) *For every $(f,\zeta)\in W$, we have that $\widetilde{\mu}(f,\zeta)\geq k$, and $\mathcal{B}=\{(f,\zeta):\widetilde{\mu}(f,\zeta)=k\}$.*
(6) *If $m=k\leq n$ then $\mathcal{G}_{n+1}$ acts transitively on $\mathcal{B}$.*
(7) *If all the $d_i's$ are equal $\mathcal{G}_m\times\mathcal{G}_{n+1}$ acts transitively on $\mathcal{B}$.*

Note that $\widetilde{\mu}$ is equivariant under the action (1.2) and hence by Proposition 2.(5), $\mathcal{B}$ is also invariant for this action of $\mathcal{G}_{n+1}$.

Using Proposition 2 we may identify $\mathcal{B}$ topologically. The dependence of $\mathcal{B}$ on (d) is minor! We work out some examples when $m=k\leq n$ and their homotopy groups in section 6. An example is the following result, which is immediate from the precise description of the homogeneous spaces of proposition 2, described in Section 3.

Corollary 1. *If $\mathbb{K}=\mathbb{C}$ then $\mathcal{B}$ is connected. If $\mathbb{K}=\mathbb{R}$ and k,m,n are not all equal, then $\mathcal{B}$ is connected. In the case $\mathbb{K}=\mathbb{R}$ and $k=m=n$, $\mathcal{B}$ may have one or two connected components (see Section 6 for a complete description).*

Our main theorem describes W in terms of $\mathcal{B}$.

For $b\in\mathcal{B}$, let $W^s(b)=\{x\in W:\phi_t(x)\mapsto b,t\mapsto\infty\}$, where ϕ_t is the solution flow of the vector field $V(x)=-\text{grad}\widetilde{\mu}(x)$ on W. Note that the definition of $W^s(b)$ is not very sensitive to the parametrization of ϕ_t: We may multiply $V(x)$ by a smooth positive function without changing $W^s(b)$.

Theorem 3 (Main)**.** (1) *$\widetilde{\mu}$ is an equivariant Morse function for the actions of $\mathcal{G}_{n+1}$ on W.*
(2) *$\mathcal{B}$ is the set of minima of $\widetilde{\mu}$ and the set of critical points of $\widetilde{\mu}$.*
(3) *The Hessian of $\widetilde{\mu}$ is positive definite on the normal bundle of $\mathcal{B}$.*
(4) *The $W^s(b)$ form a $\mathcal{C}^\infty$ foliation of W.*
(5) *The normal bundle $N(\mathcal{B})$ is $\mathcal{C}^\infty$ diffeomorphic to W, by a diffeomorphism $\sigma:N(\mathcal{B})\longrightarrow W$ such that*

- *$\sigma \mid_{\mathcal{B}} = Id_{\mathcal{B}}$,*
- *$D\sigma \mid_{\mathcal{B}} = Id_{T\mathcal{B}}$,*
- *σ maps the fibers of $N(\mathcal{B})$ to the W^s leaves of the foliation of W.*

The proof of this theorem can be found in section 4 for the first three items and in the appendix for the others , as well as other facts concerning conjugacies of the flow ϕ_t.

Remark 1. *Without reference to the $W^s(b)$ foliation, Theorem 3 is the simplest case of an application of the Morse-Bott Lemma (once the regularity results of Section 4 below are stated). See* [Was69] *for an equivariant version. Because we are interested in the $W^s(b)$ foliation corresponding to a particular class of conformally equivalent Riemannian structures on W we use stable manifold techniques in the appendix.*

Corollary 2. *The inclusion $i : \mathcal{B} \to W$ is a homotopy equivalence.*

Thus the homogeneous space structure for $\mathcal{B}$ given by Proposition 2 which we elaborate on in section 6 rather simply describes the diffeomorphism type of $\mathcal{B}$ and the homotopy type of W.

Let $\overline{W} = \overline{W^k_{(d)}}$ be the closure of $W \subset V$. So $\overline{W^k_{(d)}} = \mathcal{W} \cup \bigcup_{1 \le j \le k} W^j_{(d)}$ and if $k = min(m,n)$ then $\overline{W} = V$. Denote $\overline{W^k_{(d)}} - W^k_{(d)} = \Sigma'^k_{(d)} = \Sigma'$.

Proposition 3. *If $x \in W - \mathcal{B}$ then the $lim_{t \to -\infty} \phi_t(x)$ exists and is an element of Σ'. The function $\psi : W - \mathcal{B} \to \Sigma'$ defined by $\psi(x) = lim_{t \to -\infty} \phi_t(x)$ is continuous.*

ψ defines how W is attached to Σ' to form $\overline{W}$. The proof of Proposition 3 as well as comparisons of the distance between points in W and lengths of gradient curves and other useful information about the geometry of W which we prove in Section 5 follows from the next proposition (see Section 4 for a proof). The next proposition is our principal result for studying the geometry of W.

Proposition 4. *(compare to* [Dem87]*) For any pair $(f,\zeta) \in W$, the following holds.*

$$\frac{\widetilde{\mu}(f,\zeta)^2}{k}\sqrt{1 - \frac{k^2}{\widetilde{\mu}(f,\zeta)^2}} \le \|\text{grad}\widetilde{\mu}(f,\zeta)\| \le \sqrt{m}D^{3/2}\widetilde{\mu}(f,\zeta)^2.$$

In particular, $(f,\zeta) \in W$ is a regular point of $\widetilde{\mu}$ unless $(f,\zeta) \in \mathcal{B}$.

Any function satisfying Proposition 4 has many nice properties. Using Proposition 4 alone, in Section 5 below we prove with simple arguments some results for $\widetilde{\mu}$ and all rank strata that generalize results known to be true (and some times more difficult to prove) for the non-smooth condition number μ_{norm} and $m = n = k$. For example, we prove a smooth version of Proposition 1 of [BS07] (cf. Corollary 6 below), a sharp version of Theorem 1 of [SS96] (cf. Corollary 8 below), and a version of Theorem 1 of [Shu07] (cf. Corollary 5 below). All these results are valid for $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. The next proposition states some of these results, there are more in Section 5.

Proposition 5. *[See Corollaries 4, 5 and 6 below] Let $(f,\zeta) \in W$. Then,*

1 *any path in W joining (f,ζ) and some point of $\mathcal{B}$ has length in the Frobenius condition metric at least $\frac{1}{\sqrt{m}D^{3/2}} \ln \frac{\widetilde{\mu}(f,\zeta)}{k}$ while the length of the solution curve $\phi_t(f,\zeta)$ for $t \ge 0$ in the Frobenius condition metric is $\le kln|\frac{\widetilde{\mu}(f,\zeta)+\sqrt{\widetilde{\mu}(f,\zeta)^2-k^2}}{k}|$*

2 *the length of the solution curve* $\phi_t(f,\zeta)$ *for* $t \leq 0$ *in the Fubini-Study metric is* $\arcsin \frac{k}{\widetilde{\mu}(f,\zeta)}$.

3 *Let* $d((f,\zeta),\Sigma')$ *be the Fubini-Study distance from* (f,ζ) *to the set of ill-posed problems* Σ'. *Then,*

$$\frac{1}{\sqrt{m}D^{3/2}\widetilde{\mu}(f,\zeta)} \leq d((f,\zeta),\Sigma') \leq \arcsin \frac{k}{\widetilde{\mu}(f,\zeta)} \leq \frac{\pi}{2}\frac{k}{\widetilde{\mu}(f,\zeta)}.$$

4 *If* $\varepsilon = d((f,\zeta),(h,\eta))D^{3/2}\sqrt{m}\widetilde{\mu}(f,\zeta)$ *satisfies* $\varepsilon < 1$, *then*

$$\frac{\widetilde{\mu}(f,\zeta)}{1+\varepsilon} < \widetilde{\mu}(h,\eta) < \frac{\widetilde{\mu}(f,\zeta)}{1-\varepsilon}.$$

These items all have ready interpretations. Item 1 gives fairly tight bounds for the number of steps of homotopy methods starting in $\mathcal{B}$ for optimal or near optimal paths. A comparable upper bound estimate in the condition metric is the main result of [BS07]. Item 2 Establishes that the solution curve must have a limit at $t \to -\infty$. Item 3 establishes that the Frobenius condition number is comparable to the reciprocal of the distance to the ill-posed problems. Comparing the condition number to the reciprocal of the distance to the ill posed problems is a recurring theme [EY36, Dem87, SS93]. Demmel uses an estimate as proposition 4 and differential inequalities, as we do below. The analogue of item 4 for μ_{norm} was the principal new ingredient in the proof of Theorem 1. With item 4 one can prove Theorem 2 directly exactly as in the proof of Theorem 1 even for the cases $n \geq m = k$.

2. Proof of Proposition 1.

We prove the proposition for $\mathbb{K} = \mathbb{C}$, the proof being identical for $\mathbb{K} = \mathbb{R}$ substituting orthogonal groups for unitary groups. Let $\hat{W}^k = \{(f,\zeta) \in \mathcal{H}_{(d)} \times \mathbb{C}^{n+1} : f(\zeta) = 0, rank(Df(\zeta)) = k\}$ be the affine counterpart of W^k. We also denote

$$\hat{V} = \{(f,\zeta) \in \mathcal{H}_{(d)} \times \mathbb{C}^{n+1} : f(\zeta) = 0\} = \mathcal{W} \cup \bigcup_k \hat{W}^k.$$

Note that $\hat{V}$ is a submanifold of $\mathcal{H}_{(d)} \times \mathbb{C}^{n+1}$ and $\dim \hat{V} = \dim \mathcal{H}_{(d)} + n + 1 - m$ (cf. [BCSS98, pg. 194]). We use the natural notation $\hat{V}_1, \hat{W}_1^k$ for the linear case (i.e. when all the degrees $d_1,\ldots,d_m$ are equal to 1). For example, $\hat{W}_1^k = \{(M,\zeta) \in \mathcal{M}_{m\times(n+1)}(\mathbb{C}) \times \mathbb{C}^{n+1} : M\zeta = 0, rank(M) = k\}$.

Claim 1: $\hat{W}_1^k$ is a submanifold of $\hat{V}_1$ of codimension $(m-k)(n-k)$. Let $\mathcal{R}^k$ be the set of $m \times (n+1)$ matrices of rank k, which is a submanifold of codimension $(m-k)(n+1-k)$ (cf. [AVGZ86]). Then, near from $(A_0,\zeta_0) \in \hat{W}_1^k$, the set $\hat{W}_1^k$ is the pre-image of 0 under the map $\mathcal{R}^k \times \mathbb{C}^{n+1} \longrightarrow Image(A_0)$, $(A,\zeta) \mapsto \pi_{Image(A_0)}(A\zeta)$. Finally, this map is a submersion and the claim follows.

Claim 2: $\hat{W}^k$ is a submanifold of $\hat{V}$ of codimension $(m-k)(n-k)$. Consider the mapping $\phi_1 : \hat{V} \longrightarrow \hat{V}_1$, $(f,\zeta) \mapsto (Df(\zeta),\zeta)$, which is indeed a submersion. The claim follows from claim 1 as $\hat{W}^k = \phi_1^{-1}(\hat{W}_1^k)$.

Claim 3: W^k is a submanifold of V of codimension $(m-k)(n-k)$. We proceed as in [BCSS98, p. 193-4]: Note that $\hat{W}^k$ contains the product of the lines trough f and ζ for $(f,\zeta) \in W^k$. It follows that $\hat{W}^k$ is transversal to the product of the spheres of radius 1, $\mathbb{S}(\mathcal{H}_{(d)}) \times \mathbb{S}(\mathbb{C}^{n+1})$ and hence $\tilde{W}^k = \hat{W}^k \cap \mathbb{S}(\mathcal{H}_{(d)}) \times \mathbb{S}(\mathbb{C}^{n+1})$

is a smooth manifold. Then, the torus $\mathbb{S}^1 \times \mathbb{S}^1$ acts freely on $\tilde{W}^k$ and hence the quotient $W^k = \frac{\tilde{W}^k}{\mathbb{S}^1\times\mathbb{S}^1}$ is a manifold of the claimed codimension.

The assertion for $\mathcal{W}$ is proved using a similar argument, formally equivalent to the one above one with $k = 0$. Note that in the linear case, $\mathcal{W} = \emptyset$.

3. Proof of Proposition 2.

Again we do the proof for $\mathbb{K} = \mathbb{C}$, as it is identical for $\mathbb{K} = \mathbb{R}$ substituting orthogonal groups for unitary groups. We consider the affine counterpart of $\mathcal{B}^k$, $\hat{\mathcal{B}}^k \subseteq \hat{W}^k$, and its intersection with the product of the unit spheres $\tilde{\mathcal{B}}^k = \hat{\mathcal{B}}^k \cap \mathbb{S}(\mathcal{H}_{(d)}) \times \mathbb{S}(\mathbb{C}^{n+1})$. We use the natural notation $\hat{\mathcal{B}}^k_1, \tilde{\mathcal{B}}^k_1$ for the linear case (i.e. when all the degrees $d_1, \ldots, d_m$ are equal to 1).

Claim 1: $\tilde{\mathcal{B}}^k_1$ is a homogeneous space of $\mathcal{U}_m \times \mathcal{U}_{n+1}$ and is a compact real submanifold of $\hat{W}^k_1$ of (real) dimension $2mk + 2nk + 2n + 1 - 3k^2$. Note that $\tilde{\mathcal{B}}^k_1$ is the orbit of (I, e_n) where $I = \frac{1}{\sqrt{k}}\begin{pmatrix} I_k & 0 \\ 0 & 0\end{pmatrix}$, under the natural action of the compact product group $\mathcal{U}_m \times \mathcal{U}_{n+1}$ defined by $((U, V), (M, \zeta)) \mapsto (UMV^*, V\zeta)$. Thus, $\tilde{\mathcal{B}}^k_1$ is a submanifold of $\hat{W}_1$ and is diffeomorphic to $\mathcal{U}_m \times \mathcal{U}_{n+1}$ modulo the isotropy group H of I. Now,

$$H = \left\{ \left(\begin{pmatrix} U_1 & 0 \\ 0 & U_2 \end{pmatrix}, \begin{pmatrix} U_1 & 0 & 0 \\ 0 & V_2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) : U_1 \in \mathcal{U}_k, U_2 \in \mathcal{U}_{m-k}, V_2 \in \mathcal{U}_{n-k} \right\}$$

is isomorphic to $\mathcal{U}_k \times \mathcal{U}_{m-k} \times \mathcal{U}_{n-k}$ and the claim on the dimension of $\tilde{\mathcal{B}}^k_1$ follows.

Claim 2: $\tilde{\mathcal{B}}^k$ is a compact embedded submanifold of $\hat{W}^k$ diffeomorphic to $\tilde{\mathcal{B}}^k_1$ Now consider

$$\begin{array}{cccc} \phi_2 : & \tilde{\mathcal{B}}^k_1 & \longrightarrow & \hat{W}^k \\ & (M, \zeta) & \mapsto & (h, \zeta), \quad h(z) = Diag\left(d_i^{1/2}\langle z, \zeta\rangle^{d_i - 1}\right) Mz. \end{array}$$

which is an injective immersion. As $\tilde{\mathcal{B}}^k_1$ is compact, the set $\phi_2(\tilde{\mathcal{B}}^k_1)$ is a compact embedded submanifold of $\hat{W}^k$ diffeomorphic to $\tilde{\mathcal{B}}^k_1$. Finally, note that $\tilde{\mathcal{B}}^k = \phi_2(\tilde{\mathcal{B}}^k_1)$.

Claim 3: $\mathcal{B}^k$ is a compact real submanifold of W^k of (real) dimension $2mk + 2nk + 2n - 1 - 3k^2$, and it is diffeomorphic to $\frac{\tilde{\mathcal{B}}^k}{\mathbb{S}^1\times\mathbb{S}^1}$. Just note that $\mathcal{B}^k$ is the quotient of $\tilde{\mathcal{B}}^k$ under the free action of $\mathbb{S}^1 \times \mathbb{S}^1$.

Claim 4: For$(f, \zeta) \in W$, $\widetilde{\mu}(f, \zeta) \geq k$ with equality if and only if $(f, \zeta) \in \mathcal{B}$. By unitary invariance, we may assume that $\zeta = e_0$. Write $A = Diag(d_i^{-1/2})Df(e_0)$. Then,

$$\widetilde{\mu}(f, \zeta) = \|f\| \, \|A^\dagger\|_F \geq \|A\|_F \, \|A^\dagger\|_F \geq k.$$

Assume that $\widetilde{\mu}(f, \zeta) = k$ and we prove that then $(f, \zeta) \in \mathcal{B}$. Now, $\widetilde{\mu}(f, e_0) = k$ implies $\|f\| = \|A\|_F$ and $\|A\|_F\|A^\dagger\|_F = k$. In particular, f only contains non-zero monomials of the form $X_0^{d_i-1}X_j$, and the k non-zero singular values of A are equal, i.e. $(f, e_0) \in \mathcal{B}$.

Claim 5: $\mathcal{B}$ is a fiber bundle over $\mathbb{P}(\mathbb{C}^{n+1})$ with fiber a homogeneous space of $\mathcal{U}_m \times \mathcal{U}_n$. Let $\pi_2 : \mathcal{B} \longrightarrow \mathbb{P}(\mathbb{C}^{n+1})$ be the projection on the second component. Then, the action (1.2) yields an structure of fiber bundle to π_2. Note that the fiber of e_0 is the set of systems $f = (f_1, \ldots, f_m) \in \mathbb{P}(\mathcal{H}_{(d)})$ such that $f_i(z) = d_i^{1/2} \sum_{j\geq 1} z_0^{d_i-1} a_{ij} z_j$ for some matrix $A = (a_{ij}) \in \mathbb{P}(\mathcal{M}_{m\times n}(\mathbb{C}))$ such that all the singular values of A are equal. Note that $\mathcal{U}_m \times \mathcal{U}_n$ acts transitively on that space by $(U, V, A) \mapsto UAV^*$, and hence the fiber of π_2 is diffeomorphic to $\frac{\mathcal{U}_m \times \mathcal{U}_n}{H}$ where H is the isotropy group of the element $\begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$, i.e. H is the set of pairs

$$\left(\begin{pmatrix} U_1 & 0 \\ 0 & U_2 \end{pmatrix}, \begin{pmatrix} \lambda U_1 & 0 \\ 0 & V_2 \end{pmatrix} \right), U_1 \in \mathcal{U}_k, U_2 \in \mathcal{U}_{m-k}, V_2 \in \mathcal{U}_{n-k}, \lambda \in \mathbb{S}^1,$$

which is diffeomorphic to $\mathbb{S}^1 \times \mathcal{U}_k \times \mathcal{U}_{m-k} \times \mathcal{U}_{n-k}$.

Claim 6: If $m = k \leq n$ then $\mathcal{G}_{n+1}$ acts transitively on $\mathcal{B}$. Let $(f, \zeta) \in \mathcal{B}$ and choose norm 1 representatives such that

$$f(z) = \frac{1}{\sqrt{k}} Diag\left(d_i^{1/2} \langle z, \zeta \rangle^{d_i - 1} \right) U \begin{pmatrix} I_m & 0 \end{pmatrix} V^* z, \tag{3.1}$$

for $U \in \mathcal{G}_m, V \in \mathcal{G}_{n+1}$. We can choose V such that $V^* \zeta = e_n$. Then, note that

$$U \begin{pmatrix} I_m & 0 \end{pmatrix} V^* z = \begin{pmatrix} I_m & 0 \end{pmatrix} R^*, \quad R^* = \begin{pmatrix} U & 0 \\ 0 & I_{n+1-m} \end{pmatrix} V^*,$$

so that $(f, \zeta) = (g \circ R^*, Re_n)$ as wanted.

Claim 7: If all the $d_i's$ are equal $\mathcal{G}_m \times \mathcal{G}_{n+1}$ acts transitively on $\mathcal{B}$. This is clear from the expression (3.1). Note that only in this case is the action by $\mathcal{G}_m \times \mathcal{G}_{n+1}$ well defined.

□

4. The Frobenius condition number and its gradient flow

4.1. **Gradient of $\widetilde{\mu}$.** In this section we compute lower and upper bounds for the norm of the gradient of $\widetilde{\mu}$.

Proposition 6. *Let $\mathcal{R}_k$ be the manifold of $m \times (n+1)$ rank k matrices of Frobenius norm equal to some $c > 0$. Let $A \in \mathcal{R}_k$, $\dot{A} \in T_A \mathcal{R}_k$. Then, $\psi(A) = \|A^\dagger\|_F$ defined on $\mathcal{R}_k$ is a smooth function and*

$$D\psi(A)(\dot{A}) = -\frac{\mathbb{R}e \langle A^\dagger, A^\dagger \dot{A} A^\dagger \rangle_F}{\|A^\dagger\|_F}.$$

Proof. Smoothness of ψ follows from that of †. This last is a well-known folk result but we do not find an appropriate reference, so we include a short proof for completeness. Recall from (1.1) that

$$A^\dagger = (A\mid_{Ker(A)^\perp})^{-1} \circ \pi_{Image(A)},$$

where $Ker(A)^\perp$ is the orthogonal complement of the kernel of A and $Image(A)$ is its image. Now, $Ker(A)^\perp = Image(A^*)$ so all these subspaces move smoothly with A and so does $A^\dagger$.

Now we compute $D\psi$ using implicit differentiation. Recall the following well-known identities

$$AA^\dagger A = A, \quad A^\dagger A A^\dagger = A^\dagger, \quad (A^\dagger)^* = (A^\dagger)^* A^\dagger A = AA^\dagger (A^\dagger)^*. \tag{4.1}$$

Let P be any matrix of the same size of $A^\dagger$. From identities (4.1) and the elementary properties of $\langle \cdot, \cdot \rangle_F$, it is easy to prove that

$$\langle A^\dagger, P \rangle_F = \langle A^\dagger, A^\dagger A P A A^\dagger \rangle_F. \tag{4.2}$$

Let $\dot{A}$ be tangent to $\mathcal{R}_k$ at A, and let B denote the derivative of Moore-Penrose inverse at A in the direction of $\dot{A}$. From (4.1),

$$ABA = \dot{A} - \dot{A}A^\dagger A - AA^\dagger \dot{A}. \tag{4.3}$$

and hence

$$\langle A^\dagger, B\rangle_F = \langle A^\dagger, A^\dagger ABAA^\dagger\rangle_F = \langle A^\dagger, A^\dagger(\dot{A} - \dot{A}A^\dagger A - AA^\dagger \dot{A})A^\dagger\rangle_F =$$
$$\langle A^\dagger, A^\dagger\dot{A}A^\dagger - A^\dagger\dot{A}A^\dagger AA^\dagger - A^\dagger AA^\dagger\dot{A}A^\dagger\rangle_F = -\langle A^\dagger, A^\dagger\dot{A}A^\dagger\rangle_F.$$

Thus,

$$D\psi(A)(\dot{A}) = \frac{\mathbb{R}e\langle A^\dagger, B\rangle_F}{\|A^\dagger\|_F} = -\frac{\mathbb{R}e\langle A^\dagger, A^\dagger\dot{A}A^\dagger\rangle_F}{\|A^\dagger\|_F}.$$

□

Lemma 1. *For any $k \times k$ diagonal matrix P, the following inequality holds:*

$$\frac{\|P^3\|_F^2 - \|P\|_F^4}{\|P\|_F^2} \geq \frac{\|P\|_F^4}{k^2}\left(1 - \frac{k^2}{\|P\|_F^2}\right).$$

Proof. Note that the inequality of the lemma is equivalent to

$$\|P^3\|_F^2 - \|P\|_F^4 \geq \frac{\|P\|_F^6}{k^2} - \|P\|_F^4,$$

so it suffices to prove that $\|P\|_F^6 \leq k^2\|P^3\|_F^2$. But this is true for the inequality

$$(\sigma_1 + \cdots + \sigma_k)^3 \leq k^2(\sigma_1^3 + \cdots + \sigma_k^3)$$

holds for every choice of real non-negative numbers $\sigma_1, \ldots, \sigma_k$. □

Proof of Proposition 4. Note that $\widetilde{\mu}$ is unitarily invariant so it suffices to prove the result for the case that $\zeta = e_0$. Write $A = Diag(d_i^{-1/2})Df(e_0)$, and choose a representative of f such that $\|f\| = 1$, so that

$$\widetilde{\mu}(f, e_0) = \|A^\dagger\|_F = \|P^{-1}\|_F,$$

where P is the diagonal matrix whose entries are the singular values of A. Let $(\dot{f}, \dot{\zeta}) \in T_{(f,\zeta)}W$ so that $\langle f, \dot{f}\rangle = 0$. From Proposition 6,

$$D\widetilde{\mu}(f, e_0)(\dot{f}, \dot{\zeta}) = -\frac{\mathbb{R}e\langle A^\dagger, A^\dagger(B + C_{\dot{\zeta}})A^\dagger\rangle_F}{\|A^\dagger\|_F} = -\frac{\mathbb{R}e\langle (A^\dagger)^*A^\dagger(A^\dagger)^*, B + C_{\dot{\zeta}}\rangle_F}{\|A^\dagger\|_F}.$$

where $B = Diag(d_i^{-1/2})D\dot{f}(e_0)$ and $C_{\dot{\zeta}} = Diag(d_i^{-1/2})D^2f(e_0)(\dot{\zeta}, \cdot)$ is seen as a $m \times (n+1)$ matrix. Recall that from the higher derivative estimate of [SS93],

$$\|B\|_F \leq \sqrt{m}D^{1/2}\|\dot{f}\|, \quad \|C_{\dot{\zeta}}\|_F \leq \sqrt{m}D^{1/2}(D-1)\|\dot{\zeta}\|.$$

The upper bound for the norm of the gradient follows immediately:

$$\|\mathrm{grad}\widetilde{\mu}(f, e_0)\| \leq \sqrt{m}(D^{1/2} + D^{1/2}(D-1))\widetilde{\mu}(f, e_0)^2 = \sqrt{m}D^{3/2}\widetilde{\mu}(f, e_0)^2.$$

For the lower bound, let $\dot{\zeta} = 0$ and $\dot{h}$ be defined as

$$\dot{h}(z) = Diag(d_i^{1/2}z_0^{d_i-1})(A^\dagger)^*A^\dagger(A^\dagger)^*z,$$

so that $Diag(d_i^{-1/2})D\dot{h}(e_0) = (A^\dagger)^*A^\dagger(A^\dagger)^*$. Then, let

$$\dot{f} = \pi_{f^\perp}\dot{h} = \dot{h} - \langle \dot{h}, f\rangle f.$$

Note that $\dot f(e_0)=0$ and from [BCSS98, Lemma 17, chap. 12],

$$\|\dot h\|^2=\|(A^\dagger)^*A^\dagger(A^\dagger)^*\|_F^2=\|P^{-3}\|_F^2,$$
$$\langle \dot h,f\rangle=\langle (A^\dagger)^*A^\dagger(A^\dagger)^*,A\rangle=\|P^{-1}\|_F^2,$$
$$Diag(d_i^{-1/2})D\dot f(e_0)=(A^\dagger)^*A^\dagger(A^\dagger)^*-\|P^{-1}\|_F^2A.$$

Hence,

$$\|\dot f\|^2=\|\dot h\|^2-|\langle \dot h,f\rangle|^2=\|P^{-3}\|_F^2-\|P^{-1}\|_F^4,$$

and

$$D\widetilde{\mu}(f,e_0)(\dot f,0)=-\frac{\mathbb{R}e\langle (A^\dagger)^*A^\dagger(A^\dagger)^*,(A^\dagger)^*A^\dagger(A^\dagger)^*-\|P^{-1}\|_F^2A\rangle_F}{\widetilde{\mu}(f,e_0)}=$$
$$-\frac{\|P^{-3}\|_F^2-\|P^{-1}\|_F^4}{\|P^{-1}\|_F}.$$

Thus,

$$\|\mathrm{grad}\widetilde{\mu}(f,\zeta)\|=\|D\widetilde{\mu}(f,\zeta)\|\geq\left(\frac{\|P^{-3}\|_F^2-\|P^{-1}\|_F^4}{\|P^{-1}\|_F^2}\right)^{1/2},$$

and the lower bound follows from Lemma 1.

4.2. **Hessian of $\widetilde{\mu}$.** Recall that $p=(f,\zeta)\in\mathcal{B}$ is a critical point of $\widetilde{\mu}$ and

$$\mathrm{Hess}(\widetilde{\mu})(p)(w,w)=X(X(\widetilde{\mu})),$$

where $w\in T_{(f,\zeta)}W$ and X is a (local) vector field in W such that $X(p)=w$. This definition does not depend on the choice of X.

It is not easy to compute the Hessian of $\widetilde{\mu}$ but we can prove the following

Proposition 7. *For $p=(f,\zeta)\in\mathcal{B}$ and $w\in T_pW$, $\mathrm{Hess}(\widetilde{\mu})(p)(w,w)$ is positive unless $w\in T_p\mathcal{B}$.*

The proof uses several technical results.

Proposition 8. *With the notations of Proposition 6, let $A\in\mathcal{R}^k$ be such that its k non-zero singular values are equal. Then, A is a critical point of ψ and*

$$\frac{c^3}{k}D^2\psi(A)(\dot A,\dot A)=\frac{2k}{c^2}\mathbb{R}e\langle A,\dot AA^*\dot A\rangle_F-\frac{k}{c^2}\|\dot AA^*\|_F^2-\frac{k}{c^2}\|A^*\dot A\|_F^2+$$
$$\|\dot A\|_F^2+\frac{3k^2}{c^4}\|A^*\dot AA^*\|_F^2.$$

Proof. Let B,C be the first and second derivatives of the Moore-Penrose inverse along some fixed curve γ, $\gamma(0)=A$, $\dot\gamma(0)=\dot A$, $\ddot\gamma(0)=\ddot A$, and assume that A has all its non-zero singular values equal to $ck^{-1/2}$, which implies indeed that $A^\dagger=\frac{k}{c^2}A^*$. Thus, for any matrix P of the same size of A,

$$\langle A^\dagger,A^\dagger PA^\dagger\rangle_F=\langle (A^\dagger)^*A^\dagger(A^\dagger)^*,P\rangle_F=\frac{k^2}{c^4}\langle AA^\dagger A,P\rangle_F=\frac{k^2}{c^4}\langle A,P\rangle_F,$$

Note that

$$D^2\psi(A)(\dot A,\dot A)=\frac{d^2}{dt^2}\mid_{t=0}(\|A^\dagger+tB+\frac{t^2}{2}C+o(t^2)\|_F)=\frac{\mathbb{R}e\langle A^\dagger,C\rangle_F+\|B\|_F^2}{\|A^\dagger\|_F}. \tag{4.4}$$

We compute each of these terms. From (4.3),

$$ACA=\ddot A-\ddot AA^\dagger A-AA^\dagger\ddot A-2\dot ABA-2AB\dot A-2\dot AA^\dagger\dot A. \tag{4.5}$$

Again, using (4.2) and (4.1), we conclude

$$\langle A^\dagger, C\rangle_F = -\langle A^\dagger, A^\dagger \ddot{A} A^\dagger\rangle_F - \langle A^\dagger, A^\dagger(2\dot{A}BA + 2AB\dot{A} + 2\dot{A}A^\dagger\dot{A})A^\dagger\rangle_F.$$

Hence,

$$\langle A^\dagger, C\rangle_F = \frac{k^2}{c^4}\left(-\langle A, \ddot{A}\rangle_F - 2\langle A, \dot{A}BA + AB\dot{A} + \dot{A}A^\dagger\dot{A}\rangle_F\right) \tag{4.6}$$

Moreover, γ is contained in the sphere of Frobenius norm c and hence $\mathbb{R}e\langle A, \ddot{A}\rangle_F = -\|\dot{A}\|_F^2$. On the other hand, $A^* = A^\dagger A A^*$ so $\dot{A}^* = BAA^* + A^\dagger\dot{A}A^* + A^\dagger A\dot{A}^*$, which implies

$$\langle A, \dot{A}BA\rangle_F = \langle I, \dot{A}BAA^*\rangle_F = \|\dot{A}\|^2 - \langle A, \dot{A}A^\dagger\dot{A}\rangle_F - \langle \dot{A}, \dot{A}A^\dagger A\rangle_F. \tag{4.7}$$

Similarly,

$$\langle A, AB\dot{A}\rangle_F = \|\dot{A}\|^2 - \langle A, \dot{A}A^\dagger\dot{A}\rangle_F - \langle \dot{A}, AA^\dagger\dot{A}\rangle_F. \tag{4.8}$$

We put together (4.6), (4.7) and (4.8) to conclude

$$\mathbb{R}e\langle A^\dagger, C\rangle_F = \frac{2k^3}{c^6}\left(\mathbb{R}e\langle A, \dot{A}A^*\dot{A}\rangle_F + \|\dot{A}A^*\|_F^2 + \|A^*\dot{A}\|_F^2\right) - \frac{3k^2}{c^4}\|\dot{A}\|_F^2. \tag{4.9}$$

The same formulas used to obtain equation (4.9) help to compute $\|B\|_F^2$. We omit most of the details, the reader may reconstruct the following equality

$$\|B\|_F^2 = \frac{k^4}{c^8}\left\|\frac{2c^2}{k}\dot{A}^* - \dot{A}^*AA^* - A^*A\dot{A}^* - A^*\dot{A}A^*\right\|_F^2.$$

Then, expand the terms in the squared norm and simplify to conclude

$$\|B\|_F^2 = \frac{4k^2}{c^4}\|\dot{A}\|_F^2 - \frac{3k^3}{c^6}\|A^*\dot{A}\|_F^2 - \frac{3k^3}{c^6}\|\dot{A}A^*\|_F^2 + \frac{3k^4}{c^8}\|A^*\dot{A}A^*\|_F^2 \tag{4.10}$$

The proposition follows from equations (4.4), (4.9) and (4.10). □

Corollary 3. *With the notations of Proposition 6, let $A \in \mathcal{R}_k$ be such that all its non-zero singular values are equal to* 1*, and let ζ be a projective point such that $A\zeta = 0$. Let*

$$T = \{\dot{A} \in T_A\mathcal{R}_k : D^2\psi(A)(\dot{A}, \dot{A}) = 0, \langle A, \dot{A}\rangle_F = 0, \dot{A}\zeta = 0\}.$$

Then, T has real dimension

- $2kn + 2km - 3k^2 - 1$ *if* $\mathbb{K} = \mathbb{C}$,
- $kn + km - \frac{3}{2}k^2 - \frac{1}{2}k$ *if* $\mathbb{K} = \mathbb{R}$.

Proof. First, assume that $A = I$ is such that the first k entries of its main diagonal are 1 and the rest of entries are 0. So, the zero ζ is in the subspace spanned by the last $n + 1 - k$ vectors of the standard basis. Let $\dot{A} \in T_A\mathcal{R}_k$ and write

$$I = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix} \quad \dot{A} = \begin{pmatrix} \dot{A}_1 & \dot{A}_2 \\ \dot{A}_3 & 0 \end{pmatrix},$$

where the different blocks of I and $\dot{A}$ are compatible, and $\dot{A}_1$ is a square matrix. From the expression for $D^2\psi(I)(\dot{A}, \dot{A})$ of Proposition 8, we conclude that $\dot{A} \in T$ if and only if $\mathbb{R}e\langle \dot{A}_1, \dot{A}_1 + \dot{A}_1^*\rangle_F = 0$. Now, $\mathbb{R}e\langle \dot{A}_1, \dot{A}_1 + \dot{A}_1^*\rangle_F = 0$ implies $\mathbb{R}e\langle \dot{A}_1^*, \dot{A}_1 + \dot{A}_1^*\rangle_F = 0$, and hence $\dot{A}_1 + \dot{A}_1^* = 0$. We conclude that

$$T = \left\{\dot{A} = \begin{pmatrix} \dot{A}_1 & \dot{A}_2 \\ \dot{A}_3 & 0 \end{pmatrix} : \dot{A}_1 \text{ is skew-symmetric and } trace(\dot{A}_1) = 0, \dot{A}\zeta = 0\right\},$$

and the corollary follows in this case. Now, for general A, write a singular value decomposition $A = UIV^*$ with U and V unitary (orthogonal if $\mathbb{K} = \mathbb{R}$), and note that ψ is invariant under multiplication by U and V. That is, $\psi(P) = \psi(U^* PV)$ for any matrix P. Hence,

$$D^2\psi(A)(\dot{A}, \dot{A}) = D^2\psi(I)(U^*\dot{A}V, U^*\dot{A}V).$$

Namely, the dimension of T is the same that for the case $A = I$. □

Assume that representatives such that $\|f\| = \sqrt{k}$, $\|\zeta\| = 1$ are chosen.

Lemma 2. *Let $w = h + v$ where $v \in T_p\mathcal{B}$. Then,*

$$\mathrm{Hess}(\widetilde{\mu})(p)(w, w) = Hess(\widetilde{\mu})(p)(h, h).$$

Proof. Note that $v \in T_p\mathcal{B}$ implies $\mathrm{Hess}(\widetilde{\mu})(p)(v, v) = 0$. Now, the Hessian is a bilinear form and as p is a local minimum it is positive semi-definite. Thus, $\mathrm{Hess}(\widetilde{\mu})(p)(v, v) = 0$ implies $\mathrm{Hess}(\widetilde{\mu})(p)(h, v) = 0$ for every h and the lemma follows. □

Let $V_p \subseteq T_pW$ be the set of tangent vectors $(\dot{f}, 0)$ such that $\mathrm{Hess}(\widetilde{\mu})(f, \zeta)(v, v) = 0$, and let $H_p = \{(\dot{h}, 0) : \langle(\dot{h}, 0), (\dot{f}, 0)\rangle = 0 \ \forall\ (\dot{f}, 0) \in V_p\} \subseteq T_pW$. Note that the real codimension of H_p in T_pW equals

$$\mathrm{codim}_{\mathbb{R}}(H_p) = \dim_{\mathbb{R}}(\mathbb{P}(\mathbb{K}^{n+1})) + \dim_{\mathbb{R}}(V_p).$$

Lemma 3. *The real dimension of V_p equals*

- $2kn + 2km - 3k^2 - 1$ *if* $\mathbb{K} = \mathbb{C}$,
- $kn + km - \frac{3}{2}k^2 - \frac{1}{2}k$ *if* $\mathbb{K} = \mathbb{R}$.

Proof. By unitary invariance, we may assume that $\zeta = e_0$. Let $(f_t, e_0) \subseteq W$ be a $\mathcal{C}^1$ curve, $p = (f_0, e_0) = (f, e_0)$, with tangent vector $(\dot{f}_t, 0)$. Denote $A_t = Diag(d_i^{-1/2})Df_t(e_0)$ and $\dot{A}_t = Diag(d_i^{-1/2})D\dot{f}_t(e_0)$, $\ddot{A}_t = Diag(d_i^{-1/2})D\ddot{f}_t(e_0)$. Choose representatives such that $\|A_t\|_F^2 = k, \forall\ t$. In particular, all the non-zero singular values of A_0 are equal to 1, and

$$0 = \sqrt{k}\frac{d^2}{dt^2} \mid_{t=0} \|A_t\| = \|\dot{A}_0\|_F^2 + \mathbb{R}e\langle A_0, \ddot{A}_0\rangle_F.$$

Note that $p \in \mathcal{B}$ implies $\langle f, h\rangle = \langle A_0, Diag(d_i^{-1/2})Dh(e_0)\rangle_F$ for every h such that $h(e_0) = 0$. Hence, we have

$$\sqrt{k}\frac{d}{dt} \mid_{t=0} \|f_t\| = \mathbb{R}e\langle f, \dot{f}_0\rangle = \mathbb{R}e\langle A_0, \dot{A}_0\rangle_F = 0,$$

$$\sqrt{k}\frac{d^2}{dt^2} \mid_{t=0} \|f_t\| = \|\dot{f}\|^2 + \mathbb{R}e\langle f, \ddot{f}_0\rangle \geq \|\dot{A}_0\|_F^2 + \mathbb{R}e\langle A_0, \ddot{A}_0\rangle_F = 0,$$

the inequality bing strict if $\|\dot{f}_0\| \neq \|\dot{A}_0\|_F$, i.e. if $\dot{f}$ has some non-zero monomial containing $X_0^{d_i - l}$ for some i and $l \geq 2$.

Note that

$$\mathrm{Hess}(\widetilde{\mu})(f, e_0)(v, v) = \frac{d^2}{dt^2} \mid_{t=0} (\widetilde{\mu}(f_t, e_0)) = \frac{d^2}{dt^2} \mid_{t=0} (\|f_t\|\ \|A_t^\dagger\|_F) =$$

$$\sqrt{k}\frac{d^2}{dt^2} \mid_{t=0} (\|f_t\|) + \sqrt{k}\frac{d^2}{dt^2} \mid_{t=0} (\|A_t^\dagger\|_F) \geq \sqrt{k}\frac{d^2}{dt^2} \mid_{t=0} (\|A_t^\dagger\|_F).$$

The lemma then follows from Corollary 3. □

Proof of Proposition 7. From the definition of H_p, we have that $H_p \cap T_p\mathcal{B} = \{0\}$. Moreover, from Lemma 3 and Proposition 2, the codimension of H_p in T_pW is equal to the dimension of $\mathcal{B}$ (both for $\mathbb{K} = \mathbb{R}, \mathbb{C}$).

We conclude that H_p and $T_p\mathcal{B}$ are complementary subspaces of T_pW. Let $w = h + v$ where $v \in T_p\mathcal{B}$ and $h \in H_p$. Then, from Lemma 2 and the definition of H_p,

$$\mathrm{Hess}(\widetilde{\mu})(p)(w,w) = \mathrm{Hess}(\widetilde{\mu})(p)(h,h) > 0,$$

unless $h = 0$ (i.e. unless $w \in T_p\mathcal{B}$).

□

5. Distances and integral curves of the gradient flow

In this section the symbol X denotes the smooth vector field in $W \setminus \mathcal{B}$ defined by

$$X(f,\zeta) = \frac{grad\widetilde{\mu}(f,\zeta)}{\|grad\widetilde{\mu}(f,\zeta)\|^2}.$$

It is easy to see that for $p = (f_0, \zeta_0) \in W \setminus \mathcal{B}$, the integral curve $\alpha_p(t)$ is defined for $t \in (k - \widetilde{\mu}(p), \infty)$ and

$$\widetilde{\mu}(\alpha_p(t)) = \widetilde{\mu}(p) + t.$$

Lemma 4.

$$\int \frac{k}{s\sqrt{s^2 - k^2}}\, ds = \arccos\frac{k}{s}, \qquad \int \frac{k}{\sqrt{s^2 - k^2}}\, ds = k\ln(s + \sqrt{s^2 - k^2})$$

Proposition 9. *Let $\alpha : [a,b] \longrightarrow W$ be a piecewise $\mathcal{C}^1$ curve, and let $L(a,b)$ be the length of α in the Fubini-Study metric. Then,*

$$L(a,b) \geq \frac{1}{\sqrt{m}D^{3/2}} \left| \frac{1}{\widetilde{\mu}(\alpha(a))} - \frac{1}{\widetilde{\mu}(\alpha(b))} \right|.$$

Moreover, if α is an integral curve of X, then

$$L(a,b) \leq \left| \arccos\frac{k}{\widetilde{\mu}(\alpha(b))} - \arccos\frac{k}{\widetilde{\mu}(\alpha(a))} \right|$$

Proof. For the lower bound, let α be identified with its image and let

$$\phi = \widetilde{\mu}\,|_\alpha : \begin{array}{ccc} \alpha & \longrightarrow & \mathbb{R} \\ (f,\zeta) & \mapsto & \widetilde{\mu}(f,\zeta), \end{array}$$

and change variables with ϕ to conclude

$$L(a,b) \geq \int_{s\in[\widetilde{\mu}(\alpha(a)),\widetilde{\mu}(\alpha(b))]} \frac{1}{J(\phi)(f_s,\zeta_s)}\, ds, \tag{5.1}$$

where (f_s, ζ_s) is some pair in α such that $\phi(f_s,\zeta_s) = s$ and $J(\phi)(f_s,\zeta_s)$ is the Jacobian of ϕ at (f_s,ζ_s). Note that we have implicitly restricted to the subset of points of α that are regular points of ϕ. Now, a lower bound for $J(\phi)(f_s,\zeta_s)$ is easy to obtain:

$$J(\phi)(f_s,\zeta_s) = \|D\phi(f_s,\zeta_s)\| \leq \|D\widetilde{\mu}(f_s,\zeta_s)\| = \|\mathrm{grad}\widetilde{\mu}(f_s,\zeta_s)\|,$$

and from Proposition 4 we conclude that

$$L(a,b) \geq \int_{s\in[\widetilde{\mu}(\alpha(a)),\widetilde{\mu}(\alpha(b))]} \frac{1}{\sqrt{m}D^{3/2}s^2}\, ds,$$

and the lower bound of the Proposition follows.

For the upper bound, note that if α is an integral curve for X,

$$L(a,b)=\int_a^b \|\dot\alpha(s)\|\,ds=\int_a^b \frac{1}{\|\mathrm{grad}\widetilde{\mu}(\alpha(s))\|}\,ds,$$

and again from Proposition 4,

$$L(a,b)\le\int_a^b \frac{k}{\widetilde{\mu}(\alpha(s))^2\sqrt{1-\frac{k^2}{\widetilde{\mu}(\alpha(s))^2}}}\,ds=$$

$$\int_a^b \frac{k}{(\widetilde{\mu}(\alpha(a))+s-a)^2\sqrt{1-\frac{k^2}{(\widetilde{\mu}(\alpha(a))+s-a)^2}}}\,ds=\int_{\widetilde{\mu}(\alpha(a))}^{\widetilde{\mu}(\alpha(b))}\frac{k}{s\sqrt{s^2-k^2}}\,ds$$

and the upper bound of the Proposition follows from Lemma 4. □

Corollary 4. *[See Prop. 5, item 4] For any two pairs $(f,\zeta),(h,\eta)\in W$,*

$$\left|\frac{1}{\widetilde{\mu}(f,\zeta)}-\frac{1}{\widetilde{\mu}(h,\eta)}\right|\le d((f,\zeta),(h,\eta))\sqrt{m}D^{3/2}.$$

In particular, if $\varepsilon=d((f,\zeta),(h,\eta))D^{3/2}\sqrt{m}\widetilde{\mu}(f,\zeta)$ satisfies $\varepsilon<1$, then

$$\frac{\widetilde{\mu}(f,\zeta)}{1+\varepsilon}<\widetilde{\mu}(h,\eta)<\frac{\widetilde{\mu}(f,\zeta)}{1-\varepsilon}.$$

Proof. Immediate using the lower bound for the length of a curve given by Proposition 9. □

Proposition 10. *[See Prop. 5, item 1] Let $\alpha:[a,b]\longrightarrow W$ be curve, and let $L_F(a,b)$ be the length of α in the Frobenius condition metric. Then,*

$$L_F(a,b)\ge\frac{1}{\sqrt{m}D^{3/2}}\left|\ln\frac{\widetilde{\mu}(\alpha(b))}{\widetilde{\mu}(\alpha(a))}\right|.$$

Moreover, if α is an integral curve of X, then

$$L_F(a,b)\le k\left|\ln\frac{\widetilde{\mu}(\alpha(b))+\sqrt{\widetilde{\mu}(\alpha(b))^2-k^2}}{\widetilde{\mu}(\alpha(a))+\sqrt{\widetilde{\mu}(\alpha(a))^2-k^2}}\right|$$

Proof. The proof is the same as that of Proposition 9, but now we use the Frobenius condition metric so we must bound

$$L_F(a,b)=\int_{(f,\zeta)\in\alpha}\widetilde{\mu}(f,\zeta)\,d\alpha,$$

That is, for any curve α,

$$L_F(a,b)\ge\int_{s\in[\widetilde{\mu}(\alpha(a)),\widetilde{\mu}(\alpha(b))]}\frac{1}{\sqrt{m}D^{3/2}s}\,ds,$$

and for α an integral curve of X,

$$L_F(a,b)\le\int_{\widetilde{\mu}(\alpha(a))}^{\widetilde{\mu}(\alpha(b))}\frac{k}{\sqrt{s^2-k^2}}\,ds,$$

which is computed using Lemma 4. □

Corollary 5. *[See Prop 5, items 2 and 3] Let* $(f,\zeta)\in W=W^k$ *and let* $d((f,\zeta),\Sigma')$ *be the Fubini-Study distance from* (f,ζ) *to the set of ill-posed problems* $\Sigma'=\Sigma'^k=\{(f,\zeta)\in V: rank(Df(\zeta))\le k-1\}$. *Then,*

$$\frac{1}{\sqrt{m}D^{3/2}\widetilde{\mu}(f,\zeta)}\le d((f,\zeta),\Sigma')\le \arcsin\frac{k}{\widetilde{\mu}(f,\zeta)}\le\frac{\pi}{2}\frac{k}{\widetilde{\mu}(f,\zeta)}.$$

Moreover, let $d((f,\zeta),\mathcal{B})$ *be the Fubini-Study distance from* (f,ζ) *to the set* $\mathcal{B}$. *Then,*

$$\frac{1}{\sqrt{m}D^{3/2}}\left(\frac{1}{k}-\frac{1}{\widetilde{\mu}(f,\zeta)}\right)\le d((f,\zeta),\mathcal{B})\le \arccos\frac{k}{\widetilde{\mu}(f,\zeta)}.$$

Proof. Immediate from Proposition 9. □

Corollary 6. *Let* $(f,\zeta)\in W$ *and let* $d((f,\zeta),\mathcal{B})$ *be the distance in the Frobenius condition metric from* (f,ζ) *to the set* $\mathcal{B}$. *Then,*

$$\frac{1}{\sqrt{m}D^{3/2}}\ln\frac{\widetilde{\mu}(f,\zeta)}{k}\le d((f,\zeta),\mathcal{B})\le k\ln\frac{\widetilde{\mu}(f,\zeta)+\sqrt{\widetilde{\mu}(f,\zeta)^2-k^2}}{k}.$$

Proof. Immediate from Proposition 10. □

Corollary 7. *Let* $\widetilde{\mu}_j=\widetilde{\mu}(f_j,\zeta_j)$, $j=1,2$, *and let* $\gamma\subseteq W$ *be a* $\mathcal{C}^1$ *embedded curve joining* (f_1,ζ_1) *and* (f_2,ζ_2). *Assume that* $\widetilde{\mu}_1\le\widetilde{\mu}_2$ *and let* $MAX=\max_{(f,\zeta)\in\gamma}\widetilde{\mu}(f,\zeta)$ *and* $MIN=\min_{(f,\zeta)\in\gamma}\widetilde{\mu}(f,\zeta)$. *Then,*

$$\frac{MAX}{MIN}\le\sqrt{\frac{\widetilde{\mu}_2}{\widetilde{\mu}_1}}\,e^{\frac{\sqrt{m}D^{3/2}Length_F(\gamma)}{2}},$$

where $Length_F(\gamma)$ *is the length of* γ *in the Frobenius condition metric.*

Proof. Note that for every $t\in(MIN,\widetilde{\mu}_1)\cup(\widetilde{\mu}_2,MAX)$ there are at least two points p_1,p_2 in γ such that $\widetilde{\mu}(p_1)=\widetilde{\mu}(p_2)=t$. Hence, changing variables as in the proof of Proposition 10 we obtain

$$\sqrt{m}D^{3/2}Length_F(\gamma)\ge 2\int_{s\in(MIN,\widetilde{\mu}_1)}\frac{1}{s}\,ds+\int_{s\in[\widetilde{\mu}_1,\widetilde{\mu}_2]}\frac{1}{s}\,ds+2\int_{s\in(\widetilde{\mu}_2,MAX)}\frac{1}{s}\,ds=$$

$$\ln\frac{MAX^2\widetilde{\mu}_1}{MIN^2\widetilde{\mu}_2},$$

and the corollary follows. □

Corollary 8. *Let* $k=m\le n$, $(f,\zeta)\in W$ *and let* $h\in\mathbb{P}(\mathcal{H}_{(d)})$ *be a system such that*

$$\varepsilon=2D^{3/2}\sqrt{\frac{m^2+1}{m}}\,d(f,h)\widetilde{\mu}(f,\zeta)^2<1.$$

Then, ζ *can be continued to a zero* η *of* h *and*

$$\widetilde{\mu}(h,\eta)\le\frac{\widetilde{\mu}(f,\zeta)}{\sqrt{1-\varepsilon}}.$$

Proof. Let $\{f_t:0\le t\le d(f,h)\}\subseteq\mathbb{P}(\mathcal{H}_{(d)})$ be a geodesic with unit length tangent vector joining f and h. Let $\alpha(t)=(f_t,\zeta_t)$ be the horizontal lift of that curve to W, so that

$$\|\dot{\zeta}_t\|\le\widetilde{\mu}(f_t,\zeta_t)\|\dot{f}_t\|. \tag{5.2}$$

From the implicit function theorem, $\alpha(t)$ is defined for $t \in (0, s)$ for some $s > 0$. Note that from Proposition 4, if $t \in (0, s)$,

$$\frac{d}{dt}\widetilde{\mu}(\alpha(t)) = D\widetilde{\mu}(f_t, \zeta_t)\dot{\alpha}(t) \leq \sqrt{m}D^{3/2}\widetilde{\mu}(\alpha(t))^2\|(\dot{f}_t, \dot{\zeta}_t)\|,$$

and from (5.2) we conclude

$$\frac{d}{dt}\widetilde{\mu}(\alpha(t)) \leq \sqrt{m}D^{3/2}\widetilde{\mu}(\alpha(t))^2\sqrt{1+\widetilde{\mu}(\alpha(t))^2}\|\dot{f}_t\| \leq D^{3/2}\widetilde{\mu}(\alpha(t))^3\sqrt{\frac{m^2+1}{m}}.$$

The standard proof for Gronwall's Inequality can be repeated here:

$$D^{3/2}\sqrt{\frac{m^2+1}{m}} \geq \frac{1}{\widetilde{\mu}(\alpha(t))^3}\frac{d}{dt}\widetilde{\mu}(\alpha(t)) = -\frac{1}{2}\frac{d}{dt}\left(\frac{1}{\widetilde{\mu}(\alpha(t))^2}\right).$$

Hence,

$$\frac{d}{dt}\left(\frac{1}{\widetilde{\mu}(\alpha(t))^2}\right) \geq -2D^{3/2}\sqrt{\frac{m^2+1}{m}},$$

and

$$\frac{1}{\widetilde{\mu}(\alpha(t))^2} \geq \frac{1}{\widetilde{\mu}(f,\zeta)^2} - 2tD^{3/2}\sqrt{\frac{m^2+1}{m}}.$$

Equivalently,

$$\widetilde{\mu}(\alpha(t))^2 \leq \frac{\widetilde{\mu}(f,\zeta)^2}{1-2t\widetilde{\mu}(f,\zeta)^2D^{3/2}\sqrt{\frac{m^2+1}{m}}}.$$

In particular, as far as

$$t < \frac{1}{2\widetilde{\mu}(f,\zeta)^2D^{3/2}\sqrt{\frac{m^2+1}{m}}},$$

we have that $\alpha(t)$ can be continued and hence it is defined for every $t \in [0, d(f, h)]$ and the inequality of the lemma holds. □

6. Topology of $\mathcal{B}$ in the case that $k = m \leq n$

Theorem 3 describes the topology of W in quite a explicit way, as it suffices to study the topology of $\mathcal{B}$. The homotopy, homology and cohomology groups of W coincide with those of $\mathcal{B}$, that may be easier to compute. For example, if $\mathbb{K} = \mathbb{R}$ and $n = 1$ then $\mathcal{B}$ is connected and diffeomorphic to the unit circle $\mathbb{S}^1$. We compute some of these groups for the particular case that $k = m \leq n$. These results are summarized in the following table.

	$\mathbb{K}=\mathbb{R}$ $m=n>1$ n and $d_1+\cdots+d_n-1$ even	$\mathbb{K}=\mathbb{R}$ $m=n>1$ other cases	$\mathbb{K}=\mathbb{C}$ $m<n$	$\mathbb{K}=\mathbb{C}$ $m=n$
$\pi_0(\mathcal{B})$	$\{0,1\}$	$\{0\}$	$\{0\}$	$\{0\}$
$\pi_1(\mathcal{B})$	8 elements	4 elements	$\{0\}$	$\mathbb{Z}/a\mathbb{Z}$
$\pi_2(\mathcal{B})$	$\{0\}$	$\{0\}$	$\mathbb{Z}\oplus\mathbb{Z}$	$\mathbb{Z}$
$\pi_k(\mathcal{B}), k\geq 3$	$\pi_k(\mathcal{SO}_{n+1})$	$\pi_k(\mathcal{SO}_{n+1})$	?	$\pi_k(\mathcal{SU}_{n+1})$

where $a = \gcd(n, d_1 + \cdots + d_n - 1)$ and $\mathbb{Z}/a\mathbb{Z}$ is the finite cyclic group of a elements.

For other values of m, n, k, similar results may be achieved using the techniques of this section.

Remark 2. *The topological structure of W seems surprisingly simple. For example, the table implies that if $\mathbb{K}=\mathbb{C}$, $m=n=k$ and all the degrees $d_1,\dots,d_n$ are equal, then W is a simply connected manifold. The topology of V can also be studied from the fiber bundle given by $\pi_2 : V \to \mathbb{P}(\mathbb{K}^{n+1})$. The long exact sequence for a fibration* [Hat02, Th. 4.41] *yields for example that V is simply connected if $\mathbb{K}=\mathbb{C}$ and $\pi_1(V)$ has 4 elements if $\mathbb{K}=\mathbb{R}$, $n>1$.*

The cellular structure and homology theory of $\mathcal{B}$ should also be understandable. This is expressed for an example via Morse theory in the next proposition, as was pointed out by our colleagues Megumi Harada, Yael Karshon and Paul Selick.

Proposition 11. *Fix $\mathbb{K}=\mathbb{C}$ and $k=m=n$. The minimum number of critical points of a non-degenerate Morse function defined on $\mathcal{B}$ is at most $2^{n-1}(n+1)!$.*

We delay the proof of this proposition until we prove the next lemma and proposition.

From Proposition 2, $\mathcal{B}$ is the orbit of (g,e_0) under the action of $\mathcal{G}_{n+1}$, where g is the (normalized) homogenization of the identity, i.e. $g=(g_1,\dots,g_m)\in\mathbb{P}(\mathcal{H}_{(d)})$, $g_i = d_i^{1/2}X_0^{d_i-1}X_i$, $i=1\dots m$. Thus, $\mathcal{B}$ is diffeomorphic to $\mathcal{G}_{n+1}/H_1$ where

$$H_1 = \left\{ \begin{pmatrix} \rho & & \\ & Diag(\lambda\rho^{1-d_i}) & \\ & & Q \end{pmatrix}, \rho,\lambda\in\mathcal{G}_1, Q\in\mathcal{G}_{n-m} \right\}.$$

is the stabilizer of (g,e_0).

6.1. **Case** $\mathbb{K}=\mathbb{C}$. Note that the stabilizer H_1 is a connected set, and $\pi_1(\mathcal{U}_{n+1})=\mathbb{Z}$. The long exact sequence for the homotopy groups [Hat02, Th. 4.41] shows then that $\pi_1(\mathcal{B})$ is the homomorphic image of $\mathbb{Z}$ and hence it is a cyclic group.

Clearly, $\mathcal{B}$ is also the orbit of (g,e_0) under the action of the special unitary group $\mathcal{SU}_{n+1}$ on W. Let H_2 be the stabilizer of this action, so that $\mathcal{B}$ is diffeomorphic to $\frac{\mathcal{SU}_{n+1}}{H_2}$. Moreover, H_2 is the set of matrices

(6.1)
$$\begin{pmatrix} \rho & & \\ & Diag(\lambda\rho^{1-d_i}) & \\ & & Q \end{pmatrix}, \rho,\lambda\in\mathbb{S}^1, Q\in\mathcal{U}_{n-m}, \rho^{m+1-d_1-\cdots-d_m}\lambda^m\det(Q)=1.$$

Lemma 5. *Let $m<n$. Then, $\mathcal{B}$ is simply connected and $\pi_2(\mathcal{B})=\mathbb{Z}\oplus\mathbb{Z}$.*

Proof. Note that if $m<n$, we can define a surjection $H_2 \longrightarrow \mathbb{T}^2=\mathbb{S}^1\times\mathbb{S}^1$ by sending and element of H_2 in the form (6.1) to (ρ,λ). This defines on H_2 a structure of fiber bundle with fiber $\mathcal{SU}_{n-m}$. The long exact sequence for the homotopy groups reads

$$\pi_2(\mathcal{SU}_{n-m})\to\pi_2(H_2)\to\pi_2(\mathbb{T}^2)\to\pi_1(\mathcal{SU}_{n-m})\to\pi_1(H_2)\to\pi_1(\mathbb{T}^2)\to\{0\}.$$

That is,

$$\{0\}\to\pi_2(H_2)\to\{0\}\to\{0\}\to\pi_1(H_2)\to\mathbb{Z}\oplus\mathbb{Z}\to\{0\}.$$

Here we have used that $\pi_2(\mathcal{SU}_j)=\pi_1(\mathcal{SU}_j)=\{0\}$ for $j>0$ (cf. for example [Hus75, Ch. 7]). We conclude that $\pi_2(H_2)=\{0\}$ and $\pi_1(H_2)=\mathbb{Z}\oplus\mathbb{Z}$. So, the long exact sequence for $\mathcal{B}=\frac{\mathcal{SU}_{n+1}}{H_2}$,

$$\pi_2(H_2)\to\pi_2(\mathcal{SU}_{n+1})\to\pi_2(\mathcal{B})\to\pi_1(H_2)\to\pi_1(\mathcal{SU}_{n+1})\to\pi_1(\mathcal{B})\to\{0\},$$

reads

$$\{0\} \to \{0\} \to \pi_2(\mathcal{B}) \to \mathbb{Z} \oplus \mathbb{Z} \to \{0\} \to \pi_1(\mathcal{B}) \to \{0\},$$

and the lemma follows. □

Lemma 6. *Let $m = n$ and $a = gcd(n, d_1 + \cdots + d_n - 1)$. Then, there is a Lie isomorphism $H_2 \to G_a \times \mathbb{S}^1$, where $G_a \subseteq \mathbb{S}^1$ is the group of a-th roots of 1.*

Proof. Let $b = n+1-d_1-\cdots-d_n$, so that $a = gcd(n, b)$. Note that H_2 is isomorphic to $G_{b,n} = \{(\rho, \lambda) \in \mathbb{S}^1 \times \mathbb{S}^1 : \rho^b \lambda^n = 1\} \subseteq \mathbb{T}^2$. By considering the Lie epimorphism $G_{b,n} \longrightarrow G_a$, $(\rho, \lambda) \mapsto \rho^{b/a}\lambda^{n/a}$, it suffices to prove the lemma for $a = 1$. Now, in that case $G_{b,n}$ can be parametrized by $\{(t^n, t^{-b}) : t \in \mathbb{S}^1\}$ and is a torus knot that goes around $\mathbb{T}^2$, b times in one direction and n times in the other direction ([Mum76, Pg. 13-14]). This finishes the proof. □

Proposition 12. *Let $m = n$ and $a = gcd(n, d_1 + \cdots + d_n - 1)$. Then,*

$$\begin{array}{rcl} \pi_k(\mathcal{B}) & \equiv & \pi_k(\mathcal{SU}_{n+1}) \quad for\ \ k \geq 3 \\ \pi_2(\mathcal{B}) & \equiv & \mathbb{Z} \\ \pi_1(\mathcal{B}) & \equiv & \mathbb{Z}/a\mathbb{Z} \end{array}$$

where $\mathbb{Z}/a\mathbb{Z}$ is the finite cyclic group of order a.

Proof. The long exact sequence for the fibration $\mathcal{B} = \frac{\mathcal{SU}_{n+1}}{H_2}$ reads

$$\cdots \longrightarrow \pi_k(H_2) \longrightarrow \pi_k(\mathcal{SU}_{n+1}) \xrightarrow{\pi_*} \pi_k(\mathcal{B}) \longrightarrow \pi_{k-1}(H_2) \longrightarrow \cdots$$

Now, for $k \geq 2$ we have that $\pi_k(H_2) = \{0\}$ and we conclude that

$$\pi_k(\mathcal{B}) \equiv \pi_k(\mathcal{SU}_{n+1}), \quad \forall\ k \geq 3,$$

while the last terms of the long exact sequence above read

$$\{0\} \xrightarrow{\pi_*} \pi_2(\mathcal{B}) \longrightarrow \mathbb{Z} \longrightarrow \{0\} \xrightarrow{\pi_*} \pi_1(\mathcal{B}) \longrightarrow \pi_0(H_2) \longrightarrow \{0\}$$

We conclude that $\pi_2(\mathcal{B}) \equiv \mathbb{Z}$ and $\pi_1(\mathcal{B})$ has finite cardinal a. As stated as the beginning of this section, $\pi_1(\mathcal{B})$ is cyclic so it equals $\mathbb{Z}/a\mathbb{Z}$. □

Proof of Proposition 11. Consider the fiber bundle $\mathbb{T}^n/H_2 \to \mathcal{B} \xrightarrow{p} \mathcal{SU}_{n+1}/\mathbb{T}^n$ where $\mathbb{T}^n$ is the maximal torus in $\mathcal{SU}_{n+1}$. $\mathcal{SU}_{n+1}/\mathbb{T}^n$ is a flag manifold and has a perfect Morse function with $(n+1)!$ critical points. Composing with p we have a Morse-Bott function with $(n+1)!$ critical tori. Perturbing, we get a Morse function with $2^{n-1}(n+1)!$ non-degenerate critical points.

6.2. **Case $\mathbb{K} = \mathbb{R}$.** Now we introduce the analogue lemmas on the topology of $\mathcal{B}$ for the real case, and we include a proof when necessary. The stabilizer H_1 of (g, e_0) under the action of the orthogonal group on W is now the set of matrices

$$\begin{pmatrix} \rho & & \\ & Diag(\rho^{1-d_i})\lambda & \\ & & Q, \end{pmatrix}$$

where $\rho, \lambda \in \mathbb{S}^0 = \{-1, 1\}$ and $Q \in \mathcal{O}_{n-m}$. In particular, H_1 is diffeomorphic to $\mathbb{S}^0 \times \mathbb{S}^0 \times \mathcal{O}_{n-m}$.

Lemma 7. *Let $H_2 \subseteq \mathcal{SO}_{n+1}$ be the stabilizer of (g, e_0) under the action of $\mathcal{SO}_{n+1}$ on W. If $m < n$ or one of $n, d_1 + \cdots + d_n - 1$ is odd, then $\mathcal{B}$ equals the orbit of (g, e_0) under the action of $\mathcal{SO}_{n+1}$ on W and hence $\mathcal{B}$ is diffeomorphic to $\mathcal{SO}_{n+1}/H_2$. If $m = n$ and both $n, d_1 + \cdots + d_n - 1$ are even, $\mathcal{B}$ has two diffeomorphic connected components, and each of them is diffeomorphic to $\mathcal{SO}_{n+1}/H_2$.*

Proof. Consider the action of the special orthogonal group $\mathcal{SO}_{n+1}$ on W. Let $(f, \zeta) = (g \circ U^*, Ue_0)$ where $U \in \mathcal{O}_{n+1}$ and let $V = UDiag(\rho, \rho^{1-d_1}\lambda, \ldots, \rho^{1-d_m}\lambda, Q)$ where $\rho, \lambda \in \mathbb{S}^0$ and $Q \in \mathcal{O}_{n-m}$ are generic. Note that $(g \circ V^*, Ve_0) = (f, \zeta)$, and $V \in \mathcal{SO}_{n+1}$ when $\rho^{1+m-d_1-\cdots-d_m}\lambda^m \det(Q) = \frac{1}{\det(U)}$. Hence, if $m < n$ or one of $n, d_1 + \cdots + d_n - 1$ is odd, we have that $\mathcal{B}$ equals the orbit of (g, e_0) under the action of $\mathcal{SO}_{n+1}$ on W. Otherwise, $\mathcal{B}$ is diffeomorphic to $\mathcal{O}_{n+1}/H_1$, where H_1 is the discrete subgroup of matrices of the form $Diag(\rho, \rho^{1-d_1}\lambda, \ldots, \rho^{1-d_n}\lambda)$. Now, $\mathcal{O}_{n+1}/H_1$ has two diffeomorphic connected components as $H_1 \subseteq \mathcal{SO}_{n+1}$. This finishes the proof. □

Recall (cf. for example [Hus75, Ch. 7]) that

$$\pi_0(\mathcal{SO}_r) = \{0\}, \;\; \pi_2(\mathcal{SO}_r) = 0, \forall\, r \geq 1,$$

$$\pi_1(\mathcal{SO}_1) = \{0\}, \;\; \pi_1(\mathcal{SO}_2) = \mathbb{Z}, \;\; \pi_1(\mathcal{SO}_r) = \mathbb{Z}_2, \forall\, r \geq 3.$$

and note that if $m = n$,

- $\sharp\pi_0(H_2) = 2$ except if both $n, d_1 + \cdots + d_n - 1$ are even, and
- in that case, $\sharp\pi_0(H_2) = 4$.

The long exact sequence for the homotopy yields some information about the homotopy groups of $\mathcal{B}$ in the different cases:

- Let $m = n = 1$. Then, $\mathcal{B}$ is diffeomorphic to a $\mathbb{S}^1$.
- Let $m = n > 1$. Then, $\pi_k(\mathcal{B}) = \pi_k(\mathcal{SO}_{n+1})$, for every $k \geq 2$, and $\sharp(\pi_1(\mathcal{B})) = 2\sharp(\pi_0(H_2))$.
- Let $n - m \geq 3$. Then, $\pi_2(\mathcal{B})$ has 1 or 2 elements, and $\pi_1(\mathcal{B})$ is not trivial and has at most 16 elements.

7. Appendix: Proof of Theorem 3

The proof of Theorem 3 uses Dynamical Systems Theory, more especifically theory of invariant manifolds. The reader may find background in [HPS77].

Let M be a finite dimensional manifold and $\mathcal{B} \subseteq M$ be a compact submanifold of M. Let TM be the tangent bundle to M and $T_{\mathcal{B}}M$ the restriction of TM to $\mathcal{B}$. Let $X : M \longrightarrow TM$ be a smooth vector field which vanishes on $\mathcal{B}$. The derivative of X at $b \in \mathcal{B}$ is then a well-defined map $T_bX : T_bM \longrightarrow T_bM$ that can be computed in any coordinate system. We write $T_{\mathcal{B}}X : T_{\mathcal{B}}M \longrightarrow T_{\mathcal{B}}M$ for the global version.

For all $b \in \mathcal{B}$ and for every eigenvector $v \in T_bW \setminus T_b\mathcal{B}$, we now assume that the real part of the corresponding eigenvalue is not zero. Let $N^s \subseteq T_{\mathcal{B}}M$ (resp. $N^u \subseteq T_{\mathcal{B}}M$) be the vector subbundle such that N^s_b is the generalized eigenspace of the eigenvalues of T_bX with negative (resp. positive) real part. Using local coordinates, it is easy to see that N^s, N^u and $N = N^s \oplus N^u$ are smooth vector subbundles of $T_{\mathcal{B}}M$, invariant under $T_{\mathcal{B}}X$. We write

$$T^sX = T_{\mathcal{B}}X\,|_{N^s} : N^s \longrightarrow N^s,$$

$$T^uX = T_{\mathcal{B}}X\,|_{N^u} : N^u \longrightarrow N^u.$$

We define a flow $\psi_t : N \longrightarrow N$ by $\psi_t(b, v) = (b, \exp(tT_bX)(v))$ where $\exp(\cdot)$ is the usual exponential of a linear map. Then, ψ_t leaves invariant N^u and N^s and we can choose metrics so that $\psi_t^s = \psi_t \mid_{N^s}$ contracts vector lengths and $\psi_t^u = \psi_t \mid_{N^u}$ expands vector lengths (note that this choice of metric may disagree with any previously imposed Riemannian structure on M. In particular, if X is defined as a gradient flow, the new metric might be different from the one defining X).

Recall that ϕ_t is the solution flow of the vector field $V(x) = -\mathrm{grad}\widetilde{\mu}(x)$ on W. For simplicity's sake we assume that ϕ_t is defined for all t. The manifold $\mathcal{B}$ is normally hyperbolic (see [HPS77, Sect. 1]) for the flow ϕ_t. Here we restrict the application of the theory of normally hyperbolic manifolds to the case that the manifold $\mathcal{B}$ consists of fixed points. Much of what we say applies with greater generality.

We consider the stable and unstable manifolds of $\mathcal{B}$, $W^s(\mathcal{B}) = \{x \in M : \phi_t(x) \mapsto \mathcal{B}, t \mapsto \infty\}$ and $W^u(\mathcal{B}) = \{x \in M : \phi_t(x) \mapsto \mathcal{B}, t \mapsto -\infty\}$. Recall that for $b \in \mathcal{B}$ we also have $W^s(b) = \{x \in M : \phi_t(x) \mapsto b, t \mapsto \infty\}$ and $W^u(b) = \{x \in M : \phi_t(x) \mapsto b, t \mapsto -\infty\}$. Finally, two flows ϕ_t^1 and ϕ_t^2 are (locally) topologically conjugate if there is a (local) homeomorphism h such that $h\phi_t^1 = \phi_t^2 h$ on the domain of definition of h and when defined.

We state the following result for stable manifolds, being identical for unstable manifolds (changing s to u). Note that we are in the situation of [HPS77, Sect. 1], with the extra assumption that the invariant manifold $\mathcal{B}$ consists of fixed points of the flow. This extra assumption will imply smoothness (instead of the usual continuity) of the foliation W^s.

Theorem 4. *In the conditions above,*

1. *$W^s(\mathcal{B})$ is a $\mathcal{C}^\infty$ manifold tangent to N^s at $\mathcal{B}$.*
2. *$W^s(\mathcal{B})$ is $\mathcal{C}^\infty$ foliated by the $W^s(b)$ ($b \in \mathcal{B}$), which are tangent to N_b^s at $b \in \mathcal{B}$.*
3. *N^s is $\mathcal{C}^\infty$ diffeomorphic to $W^s(\mathcal{B})$ by a diffeomorphism σ which is the identity on $\mathcal{B}$ and which maps the fibration of N^s as a vector bundle over $\mathcal{B}$ to the foliation of $W^s(\mathcal{B})$ by the $W^s(b)$ leaves. The derivative $D\sigma \mid_{N_b^s}$ is the identity for all $b \in \mathcal{B}$.*
4. *ψ_t^s is topologically conjugate to $\phi_t \mid_{W^s(\mathcal{B})}$ by a $\mathcal{C}^0$ conjugacy h^s which is the identity on $\mathcal{B}$.*
5. *ϕ_t and ψ_t are topologically conjugate on a neighborhood of $\mathcal{B}$.*

Proof. (1) See [HPS77, Th. 4.1].

(2) By the overflowing property on $W^s(\mathcal{B})$ locally near $\mathcal{B}$ and the fact that $\phi_t \mid_{\mathcal{B}} = Id_{\mathcal{B}}$, by the $\mathcal{C}^r$-section theorem [Shu87, Th. 5.18] applied to the graph transform whose fixed section is the tangent bundle to the manifolds $W^s(b)$, for every $r > 0$ there is a neighborhood U_r^s of $\mathcal{B}$ such that the $W^s(b)$ foliation of $W^s(\mathcal{B})$ is $\mathcal{C}^r$ near $\mathcal{B}$. Saturating by the flow shows that the global foliations are $\mathcal{C}^r$ for all $r > 0$ hence $\mathcal{C}^\infty$.

(3) To see that σ exists locally near $\mathcal{B}$, consider the manifold $\{(x, b) : x \in W^s(b)\} \subseteq W^s(\mathcal{B}) \times \mathcal{B}$ which is the graph of a smooth function and hence smooth. Then consider the smooth mapping $\varphi : x \mapsto (x, b) \mapsto \pi_{N_b^s} exp_b^{-1}(x)$ with exp the exponential map in M and $\pi_{N_b^s}$ the orthogonal projection. Clearly, the derivative of φ at b is non-singular and hence, locally near every $b \in \mathcal{B}$, φ is a smooth diffeomorphism. By compactness of $\mathcal{B}$, we conclude that φ is a smooth diffeomorphism locally near $\mathcal{B}$. Take $\sigma = \varphi^{-1}$.

Now we globalize σ. Put a metric on N^s so that the orbits of ψ_t^s are transversal to $\mathbb{S}_b^s(r)$ for all $r < r_0$, where $\mathbb{S}_b^s(r)$ is the sphere of radius r in N_b^s, and the same is true for $\sigma^{-1}(\phi_t)$. Here, r_0 is small enough such that σ is defined. Now choose $0 < a_1 < a_2 < r_0$ and alter the vector field defining the ψ_t^s flow to a $\mathcal{C}^\infty$ vector field such that it remains transversal to the spheres for $r < r_0$, and the new flow $\tilde{\psi}_t^s$ and ϕ_t are σ-conjugate on the set $V_{a_1}^{a_2} = \{(b,v) \in N^s : a_1 < \|v\| < a_2\}$. Finally define

$$\begin{array}{llll} \tilde{\sigma}: & N^s & \longrightarrow & W^s(\mathcal{B}) \\ & (b,v) & \mapsto & \sigma(b,v) \ \ if \ \ \|v\| < a_2, \\ & (b,v) & \mapsto & \phi_{-t}\sigma\tilde{\psi}_t(b,v) \ \ if \ \ \|v\| \geq a_2, \end{array}$$

where t in this last formula is any positive number such that $\tilde{\psi}_t(v) \in V_{a_1}^{a_2}$. Note that as $\tilde{\psi}_t$ and ϕ_t are σ-conjugate on $\sigma(V_{a_1}^{a_2})$, $\tilde{\sigma}$ is well defined, smooth and indeed a diffeomorphism. Moreover, $\tilde{\sigma} = \sigma$ on a neighborhood of $\mathcal{B}$ and satisfies the claimed properties.

(4) The local version is proven in [PS70]. To get the global version from a local conjugacy h_{loc} define $h(x) = \psi_t h_{loc} \phi_t^{-1}(x)$ for any t such that $\phi_t^{-1}(x)$ is in the domain of definition of h_{loc}.

(5) Same as (4).

□

Remark 3. *By a result of Hartman* [Har60] *(see also* [McS96]*), for any fixed* $b \in \mathcal{B}$, $\phi_t \mid_{W^s(b)}$ *is* $\mathcal{C}^1$ *conjugate to* $\psi_t^s \mid_{N_b^s}$. *We don't know how such a* $\mathcal{C}^1$ *conjugacy may be made global so as to be* $\mathcal{C}^1$ *in* b. *In the context of equivariant gradient flow as we are considering, the proof in* [McS96] *might be adaptable.*

7.1. **Proof of Theorem 3.** Note that $\widetilde{\mu}$ is proper and bounded from below. Thus $W^s(\mathcal{B}) = W$. The derivative at $\mathcal{B}$ satisfies $-D\text{grad}\widetilde{\mu} = -D^2\widetilde{\mu}$, which from Proposition 7 is symmetric negative definite on the normal bundle to $\mathcal{B}$ and its eigenvalues are pointwise real and negative. Hence we can apply Theorem 4, from which Theorem 3 follows.

References

[AVGZ86] V. Arnold, A. Varchenko, and S. Goussein-Zadé, *Singularités des aplications différentiables*, Éditions Mir, Moscou, 1986.

[BCSS98] L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and real computation*, Springer-Verlag, New York, 1998.

[BP06] C. Beltrán and L.M. Pardo, *On Smale's 17th problem: A probabilistic positive answer.*, Found. Comput. Math. **Online First DOI 10.1007/s10208-005-0211-0** (2006).

[BP07] ———, *Smale's 17th problem: Average polynomial time to compute affine and projective solutions.*, preprint (2007).

[BS07] C. Beltrán and M. Shub, *Complexity of Bezout's Theorem VII: Distance estimates in the condition metric*, Found. Comput. Math. **DOI 10.1007/s10208-007-9018-5** (2007).

[Dem87] J. W. Demmel, *The geometry of ill-conditioning*, J. Complexity **3** (1987), no. 2, 201–229.

[EY36] C. Eckart and G. Young, *The approximation of one matrix by another of lower rank*, Psychometrika **1** (1936), 211–218.

[Har60] P. Hartman, *On local homeomorphisms of Euclidean spaces*, Bol. Soc. Mat. Mexicana (2) **5** (1960), 220–241.

[Hat02] Allen Hatcher, *Algebraic topology*, Cambridge University Press, Cambridge, 2002. MR MR1867354 (2002k:55001)

[HPS77] M. W. Hirsch, C. C. Pugh, and M. Shub, *Invariant manifolds*, Springer-Verlag, Berlin, 1977, Lecture Notes in Mathematics, Vol. 583.

[Hus75] Dale Husemoller, *Fibre bundles*, second ed., Springer-Verlag, New York, 1975, Graduate Texts in Mathematics, No. 20. MR MR0370578 (51 #6805)

[McS96] P. D. McSwiggen, *A geometric characterization of smooth linearizability*, Michigan Math. J. **43** (1996), no. 2, 321–335. MR MR1398157 (97i:58128)

[Mum76] D. Mumford, *Algebraic geometry. I*, Springer-Verlag, Berlin, 1976, Complex projective varieties, Grundlehren der Mathematischen Wissenschaften, No. 221.

[PS70] Charles Pugh and Michael Shub, *Linearization of normally hyperbolic diffeomorphisms and flows*, Invent. Math. **10** (1970), 187–198.

[Shu87] Michael Shub, *Global stability of dynamical systems*, Springer-Verlag, New York, 1987, With the collaboration of Albert Fathi and Rémi Langevin, Translated from the French by Joseph Christy. MR MR869255 (87m:58086)

[Shu07] M. Shub, *Complexity of Bézout's theorem. VI: Geodesics in the condition (number) metric*, J. Foundations of Computational Mathematics **DOI 10.1007/s10208-007-9017-6** (2007).

[SS93] M. Shub and S. Smale, *Complexity of Bézout's theorem. I. Geometric aspects*, J. Amer. Math. Soc. **6** (1993), no. 2, 459–501.

[SS96] ———, *Complexity of Bezout's theorem. IV. Probability of success; extensions*, SIAM J. Numer. Anal. **33** (1996), no. 1, 128–148.

[Was69] Arthur G. Wasserman, *Equivariant differential topology*, Topology **8** (1969), 127–150. MR MR0250324 (40 #3563)

(C. Beltrán and M. Shub) Department of Mathematics, University of Toronto, Toronto, Ontario, Canada M5S 2E4

E-mail address, C. Beltrán: `beltranc@math.toronto.edu`

E-mail address, M. Shub: `michael.shub@utoronto.ca`