Conditioned invariant subspaces and the geometry of nilpotent matrices

Uwe Helmke[1] and Jochen Trumpf[2]

[1] Mathematisches Institut
Universität Würzburg
Am Hubland, 97074 Würzburg, Germany
helmke@mathematik.uni-wuerzburg.de

[2] Department of Information Engineering
The Australian National University
Canberra ACT 0200, Australia
and
National ICT Australia Ltd.**
Jochen.Trumpf@anu.edu.au

Dedicated to Clyde F. Martin, on occasion of his 60th birthday

1 Introduction

Invariant subspaces of linear operators have long played an important role in pure and applied mathematics, including areas such as e.g. operator theory and linear algebra [7],[8],[11], as well as algebraic groups, representation theory and singularity theory [4],[25],[27],[28]. Their role in control theory in connection with matrix Riccati equations and linear optimal control is now well-understood and has been the subject of extensive research during the past decades; see e.g. [22],[16]. We also mention the important connection to geometric control theory and the fundamental concept of conditioned and controlled invariant subspaces [31]. In fact, the so-called "quaker lemma" completely characterizes conditioned invariant subspaces solely in terms of invariant subspaces through the concept of a friend.

The focus of this work is on certain geometric aspects of the classification problems for invariant and conditioned invariant subspaces. The investigation of the geometry of the algebraic variety $\mathrm{Inv}_k(A)$ of k-dimensional A-invariant subspaces of a vector space V goes back to the early work of Steinberg [27].

**National ICT Australia Ltd. is funded by the Australian Government's Department of Communications, Information Technology and the Arts and the Australian Research Council through Backing Australia's Ability and the ICT Centre of Excellence Program.

Motivated by applications to singularity theory, Steinberg raised the problem of analyzing the geometry of $\mathrm{Inv}_k(A)$ and derived important results. He, as well as Springer [25], showed that the geometry of $\mathrm{Inv}_k(A)$ could be used to construct resolutions of singularities for the set of nilpotent matrices. It also led to a new geometric construction of Weyl group representations [26]. Subsequent work by Spaltenstein [24] and others established basic geometric facts about $\mathrm{Inv}_k(A)$, such as the enumeration of irreducible components via Young diagrams, or the computation of topological invariants, such as Borel-Moore homology groups and intersection homologies [3],[13].

Control theory has provided a different and new entry point to this circle of ideas, as was first realized by Shayman, and Hermann and Martin [22],[16],[18]. In fact, the projective variety $\mathrm{Inv}_k(A)$ can be interpeted as a compactification for the solution set of the algebraic Riccati equation and this link deepened considerably the further understanding of the Riccati equation. In [21],[22], Shayman studied the geometric properties of the solution set of the algebraic Riccati equations, by connecting it to the geometry of $\mathrm{Inv}_k(A)$ and the Grassmann manifold. Interesting applications of the Grassmannian approach to numerical linear algebra appeared in [1] and a whole circle of ideas, centering around nilpotent matrices, representations of the symmetric group, Schubert cycles and the classification of state feedback orbits, has been masterfully presented in [12]. Already in the late 70s, the link between invariant subspaces and geometric control objects, such as conditioned invariant subspaces, was well understood. A driving force for their analysis has been their ubiquitous role in e.g. spectral factorization, linear quadratic control, H_∞ and game theory, as well as observer theory, filtering and estimation. However, it is only until recently, that first attempts have been made towards a better understanding of the geometry of the set of conditioned invariant subspaces $\mathrm{Inv}_k(C, A)$; see [9],[10],[19],[6],[17]. The recent Ph.D. thesis [29] contains a comprehensive summary. Nevertheless, it is fair to say that our current knowledge on $\mathrm{Inv}_k(C, A)$ remains limited, with several basic questions unsolved. For example, it is not known, whether or not $\mathrm{Inv}_k(C, A)$ is homotopy equivalent to $\mathrm{Inv}_k(A)$, or if $\mathrm{Inv}_k(C, A)$ is a manifold. Generally speaking, the interactions between linear systems theory and geometry or physics, despite first steps [14],[23], have not been explored to the depth that they deserve and remain a challenging task for future research. We are convinced, that conditioned invariant subspaces are bound to play an important role here.

In this paper, we make an attempt to illustrate the interplay between geometry and control, by focussing on the connections between partial state observers, spaces of invariant and conditioned invariant subspaces, and nilpotent matrices. The two crucial players in our story are on the one hand the set of pairs (A, V) of linear operators and invariant flags, and on the other hand the set of pairs (J, V), of friends J and conditioned invariant flags V, for a given observable pair (C,A) in dual Brunovsky form. We prove that, despite the actual and potential singularities of $\mathrm{Inv}_k(A)$ and $\mathrm{Inv}_k(C, A)$, respectively, these sets of pairs are actually smooth manifolds, being closely related to classical geometric objects such as the cotangent bundle of the flag manifold. After having introduced these spaces and established their manifold property, we then consider desingularizations of the set of nilpotent matrices. Here we make contact with symplectic geometry and the moment map. State feedback transformations enable us to construct suitable transversal slices to nilpotent similarity orbits. Our construction of these Brunovsky slices extends that of Steinberg [28]. The intersections of the similarity orbits with the variety of nilpotent orbits exhibit interesting singularities, including Kleinian singularities of complex surfaces. We show, at least in a generic situation, that the proposed desingularization via conditioned invariant subspaces restricts to a simultaneous desingularization for all the intersection varieties. We regard this as one of our most surprising results, i.e., that conditioned invariant subspaces provide a natural desingularization for Kleinian singularities. Unfortunately, we cannot really give a deeper explanation of this fact, as this would require to extend the parametrization of conditioned invariant subspaces to systems defined on arbitrary semi-simple Lie groups. The link between linear systems and geometry is also visible at the proof level. In fact, in order to construct a miniversal deformation of the variety of nilpotent matrices via Brunovsky slices, we have to extend the well–known fact (commonly referred to as the *Hermann-Martin Lemma*) that state feedback transformations define a submersion on the state feedback group.

From what has been said above, it becomes evident that the pioneering contributions of Clyde F. Martin to systems theory had a great impact on a wide range of research areas. The influential character of his work on other researchers is also visible in this paper at crucial steps. In fact, Clyde has always been a source of inspiration, and a friend. It is a great pleasure to dedicate this paper to him.

2 Smoothness criteria for vector bundles

A well-known consequence of the implicit function theorem is that the fibres $f^{-1}(y)$ of a smooth map $f : M \longrightarrow N$ are smooth manifolds, provided the rank of the differential $Df(x)$ is constant on a neighborhood of $f^{-1}(y)$. The situation becomes more subtle if the rank of $Df(x)$ is assumed to be constant only *on* the fibre, but is allowed to vary outside of it. Then we cannot conclude that $f^{-1}(y)$ is smooth and more refined techniques than the implicit function theorem are needed for. In fact, this is exactly the situation that arises when one tries to prove that certain families of conditioned invariant subspaces are smooth manifolds. One technique that can be applied in such a situation where the implicit function theorem fails is by realizing the space as an abstract vector bundle X over a base manifold, suitably embedded into a smooth vector bundle. Then one can try to employ group action arguments to prove that X is indeed a smooth vector bundle. This therefore requires an easily applicable criterion, when a vector subbundle qualifies as a smooth vector bundle. In this section, we will derive a sufficient condition for a (topological) vector bundle to be a smooth vector bundle. The result also provides a sufficient condition in certain situations for the pre-image of a smooth submanifold being a smooth submanifold again. Second, quotients of smooth vector bundles with respect to free and proper Lie group actions are shown to be smooth vector bundles. Finally, we give a sufficient criterion for a projection map from a subset of a product manifold onto one of the factors being a smooth vector bundle. The criterion involves a Lie group that operates on both factors.

We begin with some standard terminology and notations. Throughout the paper let $\mathbb{F}$ be the field $\mathbb{R}$ or $\mathbb{C}$. Recall, that a *Lie group action* of a Lie group G on a manifold M is a smooth map

$$\sigma : G \times M \longrightarrow M,\ (g,x) \mapsto \sigma(g,x) = g \cdot x$$

such that $e \cdot x = x, g \cdot (h \cdot x) = gh \cdot x$ holds for all $g, h \in G, x \in M$. The *orbit space* of σ is defined as the quotient space $M/\sim_\sigma$ for the associated equivalence relation on M, where

$$m \sim_\sigma m' \text{ if and only if there is } g \in G \text{ with } m' = g \cdot m.$$

Thus the points of M/G are the equivalence classes

$$[m]_{\sim_\sigma} := \{m' \mid m \sim_\sigma m'\}.$$

The orbit space M/G carries a canonical topology, the quotient topology, which is defined as the finest topology for which the *natural projection*

$$\pi : M \longrightarrow M/G,\ m \mapsto [m]_{\sim_\sigma}$$

is continous. In order to study geometric properties of the orbit space M/G one has to consider the *graph map* $\tilde{\sigma}$ of the action. This is defined as

$$\tilde{\sigma} : G \times M \longrightarrow M \times M,\ (g,x) \mapsto (x, g \cdot x).$$

Therefore the image of $\tilde{\sigma}$ is nothing else but the relation $\sim_\sigma$ seen as a subspace of $M \times M$. Under certain circumstances the orbit space M/G is a manifold again. The following necessary and sufficient condition can be found in [5, Theorem 16.10.3].

Proposition 1. *There is a unique manifold structure on M/G such that the natural projection π is a submersion if and only if the image of the graph map $\operatorname{Im}\tilde{\sigma}$ is a closed submanifold of $M \times M$.*

Recall that a group action σ is called *free* if the stabilizer subgroups

$$G_x := \{g \in G \mid g \cdot x = x\}$$

are trivial, i.e., $G_x = \{e\}$ for all $x \in M$. Moreover, the action is called *proper* if the graph map is proper, i.e., the inverse image $\tilde{\sigma}^{-1}(K)$ of any compact subset $K \subset M \times M$ is a compact subset of M. Suppose, σ is a free Lie group action of G on M. Then it is easily seen that $\operatorname{Im}\tilde{\sigma}$ is a closed submanifold of $M \times M$ if and only if $\tilde{\sigma}$ is a closed map, i.e., maps closed sets to closed sets. Thus using the above quotient manifold criterion we arrive at the following well-known manifold criterion for an orbit space.

Theorem 1. *Let σ be a free and proper Lie group action of G on M. Then the orbit space M/G is a smooth manifold of dimension $dim M - dim G$. Moreover, the quotient map $\pi : M \longrightarrow M/G$ is a principal fibre bundle with structure group G.*

After these basic facts we turn to the discussion of vector bundles. The following definition is standard, but repeated here for convenience. Let X and B be Hausdorff spaces and let

$$f: \ X \longrightarrow B$$

be a continuous surjection. Let $p \in \mathbb{N}$ be fixed. For each point $x \in B$, assume there exists an open neighborhood U and a homeomorphism

$$\phi_U: \ U \times \mathbb{F}^p \longrightarrow f^{-1}(U)$$

such that

$$f(\phi_U(x,y)) = x$$

holds for all $x \in U$ and all $y \in \mathbb{F}^p$. Then ϕ_U is called a *local trivialization* of f. For each pair ϕ_U and ϕ_V of local trivializations and each point $x \in U \cap V$ suppose in addition that there exists a map $\theta_{V,U,x} \in \mathrm{GL}_p(\mathbb{F})$ such that

$$\phi_V^{-1} \circ \phi_U(x,y) = (x, \theta_{V,U,x}(y))$$

for all $y \in \mathbb{F}^p$, i.e., the induced change of coordinates function on $\mathbb{F}^p$ is linear. Then f is called a *topological vector bundle* with *fiber* $\mathbb{F}^p$. If X and B are smooth manifolds, f is a smooth map, and each ϕ_U is a diffeomorphism, then the bundle is called *smooth.*

The following result constitutes our proposed extension of the implicit function theorem for fibres that are given by topological vector bundles.

Theorem 2. *Let N and M be smooth manifolds, let $X \subset N$ be a topological subspace, let $B \subset M$ be a q-dimensional smooth submanifold and let $f : X \longrightarrow B$ be a topological vector bundle with fiber $\mathbb{F}^p$ such that f is the restriction of a smooth map*

$$F: \ U_X \longrightarrow M$$

on an open neighborhood U_X of X in N. Suppose, that each local trivialization $\phi_U : U \times \mathbb{F}^p \longrightarrow f^{-1}(U)$ of f is smooth as a map into N and such that $\phi_U^{-1} : f^{-1}(U) \longrightarrow U \times \mathbb{F}^p$ is the restriction of a smooth map

$$\Phi_{U,\mathrm{inv}}: \ U_{f^{-1}(U)} \longrightarrow M \times \mathbb{F}^p,$$

where $U_{f^{-1}(U)}$ is an open neighborhood of $f^{-1}(U)$ in N. Then X is a $(q+p)$-dimensional smooth submanifold of N and f is a smooth vector bundle.

Proof. Given $x_0 \in X$, there exists an open neighborhood U_1 of $f(x_0)$ in B and a local trivialization $\phi_{U_1} : U_1 \times \mathbb{F}^p \longrightarrow f^{-1}(U_1)$ of f. Furthermore, there exists an open neighborhood U_2 of $f(x_0)$ in B and a local coordinate chart $\varphi_{U_2} : U_2 \longrightarrow \varphi_{U_2}(U_2) \subset \mathbb{F}^q$ of B around $f(x_0)$. Set $U := U_1 \cap U_2$. Then U is open in B and $\phi_U := \phi_{U_1}|_{U\times\mathbb{F}^p} : U \times \mathbb{F}^p \longrightarrow f^{-1}(U)$ is a local trivialization of f. Furthermore, $\varphi_U := \varphi_{U_2}|_U : U \longrightarrow \varphi_U(U) = \varphi_{U_2}(U)$ is a local coordinate chart of B around $f(x_0)$. Define a local coordinate chart $\psi_{f^{-1}(U)} : f^{-1}(U) \longrightarrow \varphi_U(U) \times \mathbb{F}^p$ of X around x_0 by

$$x \mapsto (\varphi_U \circ \mathrm{pr}_1 \circ \phi_U^{-1}(x), \mathrm{pr}_2 \circ \phi_U^{-1}(x)).$$

Here pr_1 and pr_2 denote the projections on the first and second factor of $U \times \mathbb{F}^p$, respectively. Note that $f^{-1}(U)$ is open in X, since f is continuous. Note further that $\varphi_U(U) \times \mathbb{F}^p$ is open in $\mathbb{F}^q \times \mathbb{F}^p$, since φ_U is a local coordinate chart of B. Since ϕ_U^{-1} and φ_U are both bijective, so is $\psi_{f^{-1}(U)}$. Furthermore, $\psi_{f^{-1}(U)}$ is continuous as a concatenation of continuous maps, and hence it is a homeomorphism.

Now let $\psi_{f^{-1}(U)}$ and $\psi_{f^{-1}(V)}$ be two such local coordinate charts of X with $f^{-1}(U) \cap f^{-1}(V) \neq \emptyset$. Then

$$\psi_{f^{-1}(V)} \circ \psi_{f^{-1}(U)}^{-1} : \; \varphi_U(U \cap V) \times \mathbb{F}^p \longrightarrow \varphi_V(U \cap V) \times \mathbb{F}^p,$$

which is given by

$$\begin{aligned}(z, y) \;\mapsto\; & (\varphi_V \circ \mathrm{pr}_1 \circ \phi_V^{-1} \circ \phi_U(\varphi_U^{-1}(z), y), \mathrm{pr}_2 \circ \phi_V^{-1} \circ \phi_U(\varphi_U^{-1}(z), y)) \;= \\ & (\varphi_V \circ \varphi_U^{-1}(z), \theta_{V,U,\varphi_U^{-1}(z)}(y)),\end{aligned}$$

is a diffeomorphism, since $\varphi_V \circ \varphi_U^{-1}$ is a diffeomorphism and $\theta_{V,U,\varphi_U^{-1}(z)} \in \mathrm{GL}_p(\mathbb{F})$. It follows that X is a $(p+q)$-dimensional smooth manifold. Since the local coordinate charts of X are continuous, the preimage of any open set in $\mathbb{F}^q \times \mathbb{F}^p$ under any chart is open in X. These preimages form a basis of the topology τ induced on X by its differentiable structure, and thus τ coincides with the subspace topology on X, that is induced by the topology on N. Thus X is a submanifold of N.

Since f and the inverse maps ϕ_U^{-1} of all local trivializations ϕ_U of f are restrictions of smooth maps defined on open subsets of N, they are smooth themselves. Since each ϕ_U is also smooth, it follows that f is a smooth vector bundle. □

The following theorem shows that quotients of smooth vector bundles with respect to free and proper Lie group actions are again smooth vector bundles provided the natural compatibility condition (1) holds.

Theorem 3. *Let $f : X \longrightarrow B$ be a smooth vector bundle with fiber $\mathbb{F}^p$ and let*

$$\sigma_X : \; G \times X \longrightarrow X$$

and

$$\sigma_B : \ G \times B \longrightarrow B$$

be free and proper actions of the Lie group G on X and B, respectively. For every local trivialization ϕ_U of f, let U consist of full G-orbits (i.e., $x \in U$ implies $\sigma_B(g,x) \in U$ for all $g \in G$) and let

$$\phi_U(\sigma_B(g,x),y) = \sigma_X(g,\phi_U(x,y)) \tag{1}$$

for all $g \in G$, $x \in U$ and $y \in \mathbb{F}^p$. Then

$$\begin{aligned} \bar{f} : X/G &\longrightarrow B/G, \\ [x]_{\sim_{\sigma_X}} &\mapsto [f(x)]_{\sim_{\sigma_B}} \end{aligned}$$

is a smooth vector bundle with fiber $\mathbb{F}^p$.

Proof. Let $x \in X$ and $g \in G$ be arbitrary. Then $f(x) \in B$ and hence there exists a neighborhood U of $f(x)$ and a local trivialization ϕ_U such that $x = \phi_U(z,y)$ for appropriate $z \in U$ and $y \in \mathbb{F}^p$. But then (1) implies $f(\sigma_X(g,x)) = f(\sigma_X(g,\phi_U(z,y))) = f(\phi_U(\sigma_B(g,z),y)) = \sigma_B(g,z)$. Taking $g = e$ yields $f(x) = z$. But this means

$$f \circ \sigma_X(g,x) = \sigma_B(g,f(x)) \tag{2}$$

for all $g \in G$ and $x \in X$.
From (2) it follows that $\bar{f}$ is well defined. By Theorem 1, the spaces X/G and B/G are both smooth manifolds. Now consider the following commutative diagram:

$$\begin{array}{ccc} X & \xrightarrow{\ f\ } & B \\ {\scriptstyle \pi_X}\downarrow & \searrow & \downarrow{\scriptstyle \pi_B} \\ X/G & \xrightarrow{\ \bar{f}\ } & B/G \end{array}$$

Obviously, the map $\pi_B \circ f$ is smooth and hence $\bar{f}$ is smooth by the universal property of quotients ([5, Proposition 16.10.4]).
For every local trivialization ϕ_U of f define a local trivialization of $\bar{f}$ by

$$\begin{aligned} \bar{\phi}_U : \pi_B(U) \times \mathbb{F}^p &\longrightarrow \pi_X(f^{-1}(U)), \\ ([x]_{\sim_{\sigma_B}}, y) &\mapsto [\phi_U(x,y)]_{\sim_{\sigma_X}}. \end{aligned}$$

From (1) it follows that $\bar{\phi}_U$ is well defined. Since π_B is an open map, it follows that $\pi_B(U)$ is open in B/G. Since π_B is surjective, the sets $\pi_B(U)$ cover B/G. As before, the commutative diagram

$$
\begin{array}{ccc}
U \times \mathbb{F}^p & \xrightarrow{\phi_U} & f^{-1}(U) \\
\pi_B \times \mathrm{id} \downarrow & \searrow & \downarrow \pi_X \\
\pi_B(U) \times \mathbb{F}^p & \xrightarrow{\bar{\phi}_U} & \pi_X(f^{-1}(U))
\end{array}
$$

implies that $\bar{\phi}_U$ is smooth. Since π_X and ϕ_U are both surjective, so is $\pi_X \circ \phi_U$, and hence $\bar{\phi}_U$ is surjective. To see that $\bar{\phi}_U$ is also injective consider $x, x' \in U$ and $y, y' \in \mathbb{F}^p$ with $[\phi_U(x,y)]_{\sim_{\sigma_X}} = [\phi_U(x',y')]_{\sim_{\sigma_X}}$. Then there exists $g \in G$ with $\phi_U(x,y) = \sigma_X(g, \phi_U(x',y')) = \phi_U(\sigma_B(g,x'), y')$, i.e., $x = \sigma_B(g,x')$ and $y = y'$, since ϕ_U is injective. It follows that $([x]_{\sim_{\sigma_B}}, y) = ([x']_{\sim_{\sigma_B}}, y')$ and $\bar{\phi}_U$ is injective. Now the commutative diagram

$$
\begin{array}{ccc}
f^{-1}(U) & \xrightarrow{\phi_U^{-1}} & U \times \mathbb{F}^p \\
\pi_X \downarrow & \searrow & \downarrow \pi_B \times \mathrm{id} \\
\pi_X(f^{-1}(U)) & \xrightarrow{\bar{\phi}_U^{-1}} & \pi_B(U) \times \mathbb{F}^p
\end{array}
$$

implies that $\bar{\phi}_U^{-1}$ is smooth, hence $\bar{\phi}_U$ is a diffeomorphism.
Now let $x \in U$ and $y \in \mathbb{F}^p$. Then

$$
\begin{aligned}
\bar{f}(\bar{\phi}_U([x]_{\sim_{\sigma_B}}, y)) &= \bar{f}([\phi_U(x,y)]_{\sim_{\sigma_X}}) \\
&= [f(\phi_U(x,y))]_{\sim_{\sigma_B}} \\
&= [x]_{\sim_{\sigma_B}}.
\end{aligned}
$$

If ϕ_V is another local trivialization of f, $x \in U \cap V$ and $y \in \mathbb{F}^p$, then

$$
\begin{aligned}
&\bar{\phi}_V([\mathrm{pr}_1 \circ \phi_V^{-1} \circ \phi_U(x,y)]_{\sim_{\sigma_B}}, \mathrm{pr}_2 \circ \phi_V^{-1} \circ \phi_U(x,y)) = \\
&[\phi_V(\mathrm{pr}_1 \circ \phi_V^{-1} \circ \phi_U(x,y), \mathrm{pr}_2 \circ \phi_V^{-1} \circ \phi_U(x,y))]_{\sim_{\sigma_X}} = \\
&\qquad\qquad [\phi_U(x,y)]_{\sim_{\sigma_X}}
\end{aligned}
$$

implies

$$
\begin{aligned}
\bar{\phi}_V^{-1} \circ \bar{\phi}_U([x]_{\sim_{\sigma_B}}, y) &= \bar{\phi}_V^{-1}([\phi_U(x,y)]_{\sim_{\sigma_X}}) \\
&= ([\mathrm{pr}_1 \circ \phi_V^{-1} \circ \phi_U(x,y)]_{\sim_{\sigma_B}}, \mathrm{pr}_2 \circ \phi_V^{-1} \circ \phi_U(x,y)) \\
&= ([x]_{\sim_{\sigma_B}}, \theta_{V,U,x}(y)).
\end{aligned}
$$

Let furthermore $g \in G$ be arbitrary, then successive use of (1) implies

$$\begin{aligned}
\theta_{V,U,\sigma_B(g,x)}(y) &= \mathrm{pr}_2 \circ \phi_V^{-1} \circ \phi_U(\sigma_B(g,x), y) \\
&= \mathrm{pr}_2 \circ \phi_V^{-1} \circ \sigma_X(g, \phi_U(x,y)) \\
&= \mathrm{pr}_2 \circ \phi_V^{-1} \circ \sigma_X(g, \phi_V(\phi_V^{-1} \circ \phi_U(x,y))) \\
&= \mathrm{pr}_2 \circ \phi_V^{-1} \circ \sigma_X(g, \phi_V(x, \theta_{V,U,x}(y))) \\
&= \mathrm{pr}_2 \circ \phi_V^{-1} \circ \phi_V(\sigma_B(g,x), \theta_{V,U,x}(y)) \\
&= \theta_{V,U,x}(y).
\end{aligned}$$

Hence

$$\theta_{V,U,[x]_{\sim_{\sigma_B}}} := \theta_{V,U,x}$$

is well defined and

$$\bar{\phi}_V^{-1} \circ \bar{\phi}_U([x]_{\sim_{\sigma_B}}, y) = ([x]_{\sim_{\sigma_B}}, \theta_{V,U,[x]_{\sim_{\sigma_B}}}(y)).$$

It follows that $\bar{f}$ is a smooth vector bundle with fiber $\mathbb{F}^p$. □

We now present a variant of Theorem 2, to be able to treat product space situations.

Theorem 4. *Let E be a vector space over $\mathbb{F}$ and B be a q-dimensional smooth manifold. Assume that the Lie group G operates smoothly and linearly on E via*

$$\sigma_E : \; G \times E \longrightarrow E, \quad (g,v) \mapsto g \cdot v$$

and smoothly and transitively on B via

$$\sigma_B : \; G \times B \longrightarrow B, \quad (g,b) \mapsto g \cdot b.$$

Let

$$\sigma_{E\times B} : \; G \times (E \times B) \longrightarrow E \times B, \quad (g,(v,b)) \mapsto g \cdot (v,b) := (g\cdot v, g \cdot b)$$

denote the induced action of G on the product manifold $E \times B$ and let $X \subset E \times B$ be a topological subspace which is closed under the action of G, i.e., $x \in X$ implies $g \cdot x \in X$ for all $g \in G$. Suppose, that the continuous map

$$f : \; X \longrightarrow B, \quad (v,b) \mapsto b$$

is surjective. Let $b_0 \in B$ and let

$$E_0 := \{v \in E \,|\, (v,b_0) \in X\}$$

be a p-dimensional vector subspace of E. Assume further, there exists a submanifold $S \subset G$ and an open neighborhood U of $b_0 \in B$ such that

$$\sigma_{b_0} : \; S \longrightarrow U, \quad s \mapsto s \cdot b_0 = \sigma_B(g, b_0)$$

is a diffeomorphism. Then X is a $(q+p)$-dimensional smooth submanifold of $E \times B$ and f is a smooth vector bundle with fiber $\mathbb{F}^p$.

Proof. We want to apply Theorem 2. Note that for all $g \in G$ and $x \in X$

$$f(g \cdot x) = g \cdot f(x) \tag{3}$$

The vector space E_0 is isomorphic to $f^{-1}(b_0) = E_0 \times \{b_0\}$, where $\{b_0\}$ is considered as a trivial vector space. Let

$$\varphi : \; \mathbb{F}^p \longrightarrow E_0$$

be a vector space isomorphism, then

$$h : \; \mathbb{F}^p \longrightarrow f^{-1}(b_0), \quad y \mapsto (\varphi(y), b_0)$$

is also a vector space isomorphism which is clearly smooth as a map into E. For every $g \in G$, the map

$$\sigma_g : \; B \longrightarrow B, \quad b \mapsto g \cdot b$$

is a homeomorphism, and thus $U_g := g \cdot U$ is open for every $g \in G$. Since G operates transitively on B, for every $b \in B$ there exists $g \in G$ such that $b = g \cdot b_0$, hence the open sets U_g, $g \in G$ cover B. For each $g \in G$, define the continuous map

$$\phi_g : \; U_g \times \mathbb{F}^p \longrightarrow f^{-1}(U_g), \quad (b, y) \mapsto g \cdot \sigma_{b_0}^{-1}\left(g^{-1} \cdot b\right) \cdot h(y),$$

then for all $b \in U_g$ and $y \in \mathbb{F}^p$, (3) and

$$\begin{aligned} g \cdot \sigma_{b_0}^{-1}\left(g^{-1} \cdot b\right) \cdot b_0 &= g \cdot \sigma_{b_0}\left(\sigma_{b_0}^{-1}\left(g^{-1} \cdot b\right)\right) \\ &= b \end{aligned} \tag{4}$$

imply that

$$\begin{aligned} f(\phi_g(b, y)) &= f\left(g \cdot \sigma_{b_0}^{-1}\left(g^{-1} \cdot b\right) \cdot h(y)\right) \\ &= g \cdot \sigma_{b_0}^{-1}\left(g^{-1} \cdot b\right) \cdot f(h(y)) \\ &= g \cdot \sigma_{b_0}^{-1}\left(g^{-1} \cdot b\right) \cdot b_0 \\ &= b, \end{aligned}$$

i.e., ϕ_g maps indeed into $f^{-1}(U_g)$. Moreover, (4) implies for $b \in U_g$ that $\left[g \cdot \sigma_{b_0}^{-1}\left(g^{-1} \cdot b\right)\right]^{-1} \cdot b = b_0$ and hence $x = (v, b) \in f^{-1}(U_g) \subset X$ implies

$$\left[g \cdot \sigma_{b_0}^{-1}\left(g^{-1} \cdot f(x)\right)\right]^{-1} \cdot x = \left(\left[g \cdot \sigma_{b_0}^{-1}\left(g^{-1} \cdot f(x)\right)\right]^{-1} \cdot v, b_0\right),$$

which lies in X, since X is closed under the action of G, and hence in $f^{-1}(b_0)$. In particular,

$$\left[g \cdot \sigma_{b_0}^{-1}\left(g^{-1} \cdot b\right)\right]^{-1} \cdot v \in E_0 \tag{5}$$

for all $(v, b) \in f^{-1}(U_g)$. But then the continuous map

$$\psi_g : \ f^{-1}(U_g) \longrightarrow U_g \times \mathbb{F}^p, \quad x \mapsto \Big(f(x), h^{-1}(\big[g \cdot \sigma_{b_0}^{-1}\left(g^{-1} \cdot f(x)\right)\big]^{-1} \cdot x)\Big)$$

is well defined and we get for all $b \in U_g$ and $y \in \mathbb{F}^p$

$$\begin{aligned}\psi_g(\phi_g(b,y)) &= \Big(f(\phi_g(b,y)), h^{-1}(\big[g \cdot \sigma_{b_0}^{-1}\left(g^{-1} \cdot f(\phi_g(b,y))\right)\big]^{-1} \cdot \phi_g(b,y))\Big) \\ &= \Big(b, h^{-1}(\big[g \cdot \sigma_{b_0}^{-1}\left(g^{-1} \cdot b\right)\big]^{-1} \cdot g \cdot \sigma_{b_0}^{-1}\left(g^{-1} \cdot b\right) \cdot h(y))\Big) \\ &= (b,y).\end{aligned}$$

Thus for all $x \in f^{-1}(U_g)$

$$\begin{aligned}\phi_g(\psi_g(x)) &= g \cdot \sigma_{b_0}^{-1}\left(g^{-1} \cdot f(x)\right) \cdot h(h^{-1}(\big[g \cdot \sigma_{b_0}^{-1}\left(g^{-1} \cdot f(x)\right)\big]^{-1} \cdot x) \\ &= x\end{aligned}$$

and therefore ϕ_g is a homeomorphism with inverse ψ_g and hence a local trivialisation of f. Moreover, ϕ_g is smooth as a map into $E \times B$, as it is a concatenation of smooth maps. Let π denote a smooth projector from E onto E_0, then the smooth map

$$k : \ E \times B \longrightarrow \mathbb{F}^p, \quad (v,b) \mapsto \varphi^{-1}(\pi(v))$$

restricts to h^{-1} on $f^{-1}(b_0) = E_0 \times \{b_0\}$. Obviously, f is the restriction of the smooth map

$$F : \ E \times B \longrightarrow B, \quad (v,b) \mapsto b$$

to $X \subset E \times B$. Note that $F^{-1}(U_g)$ is open in $E \times B$ since F is continuous and note further that for $x \in F^{-1}(U_g)$ we have $F(x) \in U_g$ and $g^{-1} \cdot F(x) \in U$. Now $\phi_g^{-1} = \psi_g$ is the restriction of the smooth map

$$\Psi_g : \ F^{-1}(U_g) \longrightarrow U_g \times \mathbb{F}^p, \quad x \mapsto \Big(F(x), k(\big[g \cdot \sigma_{b_0}^{-1}\left(g^{-1} \cdot F(x)\right)\big]^{-1} \cdot x)\Big)$$

to $f^{-1}(U_g)$. Finally, let $g_1, g_2 \in G$ and let ϕ_{g_1} and ϕ_{g_2} be the two associated local trivialisations of f, $b \in U_{g_1} \cap U_{g_2}$, and $y \in \mathbb{F}^p$. From (4),

$$\begin{aligned}\phi_{g_1}(b,y) &= g_1 \cdot \sigma_{b_0}^{-1}\left(g_1^{-1} \cdot b\right) \cdot (\varphi(y), b_0) \\ &= (g_1 \cdot \sigma_{b_0}^{-1}\left(g_1^{-1} \cdot b\right) \cdot \varphi(y), b) \in f^{-1}(U_{g_1}) \cap f^{-1}(U_{g_2})\end{aligned}$$

and hence with

$$\alpha(g_1,g_2,b) := \big[g_2 \cdot \sigma_{b_0}^{-1}\left(g_2^{-1} \cdot b\right)\big]^{-1} \cdot g_1 \cdot \sigma_{b_0}^{-1}\left(g_1^{-1} \cdot b\right)$$

$\alpha(g_1,g_2,b) \cdot \varphi(y) \in E_0$ follows from (5). By (4) then

$$\begin{aligned}\alpha(g_1,g_2,b) \cdot b_0 &= \big[g_2 \cdot \sigma_{b_0}^{-1}\left(g_2^{-1} \cdot b\right)\big]^{-1} \cdot b \\ &= \big[g_2 \cdot \sigma_{b_0}^{-1}\left(g_2^{-1} \cdot b\right)\big]^{-1} \cdot g_2 \cdot \sigma_{b_0}^{-1}\left(g_2^{-1} \cdot b\right) \cdot b_0 \\ &= b_0.\end{aligned}$$

Therefore

$$
\begin{aligned}
\phi_{g_2}^{-1}(\phi_{g_1}(b,y)) &= \Big(f(\phi_{g_1}(b,y)), h^{-1}(\big[g_2\cdot\sigma_{b_0}^{-1}\,(g_2^{-1}\cdot f(\phi_{g_1}(b,y)))\big]^{-1}\cdot \phi_{g_1}(b,y))\Big)\\
&= \Big(b, h^{-1}(\big[g_2\cdot\sigma_{b_0}^{-1}\,(g_2^{-1}\cdot b)\big]^{-1}\cdot g_1\cdot\sigma_{b_0}^{-1}\,(g_1^{-1}\cdot b)\cdot h(y))\Big)\\
&= (b, h^{-1}(\alpha(g_1,g_2,b)\cdot(\varphi(y),b_0)))\\
&= (b, \varphi^{-1}(\alpha(g_1,g_2,b)\cdot\varphi(y))).
\end{aligned}
$$

Since G acts linearly on E and since φ is a vector space isomorphism, the change of coordinates $\theta_{g_1,g_2,b}(y) := \varphi^{-1}(\alpha(g_1,g_2,b)\cdot\varphi(y))$ on $\mathbb{F}^p$ is a linear map. Hence f is a topological vector bundle and the statement follows from Theorem 2. $\square$

3 The Cotangent Bundle of the Flag Manifold

With these results on submanifold criteria being out of the way, we can now introduce our main actors on stage. The first one is the cotangent bundle of a flag manifold and its amazing relation to the geometry of nilpotent matrices. We will explain this connection in detail, using the symplectic nature of the cotangent bundle and by computing an associated momentum map. But first some basic definitions and vocabulary.

Recall, that the Grassmann manifold $\mathrm{Grass}(m,\mathbb{F}^n)$ is defined as the set of m-dimensional $\mathbb{F}$-linear subspaces of $\mathbb{F}^n$. It is a smooth, compact manifold and provides a natural generalization of the familiar projective spaces. Flag manifolds in turn provide a generalization of Grassmannians. To define them consider arbitrary integers $n, r \in \mathbb{N}$. A *flag symbol of type* (n,r) is an r-tupel $a = (a_1,\ldots,a_r)$ of numbers $a_1,\ldots,a_r \in \mathbb{N}$ with $a_1 < \cdots < a_r < n$. The *flag manifold of type* a is the set of partial flags $V_1 \subset \cdots \subset V_r$ of linear subspaces of $\mathbb{F}^n$ with prescribed dimensions $a_1,\cdots a_r$. More precisely,

$$
\mathrm{Flag}(a,\mathbb{F}^n) = \{(V_1,\ldots,V_r) \in \prod_{i=1}^{r} \mathrm{Grass}(a_i,\mathbb{F}^n) \,|\, V_1 \subset \cdots \subset V_r\}.
$$

endowed with the differentiable structure inherited from the product of Grassmannians. For convenience of notation we set $a_0 = 0$, $a_{r+1} = n$, $V_0 = \{0\}$ and $V_{r+1} = \mathbb{F}^n$. Furthermore, we define $b_i := a_i - a_{i-1}$, $i = 1,\ldots,r+1$. In the case of *complete flags*, i.e., $a_i = i$ for $i = 0,\cdots,n$, or, equivalently, $b_i = 1$ for $i = 1,\ldots,n$, we use the simplified notation $\mathrm{Flag}(\mathbb{F}^n)$ instead of $\mathrm{Flag}(a,\mathbb{F}^n)$.

We are interested in the set

$$
M(a,\mathbb{F}^n) = \{(A,(V_1,\ldots,V_r)) \in \mathfrak{gl}_n(\mathbb{F}) \times \mathrm{Flag}(a,\mathbb{F}^n) \,|\, AV_i \subset V_i, i = 1,\ldots,r\}
$$

of pairs of linear maps and the flags they leave invariant. Here $\mathfrak{gl}_n(\mathbb{F})$ denotes the vector space $\mathbb{F}^{n\times n}$ of $n \times n$-matrices. It is also a Lie algebra with the commutator $[A,B] = AB - BA$ as the Lie bracket operation.

Theorem 5. *$M(a, \mathbb{F}^n)$ is a smooth manifold of dimension n^2 and the projection map*

$$\begin{aligned} \pi : M(a, \mathbb{F}^n) &\longrightarrow \mathrm{Flag}(a, \mathbb{F}^n), \\ (A, (V_1, \ldots, V_r)) &\mapsto (V_1, \ldots, V_r) \end{aligned}$$

is a smooth vector bundle.

Proof. We will apply Theorem 4. Clearly, the Lie group $G = \mathrm{GL}_n(\mathbb{F})$ of invertible $n \times n$ matrices operates linearly on the vector space $E := \mathfrak{gl}_n(\mathbb{F})$ by similarity

$$\sigma_E : \ \mathrm{GL}_n(\mathbb{F}) \times \mathfrak{gl}_n(\mathbb{F}) \longrightarrow \mathfrak{gl}_n(\mathbb{F}), \quad (T, A) \mapsto TAT^{-1}$$

and transitively on the flag manifold $B := \mathrm{Flag}(a, \mathbb{F}^n)$ by the canonical action

$$\begin{aligned} \sigma_B : \mathrm{GL}_n(\mathbb{F}) \times \mathrm{Flag}(a, \mathbb{F}^n) &\longrightarrow \mathrm{Flag}(a, \mathbb{F}^n), \\ (T, \mathsf{V}) = (T, (V_1, \ldots, V_r)) &\mapsto T \cdot \mathsf{V} := (TV_1, \ldots, TV_r). \end{aligned} \tag{6}$$

The topological subspace $X := M(a, \mathbb{F}^n) \subset \mathfrak{gl}_n(\mathbb{F}) \times \mathrm{Flag}(a, \mathbb{F}^n)$ is closed under the induced action on the product space, since for every $A \in \mathfrak{gl}_n(\mathbb{F})$ and every $(V_1, \ldots, V_r) \in \mathrm{Flag}(a, \mathbb{F}^n)$ the inclusion $AV_i \subset V_i$ implies $TAT^{-1}TV_i \subset TV_i$ for all $i = 1, \ldots, r$. Since for every flag there exists a linear map that leaves the flag invariant (e.g. the identity), the map $f := \pi$ is surjective.
We set $b_0 := \mathsf{V}_0 \in \mathrm{Flag}(a, \mathbb{F}^n)$, where V_0 is the *standard flag*

$$\mathsf{V}_0 = (V_1, \ldots, V_r) = (\mathrm{colspan} \begin{pmatrix} I_{a_1} \\ 0 \end{pmatrix}, \ldots, \mathrm{colspan} \begin{pmatrix} I_{a_r} \\ 0 \end{pmatrix}). \tag{7}$$

Then the set E_0 of linear maps that leave V_0 invariant is the vector space

$$E_0 := \left\{ \begin{pmatrix} A_{11} & \ldots & \ldots & A_{1(r+1)} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & A_{(r+1)(r+1)} \end{pmatrix} \in \mathfrak{gl}_n(\mathbb{F}) \;\middle|\; \begin{array}{c} A_{ij} \in \mathbb{F}^{b_i \times b_j}, \\ 1 \le i \le j \le r+1 \end{array} \right\}.$$

The open Bruhat cell

$$U = \left\{ (\mathrm{colspan} \begin{pmatrix} I_{b_1} & 0 & \ldots & 0 \\ K_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \vdots & & \ddots & I_{b_i} \\ K_{(i+1)1} & \ldots & \ldots & K_{(i+1)i} \\ \vdots & & & \vdots \\ K_{(r+1)1} & \ldots & \ldots & K_{(r+1)i} \end{pmatrix})_{i=1}^r \;\middle|\; \begin{array}{c} K_{jk} \in \mathbb{F}^{b_j \times b_k}, \\ 1 \le k < j \le r+1 \end{array} \right\}$$

in $\mathrm{Flag}(a, \mathbb{F}^n)$ is an open neighborhood of $b_0 = \mathsf{V}_0$ which is trivially diffeomorphic to the Lie subgroup

$$S := \left\{ \begin{pmatrix} I_{b_1} & 0 & \dots & 0 \\ K_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ K_{(r+1)1} & \dots & K_{(r+1)r} & I_{b_{r+1}} \end{pmatrix} \in \mathrm{GL}_n(\mathbb{F}) \;\middle|\; \begin{matrix} K_{jk} \in \mathbb{F}^{b_j \times b_k}, \\ 1 \le k < j \le r+1 \end{matrix} \right\}$$

of $G = \mathrm{GL}_n(\mathbb{F})$ by

$$\sigma_{\mathsf{V}_0} : \; S \longrightarrow U, \quad T \mapsto T \cdot \mathsf{V}_0.$$

Now the manifold statement follows from Theorem 4. Concerning the dimension formula we observe that the dimension of the vector bundle $M(a, \mathbb{F}^n)$ equals the dimension of the flag manifold plus the dimension of a fibre. The dimension of a fibre is equal to the dimension of E_0, i.e., to the dimension of the space of block upper triangular matrices. On the other hand, the dimension of the flag manifold is equal to the dimension of S, i.e., to the dimension of the set of strictly lower triangular block matrices. As these dimensions add up to n^2, the result follows. □

We want to see that the bundle $M(a, \mathbb{F}^n)$ of Theorem 5 contains an isomorphic copy of the cotangent bundle $\mathrm{T}^*\mathrm{Flag}(a, \mathbb{F}^n)$. Observe that $\mathrm{GL}_n(\mathbb{F})$ acts transitively on $\mathrm{Flag}(a, \mathbb{F}^n)$, with the stabilizer group $H_n = \mathrm{Stab}(\mathsf{V}_0)$ for the standard flag V_0 of (7) given by

$$H_n = \left\{ \begin{pmatrix} A_{11} & \dots & \dots & A_{1(r+1)} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & A_{(r+1)(r+1)} \end{pmatrix} \in \mathrm{GL}_n(\mathbb{F}) \;\middle|\; \begin{matrix} A_{ij} \in \mathbb{F}^{b_i \times b_j}, \\ 1 \le i \le j \le r+1 \end{matrix} \right\}, \tag{8}$$

i.e., by the closed Lie subgroup of $\mathrm{GL}_n(\mathbb{F})$ consisting of all block upper triangular matrices. Let $\mathfrak{gl}_n(\mathbb{F})$ and $\mathfrak{h}_n$ denote the Lie algebras of $\mathrm{GL}_n(\mathbb{F})$ and H_n, respectively. Thus

$$\mathfrak{h}_n = \left\{ \begin{pmatrix} A_{11} & \dots & \dots & A_{1(r+1)} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & A_{(r+1)(r+1)} \end{pmatrix} \in \mathfrak{gl}_n(\mathbb{F}) \;\middle|\; \begin{matrix} A_{ij} \in \mathbb{F}^{b_i \times b_j}, \\ 1 \le i \le j \le r+1 \end{matrix} \right\}. \tag{9}$$

We endow $\mathfrak{gl}_n(\mathbb{F})$ with the nondegenerate symmetric bilinear form

$$(X, Y) := \mathrm{tr}(XY).$$

Note, that the orthogonal complement of $\mathfrak{gl}_n(\mathbb{F})$ with respect to this trace form is exactly the linear subspace $\mathfrak{u}^+$ of strictly upper triangular matrices

$$\mathfrak{u}^+ = \left\{ \begin{pmatrix} 0 & A_{12} & \dots & A_{1(r+1)} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & A_{r(r+1)} \\ 0 & \dots & 0 & 0 \end{pmatrix} \in \mathfrak{gl}_n(\mathbb{F}) \;\middle|\; \begin{matrix} A_{ij} \in \mathbb{F}^{b_i \times b_j}, \\ 1 \leq i < j \leq r+1 \end{matrix} \right\}. \tag{10}$$

For any $\mathsf{V} = T \cdot \mathsf{V}_0 \in \mathrm{Flag}(a, \mathbb{F}^n)$, the fibre $\sigma_{\mathsf{V}}^{-1}(T \cdot \mathsf{V}_0)$ of $\sigma_{\mathsf{V}} : \mathrm{GL}_n(\mathbb{F}) \longrightarrow \mathrm{Flag}(a, \mathbb{F}^n), g \mapsto g \cdot \mathsf{V} = gT \cdot \mathsf{V}_0$, is the stabilizer subgroup

$$\mathrm{Stab}(\mathsf{V}) = \mathrm{Stab}(T \cdot \mathsf{V}_0) = Ad(T)H_n = \{TgT^{-1} \,|\, g \in H_n\}.$$

Therefore the tangent map at the identity matrix I defines a surjective linear map

$$\mathrm{T}_I\, \sigma_{\mathsf{V}} : \mathfrak{gl}_n(\mathbb{F}) \longrightarrow \mathrm{T}_{\mathsf{V}}\, \mathrm{Flag}(a, \mathbb{F}^n)$$

that vanishes exactly on the Lie subalgebra

$$Ad(T)\mathfrak{h}_n = \{TXT^{-1} \mid X \in \mathfrak{h}_n\}.$$

By taking the duals, it follows that the associated dual map

$$\mathrm{T}_I^*\, \sigma_{\mathsf{V}} : \mathrm{T}_{\mathsf{V}}^*\, \mathrm{Flag}(a, \mathbb{F}^n) \longrightarrow \mathfrak{gl}_n^*(\mathbb{F}),\ \lambda \mapsto \lambda \circ \mathrm{T}_I\, \sigma_{\mathsf{V}},$$

maps the cotangent space $\mathrm{T}_{\mathsf{V}}^*\, \mathrm{Flag}(a, \mathbb{F}^n)$ isomorphically onto the image set

$$\{\lambda \in \mathfrak{gl}_n^*(\mathbb{F}) \,|\, Ad(T)\mathfrak{h}_n \subset \mathrm{Ker}\, \lambda\}.$$

The trace form on the Lie algebra defines an isomorphism

$$\tau : \mathfrak{gl}_n^*(\mathbb{F}) \longrightarrow \mathfrak{gl}_n(\mathbb{F}),\ \tau(\lambda) = X_\lambda,$$

where $X_\lambda \in \mathfrak{gl}_n(\mathbb{F})$ denotes the uniquely determined element satisfying $(X_\lambda, Y) = \lambda(Y)$ for all $Y \in \mathfrak{gl}_n(\mathbb{F})$. By using this isomorphism of the Lie algebra $\mathfrak{gl}_n(\mathbb{F})$ with its dual space $\mathfrak{gl}_n^*(\mathbb{F})$, then $\{\lambda \in \mathfrak{gl}_n^*(\mathbb{F}) \,|\, Ad(T)\mathfrak{h}_n \subset \mathrm{Ker}\, \lambda\}$ becomes identified with the orthogonal complement

$$\tau(\{\lambda \in \mathfrak{gl}_n^*(\mathbb{F}) \,|\, Ad(T)\mathfrak{h}_n \subset \mathrm{Ker}\, \lambda\}) = (Ad(T)\mathfrak{h}_n)^\perp = Ad(T)(\mathfrak{h}_n^\perp) = Ad(T)\mathfrak{u}^+.$$

Since $\mathfrak{u}^+$ is invariant under similarity transformations by elements of H_n this yields a well-defined smooth map

$$\mu : \mathrm{T}^*\mathrm{Flag}(a, \mathbb{F}^n) \longrightarrow \mathfrak{gl}_n(\mathbb{F}),\ (\mathsf{V}, \lambda) \mapsto \tau(\lambda \circ \mathrm{T}_I\, \sigma_{\mathsf{V}}) \tag{11}$$

that maps each cotangent space $\mathrm{T}_{\mathsf{V}}^*\, \mathrm{Flag}(a, \mathbb{F}^n)$ isomorphically onto $Ad(T)\mathfrak{u}^+$. By inspection, the image elements are seen to be exactly those matrices A that satisfy $AV_i \subset V_{i-1}$ for $i = 1, \dots, r+1$. Note that $AV_i \subset V_{i-1}$ implies $AV_i \subset V_i$, since $V_{i-1} \subset V_i$, where $i = 1, \dots, r+1$. This shows the following result.

Theorem 6. *The smooth map*

$$\Phi : \mathrm{T}^*\mathrm{Flag}(a, \mathbb{F}^n) \longrightarrow \mathfrak{gl}_n(\mathbb{F}) \times \mathrm{Flag}(a, \mathbb{F}^n), \; (\mathsf{V}, \lambda) \mapsto (\tau(\lambda \circ \mathrm{T}_I\, \sigma_{\mathsf{V}}), \mathsf{V}) \quad (12)$$

maps the cotangent bundle $\mathrm{T}^*\mathrm{Flag}(a, \mathbb{F}^n)$ *diffeomorphically onto the subbundle*

$$\{(A, (V_1, \ldots, V_r)) \in \mathfrak{gl}_n(\mathbb{F}) \times \mathrm{Flag}(a, \mathbb{F}^n) \mid AV_i \subset V_{i-1}, i = 1, \ldots, r+1\}$$

of $M(a, \mathbb{F}^n)$*. This subbundle of dimension* $2 \dim \mathrm{Flag}(a, \mathbb{F}^n)$ *will be denoted by* $N(a, \mathbb{F}^n)$ *in the sequel. In particular, there is a bundle isomorphism of* $\mathrm{T}^*\mathrm{Flag}(a, \mathbb{F}^n)$ *with the homomorphism bundle*

$$\bigoplus_{i=1}^{r} \mathrm{Hom}(V_{i+1}/V_i, V_i).$$

For $i = 1, \ldots, r$ let $V_i^{\perp}$ denote the orthogonal complement of V_i with respect to the Euclidean inner product on $\mathbb{F}^n$. Then we can identify the quotient space V_{i+1}/V_i with $V_i^{\perp} \cap V_{i+1}$ and obtain the bundle isomorphism

$$\mathrm{T}^*\mathrm{Flag}(a, \mathbb{F}^n) \simeq \bigoplus_{i=1}^{r} \mathrm{Hom}(V_i^{\perp} \cap V_{i+1}, V_i).$$

In the case of flag length $r = 1$ we recover the well known diffeomorphic descriptions of the cotangent bundle of the Grassmannian.

Corollary 1.

$$\begin{aligned} \mathrm{T}^*\mathrm{Grass}(m, \mathbb{F}^n) &\simeq \{(A, \mathsf{V}) \in \mathfrak{gl}_n(\mathbb{F}) \times \mathrm{Grass}(m, \mathbb{F}^n) \mid A\mathsf{V} = \{0\}, A\mathbb{F}^n \subset \mathsf{V}\} \\ &\simeq \mathrm{Hom}(\mathsf{V}^{\perp}, \mathsf{V}). \end{aligned}$$

Since the cotangent bundle $\mathrm{T}^*\mathrm{Flag}(a, \mathbb{F}^n)$ can be identified with the subbundle $N(a, \mathbb{F}^n)$ of the bundle $M(a, \mathbb{F}^n)$, it makes sense to consider the restriction of the projection map onto the first factor

$$\mathrm{pr}_1|_{\mathrm{T}^*\mathrm{Flag}(a,\mathbb{F}^n)} : \; \mathrm{T}^*\mathrm{Flag}(a, \mathbb{F}^n) \longrightarrow \mathfrak{gl}_n(\mathbb{F}), \; (A, (V_1, \ldots, V_r)) \mapsto A.$$

The amazing fact now is that the linear operators arising in the image (which is equal to the image of μ) are all nilpotent matrices! This is due to the fact that $AV_i \subset V_{i-1}$ for all $i = 1, \ldots, r+1$ implies $A^{r+1}\mathbb{F}^n = A^{r+1}V_{r+1} \subset V_0 = \{0\}$. Moreover, for the *complete* flag manifold $\mathrm{Flag}(\mathbb{F}^n)$ we conclude that the image of μ coincides with the set $\mathcal{N}_n(\mathbb{F})$ of arbitrary nilpotent $n \times n$-matrices over $\mathbb{F}$. This shows the following equivalent description of the cotangent bundle.

Corollary 2.

$$\mathrm{T}^*\mathrm{Flag}(\mathbb{F}^n) \simeq \{(A, (V_1, \ldots, V_n)) \in \mathcal{N}_n(\mathbb{F}) \times \mathrm{Flag}(\mathbb{F}^n) \mid AV_i \subset V_i, i = 1, \ldots, n\}$$

Of course, all this is immediate by inspection, but in order to gain a better geometric understanding of the connection between cotangent vectors and nilpotent matrices, we make contact with some basic symplectic geometry and Hamiltonian mechanics, specialized to the situation at hand.

Recall, that the cotangent bundle T^*M of an arbitrary smooth manifold M is always a symplectic manifold, implying in particular, that each of the cotangent spaces $\mathrm{T}^*_x M$ carries a canonically defined symplectic form ω (the Liouville form). Now suppose a Lie group G with Lie algebra $\mathfrak{g}$ acts smoothly on M via

$$\sigma : G \times M \longrightarrow M,\ (g, x) \mapsto g \cdot x.$$

Note that each diffeomorphism $\sigma_g : M \longrightarrow M, x \mapsto g \cdot x, g \in G$, lifts (by pull-back) to a diffeomorphism $\sigma_g^* : \mathrm{T}^*M \longrightarrow \mathrm{T}^*M$ on the cotangent bundle. By inspection, these diffeomeorphisms are seen to preserve the Liouville symplectic form on T^*M. Therefore the action σ lifts to a symplectic action on the cotangent bundle

$$\hat{\sigma} : G \times \mathrm{T}^*M \longrightarrow \mathrm{T}^*M,\ (g, (x, \lambda)) \mapsto (g \cdot x, \sigma_g^*(x, \lambda)).$$

The tangent map of the induced smooth map $\sigma_x : G \longrightarrow M, g \mapsto g \cdot x,\ x \in M$ defines a linear map

$$\mathrm{T}_e\, \sigma_x : \mathfrak{g} \longrightarrow \mathrm{T}_x\, M$$

that vanishes exactly on $\mathfrak{g}_x$, the Lie algebra of the stabilizer subgroup G_x of x. In this setting the concept of a moment map for the induced group action on T^*M is defined. It gives a map $\mu : \mathrm{T}^*M \longrightarrow \mathfrak{g}^*$ from the cotangent bundle to the dual of the Lie algebra $\mathfrak{g}$. It is simply defined by the dual of the tangent map $\mathrm{T}_e\, \sigma_x : \mathfrak{g} \longrightarrow \mathrm{T}_x\, M$.

Definition 1. *The* moment map *for the G-action $\hat{\sigma}$ on T^*M is the smooth map*

$$\mu^* : \mathrm{T}^*M \longrightarrow \mathfrak{g}^*,\ (x, \lambda) \mapsto \lambda \circ \mathrm{T}_e\, \sigma_x.$$

Duality provides us with an identification of $\mathfrak{g}$ with $\mathfrak{g}^*$, given a nondegenerate symmetric bilinear form $(\cdot, \cdot)$ on $\mathfrak{g}$. It therefore enables us to redefine the moment map as a map from the cotangent bundle T^*M onto the Lie algebra $\mathfrak{g}$. It should be emphasized that the definition below depends on the above choice of a bilinear form. In contrast, the moment map on the cotangent bundle does not require such choices and is intrinsically defined.

Definition 2. *The* dualized moment map

$$\mu : \mathrm{T}^*M \longrightarrow \mathfrak{g},\ (x, \lambda) \mapsto \mu_x(\lambda)$$

is defined by the characterizing property

$$(\mu_x(\lambda), \xi) = \lambda(\mathrm{T}_e\, \sigma_x(\xi)) \quad \textit{for all } \xi \in \mathfrak{g}.$$

We want to determine the image of the map

$$\mu_x : \mathrm{T}^*_x M \longrightarrow \mathfrak{g}$$

for a given $x \in M$. Let G_x denote the stabilizer subgroup of x in G with Lie algebra $\mathfrak{g}_x$. Let

$$\mathfrak{m}_x := \mathfrak{g}_x^{\perp} := \{\xi \in \mathfrak{g} \,|\, (\xi, \eta) = 0 \ \forall \eta \in \mathfrak{g}_x\}.$$

Since $\mathrm{T}_e\, \sigma_x$ vanishes exactly on $\mathfrak{g}_x$ we see that the image of μ_x^* is given as

$$\mathrm{Im}(\mu_x^*) = \{\lambda \in \mathfrak{g}^* \,|\, \mathfrak{g}_x \subset \mathrm{Ker}\, \lambda\}$$

and therefore

$$\mathrm{Im}(\mu_x) = \mathfrak{m}_x.$$

We now restrict generality by focussing on homogeneous spaces G/H of a Lie group G by a closed Lie subgroup H. Let $\mathfrak{g}$ and $\mathfrak{h}$ denote their Lie algebras, respectively. Thus we consider the transitive G-action on G/H that is defined by left translation

$$\sigma : G \times G/H \longrightarrow G/H, \ (g, \gamma H) \mapsto g\gamma H.$$

It lifts to an action $\hat{\sigma} : G \times \mathrm{T}^*(G/H) \longrightarrow \mathrm{T}^*(G/H)$ on the cotangent bundle. Fix a nondegenerate bilinear form $(\cdot, \cdot)$ on the Lie algebra $\mathfrak{g}$ that is $Ad(G)$-invariant, i.e.

$$(Ad(g)\xi, Ad(g)\eta) = (\xi, \eta) \quad \text{for all } g \in G,\ \xi, \eta \in \mathfrak{g}.$$

Such a form always exists on, e.g., any semisimple Lie algebra $\mathfrak{g}$ and is given by the *Killing form*

$$(\xi, \eta) = \mathrm{tr}(ad_\xi \circ ad_\eta).$$

Thus the dualized moment map

$$\mu : \mathrm{T}^*(G/H) \longrightarrow \mathfrak{g}$$

is well-defined and has image sets

$$\mathrm{Im}(\mu_{gH}) = \mathfrak{m}_{gH},$$

where

$$\mathfrak{m}_{gH} = (Ad(g)\mathfrak{h})^{\perp}$$

denotes the orthogonal complement of the Lie subalgebra $Ad(g)\mathfrak{h}$ with respect to $(\cdot, \cdot)$. By the $Ad(G)$–invariance of $(\cdot, \cdot)$ the above formula then simplifies to

$$\mathrm{Im}(\mu_{gH}) = Ad(g)(\mathfrak{h}^{\perp}).$$

After these generalities let us return to our task of interpreting the projection map pr_1 on the cotangent bundle $\mathrm{T}^*\mathrm{Flag}(a, \mathbb{F}^n)$ as a moment map. Thus we consider the Lie group $G = \mathrm{GL}_n(\mathbb{F})$ with Lie algebra $\mathfrak{gl}_n(\mathbb{F})$. We endow $\mathfrak{gl}_n(\mathbb{F})$ with the $Ad(\mathrm{GL}_n(\mathbb{F}))$–invariant nondegenerate bilinear form

$$(X, Y) := \mathrm{tr}(XY).$$

Choose H_n to denote the parabolic subgroup of $\mathrm{GL}_n(\mathbb{F})$ of all block-upper triangular matrices defined by (8), let $\mathfrak{h}_n$ denote its associated Lie algebra (9). We have observed already, that the orthogonal complement of $\mathfrak{h}_n$ with respect to the trace form is the linear subspace $\mathfrak{u}^+$ of strictly upper triangular matrices (10) satisfying

$$(Ad(T)\mathfrak{h}_n)^\perp = Ad(T)(\mathfrak{h}_n^\perp) = Ad(T)\mathfrak{u}^+.$$

It follows that at every point $\mathsf{V} = T \cdot \mathsf{V}_0 \in \mathrm{Flag}(a, \mathbb{F}^n) \simeq \mathrm{GL}_n(\mathbb{F})/H_n$ the set of image points of the dualized moment map for the canonical $\mathrm{GL}_n(\mathbb{F})$-action on the cotangent bundle $\mathrm{T}^*(\mathrm{GL}_n(\mathbb{F})/H_n)$ coincides with that of the map μ of (11). On the other hand, μ is the first factor in the isomorphism stated in Theorem 6. Thus, under the above identifications, this proves our claim that the projection map

$$\mathrm{pr}_1|_{\mathrm{T}^*\mathrm{Flag}(a,\mathbb{F}^n)}: \ \mathrm{T}^*\mathrm{Flag}(a, \mathbb{F}^n) \longrightarrow \mathfrak{gl}_n(\mathbb{F}), \ (A, (V_1, \ldots, V_r)) \mapsto A$$

coincides with the dualized moment map. We conclude:

Theorem 7. *The set $N(a, \mathbb{F}^n)$ of pairs*

$$\{(A, (V_1, \ldots, V_r)) \in \mathfrak{gl}_n(\mathbb{F}) \times \mathrm{Flag}(a, \mathbb{F}^n) \,|\, AV_i \subset V_{i-1}, i = 1, \ldots, r+1\}$$

carries the structure of a symplectic manifold of dimension $2 \dim \mathrm{Flag}(a, \mathbb{F}^n)$ such that the $\mathrm{GL}_n(\mathbb{F})$-similarity action

$$(T, (A, (V_1, \ldots, V_r)) \mapsto (TAT^{-1}, (TV_1, \ldots, TV_r))$$

becomes a symplectic action. The moment map for this action is

$$\mathrm{pr}_1|_{\mathrm{T}^*\mathrm{Flag}(a,\mathbb{F}^n)}: \ \mathrm{T}^*\mathrm{Flag}(a, \mathbb{F}^n) \longrightarrow \mathfrak{gl}_n(\mathbb{F}), \ (A, (V_1, \ldots, V_r)) \mapsto A.$$

Its image consists of nilpotent matrices.

It is possible to derive similar explicit formulas for the cotangent bundle of homogeneous spaces that are defined by other classical Lie groups. For the fun of it, we mention one further example, the *Lagrangian Grassmann manifold* $\mathrm{LG}(n)$ of n-dimensional Lagrangian subspaces of $\mathbb{F}^{2n}$. Thus $\mathrm{LG}(n)$ is a compact submanifold of the Grassmannian $\mathrm{Grass}(n, \mathbb{F}^{2n})$, consisting of all maximal isotropic subspaces of $\mathbb{F}^{2n}$ with respect to the standard symplectic form Ω. Let

$$\mathrm{Sp}_n(\mathbb{F}) := \{T \in \mathrm{GL}_{2n}(\mathbb{F}) \mid T^\top \Omega T = \Omega\}$$

denote the Lie group of symplectic transformations and

$$\mathfrak{Ham}_n(\mathbb{F}) = \{X \in \mathfrak{gl}_{2n}(\mathbb{F}) \mid X^\top \Omega + \Omega X = 0\}$$

the associated Lie algebra of Hamiltonian $2n \times 2n$-matrices. Then $\mathrm{Sp}_n(\mathbb{F})$ acts transitively on $\mathrm{LG}(n)$ and therefore $\mathrm{LG}(n)$ is a homogeneous space. Thus the above method applies and we obtain the following result. We leave the details of the proof to the reader.

Theorem 8. *The cotangent bundle of the Lagrangian Grassmannian is diffeomorphic to*

$$\mathrm{T}^*\mathrm{LG}(n) \simeq \{(A, \mathsf{V}) \in \mathfrak{Ham}_n(\mathbb{F}) \times \mathrm{LG}(n) \mid A\mathsf{V} = \{0\}, A\mathbb{F}^{2n} \subset \mathsf{V}\}$$

Moreover, the $\mathrm{Sp}_n(\mathbb{F})$*-similarity action*

$$(T, (A, \mathsf{V})) \mapsto (TAT^{-1}, T \cdot \mathsf{V})$$

is a symplectic action. The moment map for this action is

$$\mathrm{pr}_1|_{\mathrm{T}^*\mathrm{LG}(n)} : \ \mathrm{T}^*\mathrm{LG}(n) \longrightarrow \mathfrak{Ham}_n(\mathbb{F}), \ (A, \mathsf{V}) \mapsto A.$$

Its image consists of nilpotent Hamiltonian matrices.

4 A generalization of the Hermann–Martin Lemma

Lemma 1 and Theorem 9 below are taken from the classical paper [15] by Robert Hermann and Clyde Martin (their theorems 4.1-4.3 are formulated in the dual setup of controllable pairs).

Lemma 1 (Hermann-Martin Lemma). *Let* $(C, A) \in \mathbb{F}^{p\times n} \times \mathbb{F}^{n\times n}$ *and consider the map*

$$g : \mathbb{F}^{n\times n} \times \mathbb{F}^{n\times p} \longrightarrow \mathbb{F}^{n\times n},$$
$$(X, Y) \mapsto [X, A] - YC$$

The following statements are equivalent.

1. (C, A) *is observable.*
2. *Let* $Z \in \mathbb{F}^{n\times n}$ *then* $[A, Z] = 0$ *and* $CZ = 0$ *implies* $Z = 0$.
3. *The map* g *is surjective.*

As an immediate consequence we get the following theorem.

Theorem 9 (Hermann-Martin). *Let* $(C, A) \in \mathbb{F}^{p\times n} \times \mathbb{F}^{n\times n}$. *The map*

$$f : \mathrm{GL}_n(\mathbb{F}) \times \mathbb{F}^{n\times p} \longrightarrow \mathfrak{gl}_n(\mathbb{F}),$$
$$(T, J) \mapsto T(A - JC)T^{-1}$$

is a submersion if and only if (C, A) *is observable. Then the* Brunovsky orbit

$$\Gamma_{(C,A)} = \operatorname{Im} f = \{T(A - JC)T^{-1} \in \mathfrak{gl}_n(\mathbb{F}) \,|\, T \in \mathrm{GL}_n(\mathbb{F}), J \in \mathbb{F}^{n\times p}\}$$

is an open submanifold of $\mathfrak{gl}_n(\mathbb{F})$.

Note, that the elements of the Brunovsky orbit are completely characterized by Rosenbrock's theorem [20]. In the sequel we will utilize Theorem 9 to construct various manifold structures that are closely related to the flag manifolds $\mathrm{Flag}(a, \mathbb{F}^n)$ and the manifold $M(a, \mathbb{F}^n)$ of Section 3. These manifolds will turn out to be related to observers for linear control systems (Section 5), as well as to desingularisations of the variety $\mathcal{N}_n(\mathbb{F})$ of nilpotent matrices (Section 6). Fix a pair $(C, A) \in \mathbb{F}^{p\times n} \times \mathbb{F}^{n\times n}$ and recall the following standard definition from geometric control theory [31].

Definition 3. *An* $\mathbb{F}$*-linear subspace* V *of* $\mathbb{F}^n$ *is called* (C, A)-invariant (*or* conditioned invariant) *if there exists a* $J \in \mathbb{F}^{n\times p}$ *such that* $(A - JC)\mathsf{V} \subset \mathsf{V}$. *Such a* J *is called a* friend of V. *An equivalent condition is* $A(\mathsf{V} \cap \operatorname{Ker} C) \subset \mathsf{V}$.

This definition is easily extended to flags of subspaces.

Definition 4. *A flag* $\mathsf{V} = (V_1, \dots, V_r) \in \mathrm{Flag}(a, \mathbb{F}^n)$ *is called* (C, A)-invariant (*or* conditioned invariant) *if its elements have a common friend, i.e., if there exists a* $J \in \mathbb{F}^{n\times p}$ *such that* $(A - JC)V_i \subset V_i$, $i = 1, \dots, r$. *For the sake of brevity we will write* $(A - JC)\mathsf{V} \subset \mathsf{V}$.

Note that with our notation $V_0 = \{0\}$ and $V_{r+1} = \mathbb{F}^n$ a flag $\mathsf{V} = (V_1, \dots, V_r) \in \mathrm{Flag}(a, \mathbb{F}^n)$ is conditioned invariant if and only if $(A - JC)V_i \subset V_i$, $i = 0, \dots, r + 1$. Now consider the set

$$\mathrm{InvJ}(a, \mathbb{F}^n) := \{(J, \mathsf{V}) \in \mathbb{F}^{n\times p} \times \mathrm{Flag}(a, \mathbb{F}^n) \,|\, (A - JC)\mathsf{V} \subset \mathsf{V}\}$$

of conditioned invariant flags and their friends and the set

$$\mathrm{InvTJ}(a, \mathbb{F}^n) := \{(T, J, \mathsf{V}) \in P(a, \mathbb{F}^n) \,|\, T(A - JC)T^{-1}\mathsf{V} \subset \mathsf{V}\},$$

where $P(a, \mathbb{F}^n)$ denotes the product space $\mathrm{GL}_n(\mathbb{F}) \times \mathbb{F}^{n\times p} \times \mathrm{Flag}(a, \mathbb{F}^n)$. As with the flag manifolds we will write $\mathrm{InvJ}(\mathbb{F}^n)$ and $\mathrm{InvTJ}(\mathbb{F}^n)$ in the case of full flags, respectively.

Theorem 10. *Let* $(C, A) \in \mathbb{F}^{p\times n} \times \mathbb{F}^{n\times n}$ *be observable. Then* $\mathrm{InvJ}(a, \mathbb{F}^n)$ *and* $\mathrm{InvTJ}(a, \mathbb{F}^n)$ *are smooth manifolds of dimensions* pn *and* $n^2 + pn$, *respectively.*

Proof. Theorem 9 implies that the map

$$\begin{aligned}\phi : P(a, \mathbb{F}^n) &\longrightarrow \mathfrak{gl}_n(\mathbb{F}) \times \mathrm{Flag}(a, \mathbb{F}^n),\\ (T, J, \mathsf{V}) &\mapsto (T(A - JC)T^{-1}, \mathsf{V})\end{aligned}$$

is a submersion and hence the preimage $\mathrm{InvTJ}(a, \mathbb{F}^n) = \phi^{-1}(M(a, \mathbb{F}^n))$ is a smooth maniflod. Now consider the self-map $\varphi : (T, J, \mathsf{V}) \mapsto (T, J, T \cdot \mathsf{V})$ on the product manifold $P(a, \mathbb{F}^n)$, where the dot denotes the $\mathrm{GL}_n(\mathbb{F})$-action (6) on $\mathrm{Flag}(a, \mathbb{F}^n)$. Clearly, φ is a diffeomorphism and hence Theorem 9 implies that the map

$$\begin{aligned}\psi : P(a, \mathbb{F}^n) &\longrightarrow \mathfrak{gl}_n(\mathbb{F}) \times \mathrm{Flag}(a, \mathbb{F}^n),\\ (T, J, \mathsf{V}) &\mapsto (T(A - JC)T^{-1}, T \cdot \mathsf{V})\end{aligned}$$

is a submersion. Hence the preimage

$$\begin{aligned}\psi^{-1}(M(a, \mathbb{F}^n)) &= \{(T, J, \mathsf{V}) \in P(a, \mathbb{F}^n) \,|\, T(A - JC)T^{-1}T \cdot \mathsf{V} \subset T \cdot \mathsf{V}\}\\ &= \{(T, J, \mathsf{V}) \in P(a, \mathbb{F}^n) \,|\, (A - JC)\mathsf{V} \subset \mathsf{V}\}\\ &= \mathrm{GL}_n(\mathbb{F}) \times \mathrm{InvJ}(a, \mathbb{F}^n)\end{aligned}$$

is a smooth manifold. Now the canonical left action of $\mathrm{GL}_n(\mathbb{F})$ on itself induces the free and proper action

$$\begin{aligned}\sigma : \mathrm{GL}_n(\mathbb{F}) \times \psi^{-1}(M(a, \mathbb{F}^n)) &\longrightarrow \psi^{-1}(M(a, \mathbb{F}^n)),\\ (S, (T, J, \mathsf{V})) &\mapsto (ST, J, \mathsf{V})\end{aligned}$$

and hence by Theorem 1 the orbit space $\psi^{-1}(M(a, \mathbb{F}^n))/\mathrm{GL}_n(\mathbb{F}) = \mathrm{InvJ}(a, \mathbb{F}^n)$ is a smooth manifold. To verify the dimension formula, we focus on $\mathrm{InvJ}(a, \mathbb{F}^n)$. Since ψ is a submersion, the codimension of $\psi^{-1}(M(a, \mathbb{F}^n))$ in $P(a, \mathbb{F}^n)$ equals the codimension of $M(a, \mathbb{F}^n)$ in $\mathfrak{gl}_n(\mathbb{F}) \times \mathrm{Flag}(a, \mathbb{F}^n)$. Thus

$$\begin{aligned}n^2 + pn + \dim \mathrm{Flag}(a, \mathbb{F}^n) - \dim \psi^{-1}(M(a, \mathbb{F}^n)) &=\\ n^2 + \dim \mathrm{Flag}(a, \mathbb{F}^n) - \dim M(a, \mathbb{F}^n) &= \dim \mathrm{Flag}(a, \mathbb{F}^n).\end{aligned}$$

On the other hand, $\dim \psi^{-1}(M(a, \mathbb{F}^n)) = n^2 + \dim \mathrm{InvJ}(a, \mathbb{F}^n)$, and the result follows. □

Other spaces of interest are the following two subsets of $\mathrm{InvJ}(a, \mathbb{F}^n)$ and $\mathrm{InvTJ}(a, \mathbb{F}^n)$, respectively, where all the friends J yield nilpotent maps $A-JC$.

$$\mathrm{NilJ}(a, \mathbb{F}^n) := \{(J, \mathsf{V}) \in \mathbb{F}^{n\times p} \times \mathrm{Flag}(a, \mathbb{F}^n) \,|\, (A-JC)V_i \subset V_{i-1}, i = 1, \dots, r+1\}$$

and

$$\mathrm{NilTJ}(a, \mathbb{F}^n) := \{(T, J, \mathsf{V}) \in P(a, \mathbb{F}^n) \,|\, T(A-JC)T^{-1}V_i \subset V_{i-1}, i = 1, \dots, r+1\}.$$

For the case of full flags it follows from the discussion in Section 3 (cf. Corollary 2) that indeed

$$\mathrm{NilJ}(\mathbb{F}^n) = \{(J, \mathsf{V}) \in \mathrm{InvJ}(\mathbb{F}^n) \mid A - JC \in \mathcal{N}_n(\mathbb{F})\} \quad \text{and}$$
$$\mathrm{NilTJ}(\mathbb{F}^n) = \{(T, J, \mathsf{V}) \in \mathrm{InvTJ}(\mathbb{F}^n) \mid T(A - JC)T^{-1} \in \mathcal{N}_n(\mathbb{F})\}.$$

Theorem 11. *Let $(C, A) \in \mathbb{F}^{p\times n} \times \mathbb{F}^{n\times n}$ be observable. Then* $\mathrm{NilJ}(a, \mathbb{F}^n)$ *and* $\mathrm{NilTJ}(a, \mathbb{F}^n)$ *are smooth manifolds of dimensions $pn - n^2 + 2\dim \mathrm{Flag}(a, \mathbb{F}^n)$ and $pn - 2\dim \mathrm{Flag}(a, \mathbb{F}^n)$, respectively.*

Proof. The proof follows along the lines of the proof of Theorem 10 but replacing $M(a, \mathbb{F}^n)$ with the subbundle $N(a, \mathbb{F}^n)$ from Theorem 6 which is diffeomorphic to the cotangent bundle $\mathrm{T}^*\mathrm{Flag}(a, \mathbb{F}^n)$. □

In view of the Hermann-Martin Lemma 1, it is obvious that the map g is surjective if and only if the restricted map

$$g_S : \mathbb{F}^{n\times n} \times S \longrightarrow \mathbb{F}^{n\times n},$$
$$(X, Y) \mapsto [X, A] - YC$$

is surjective, where S is any vector subspace of $\mathbb{F}^{n\times p}$ with $S + L = \mathbb{F}^{n\times p}$, and

$$L := \{Y \in \mathbb{F}^{n\times p} \mid \exists X \in \mathbb{F}^{n\times n} \colon [X, A] = YC\}.$$

In [30, Lemma 1], an explicit formula for this solution set is given for an arbitrary pair (C, A) in dual Brunovsky canonical form. As an example for a subspace of minimal dimension choose $S^* := L^\perp$, where $\perp$ denotes the orthogonal complement with respect to a given inner product on $\mathbb{F}^{n\times p}$. However, if we want to derive an analogon to Theorem 9 for restricted J-sets S, matters become more complicated, as one needs to prove the simultaneous surjectivity of *all* the maps

$$g_{S,J} : \mathbb{F}^{n\times n} \times S \longrightarrow \mathbb{F}^{n\times n},$$
$$(X, Y) \mapsto [X, A - JC] - YC$$

with $J \in S$. Of course, as is stated in Theorem 9, the choice $S = \mathbb{F}^{n\times p}$ works, but one would like to know whether there is also such an S of smaller or even minimal dimension.
The relevance of this question will become clearer in Section 6, where we relate it to the construction of miniversal deformations for nilpotent similarity orbits and subsequently to resolutions of singularities in $\mathcal{N}_n(\mathbb{F})$. Here is our conjectured generalization of the Hermann-Martin Lemma 1.

Conjecture 1. *Denote by $\perp$ the orthogonal complement with respect to the inner product $(Y, Z) := \mathrm{Re}\,\mathrm{tr}(Y^*Z)$ on $\mathbb{F}^{n\times p}$, where $Y^* := \bar{Y}^\top$ denotes Hermitian transpose. Let $(C, A) \in \mathbb{F}^{p\times n} \times \mathbb{F}^{n\times n}$ be observable. Then for all $J \in S^* := L^\perp$ the map $g_{S^*,J}$ is surjective and hence the map*

$$\begin{aligned} f_{S^*} : \mathrm{GL}_n(\mathbb{F}) \times S^* &\longrightarrow \mathbb{F}^{n \times n}, \\ (T, J) &\mapsto T(A - JC)T^{-1} \end{aligned}$$

is a submersion.

In Section 6 we will prove this conjecture in some special cases. For the time being let us state some of its consequences which follow along the lines of the proofs of Theorem 10 and Theorem 11.

Theorem 12. *Let* $(C, A) \in \mathbb{F}^{p \times n} \times \mathbb{F}^{n \times n}$ *be observable and let* $P^*(a, \mathbb{F}^n)$ *denote the product space* $\mathrm{GL}_n(\mathbb{F}) \times S^* \times \mathrm{Flag}(a, \mathbb{F}^n)$*. If Conjecture 1 is true then all of the following sets are smooth manifolds.*

$$\mathrm{InvJ}^*(a, \mathbb{F}^n) := \{(J, \mathsf{V}) \in S^* \times \mathrm{Flag}(a, \mathbb{F}^n) \,|\, (A - JC)\mathsf{V} \subset \mathsf{V}\}$$
$$\mathrm{InvTJ}^*(a, \mathbb{F}^n) := \{(T, J, \mathsf{V}) \in P^*(a, \mathbb{F}^n) \,|\, T(A - JC)T^{-1}\mathsf{V} \subset \mathsf{V}\}$$
$$\mathrm{NilJ}^*(a, \mathbb{F}^n) := \{(J, \mathsf{V}) \in S^* \times \mathrm{Flag}(a, \mathbb{F}^n) \,|\, (A - JC)V_i \subset V_{i-1}, i = 1, .., r+1\}$$
$$\mathrm{NilTJ}^*(a, \mathbb{F}^n) := \{(T, J, \mathsf{V}) \in P^*(a, \mathbb{F}^n) \,|\, T(A - JC)T^{-1}V_i \subset V_{i-1}, i = 1, .., r+1\}$$

Once again we will write $\mathrm{InvJ}^*(\mathbb{F}^n)$*,* $\mathrm{InvTJ}^*(\mathbb{F}^n)$*,* $\mathrm{NilJ}^*(\mathbb{F}^n)$ *and* $\mathrm{NilTJ}^*(\mathbb{F}^n)$ *in the case of full flags, respectively. In this case we get*

$$\mathrm{NilJ}^*(\mathbb{F}^n) = \{(J, \mathsf{V}) \in \mathrm{InvJ}^*(a, \mathbb{F}^n) \,|\, A - JC \in \mathcal{N}_n(\mathbb{F})\} \quad \textit{and}$$
$$\mathrm{NilTJ}^*(\mathbb{F}^n) = \{(T, J, \mathsf{V}) \in \mathrm{InvTJ}^*(a, \mathbb{F}^n) \,|\, T(A - JC)T^{-1} \in \mathcal{N}_n(\mathbb{F})\}$$

Note that all the starred manifolds are submanifolds of the corresponding unstarred ones.

5 Conditioned invariant subspaces and observers

In the case of flag length one, i.e. for Grassmann manifolds, we give an alternative proof that $\mathrm{InvJ}((k), \mathbb{F}^n)$ is a smooth manifold. This is done by the technique developed in Theorem 2 and Theorem 3, i.e. by forming a smooth vector bundle over a smooth base manifold and subsequent quotient construction. The base manifold is the manifold of tracking observer parameters which is introduced below. This section thus highlights a surprisingly close connection of the previously developed ideas with observer theory.

We consider linear finite-dimensional time-invariant control systems of the following form.

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx, \end{aligned} \tag{13}$$

where $A \in \mathbb{F}^{n \times n}$, $B \in \mathbb{F}^{n \times m}$ and $C \in \mathbb{F}^{p \times n}$. It is known that, given $(C, A) \in \mathbb{F}^{p \times n} \times \mathbb{F}^{n \times n}$, the set

$$\mathrm{Inv}_k(C,A) = \{\mathsf{V} \in \mathrm{Grass(k,\mathbb{F}^n)} \mid \exists J \in \mathbb{F}^{n\times p} : (A-JC)\mathsf{V} \subset \mathsf{V}\}$$

of (C,A)-invariant subspaces with prescribed dimension k allows a stratification into smooth manifolds, the so-called *Brunovsky-Kronecker strata* [6]. However, it is still unclear whether $\mathrm{Inv}_k(C,A)$ is a smooth manifold itself. Consider now the set

$$\mathrm{InvJ}_k := \mathrm{InvJ}((k),\mathbb{F}^n) = \{(J,\mathsf{V}) \in \mathbb{F}^{n\times p} \times \mathrm{Grass}(k,\mathbb{F}^n) \mid (A-JC)\mathsf{V} \subset \mathsf{V}\}$$

of (C,A)-invariant subpaces with fixed dimension and their friends. We want to see that InvJ_k is a smooth manifold by relating it to the manifold of tracking observer parameters through the construction of a smooth vector bundle.

Definition 5. *A* tracking observer *for the linear function Vx of the state of system* (13)*, $V \in \mathbb{F}^{k\times n}$, is a dynamical system*

$$\dot{v} = Kv + Ly + Mu, \tag{14}$$

$K \in \mathbb{F}^{k\times k}$, $L \in \mathbb{F}^{k\times p}$ and $M \in \mathbb{F}^{k\times m}$, which is driven by the input u and by the output y of system (13) *and has the* tracking property*: For every $x(0) \in \mathbb{F}^n$, every $v(0) \in \mathbb{F}^k$ and every input function $u(.)$*

$$v(0) = Vx(0) \;\Rightarrow\; v(t) = Vx(t) \text{ for all } t \in \mathbb{R}.$$

k is called the order *of the observer.*

Note that the tracking property makes a statement about all trajectories of system (13): whatever starting point $x(0)$ and whatever input $u(t)$ is chosen, setting $v(0) := Vx(0)$ must make the observer track the given function.

Theorem 13. *System* (14) *is a tracking observer for Vx if and only if*

$$\begin{aligned} VA - KV &= LC \\ M &= VB. \end{aligned} \tag{15}$$

In this case the tracking error *$e(t) = v(t) - Vx(t)$ is governed by the differential equation $\dot{e} = Ke$.*

Proof. Let the system (14) satisfy equations (15). Set $e(t) = v(t) - Vx(t)$. Then

$$\begin{aligned} \dot{e} &= \dot{v} - V\dot{x} \\ &= (Kv + Ly + Mu) - V(Ax + Bu) \\ &= Kv + LCx + Mu - VAx - VBu \\ &= Kv - KVx + KVx + LCx - VAx + Mu - VBu \\ &= K(v - Vx) - (VA - KV - LC)x + (M - VB)u \\ &= Ke, \end{aligned}$$

where the last equation follows from (15). Now $e(0) = 0$, i.e., $v(0) = Vx(0)$ implies $e(t) = 0$, i.e., $v(t) = Vx(t)$ for all $t \in \mathbb{R}$.
Conversely let (14) be a tracking observer for Vx. Again set $e(t) = v(t) - Vx(t)$. Then

$$\dot{e} = Ke - (VA - KV - LC)x + (M - VB)u.$$

Let $x(0)$ and $u(0)$ be given and set $v(0) = Vx(0)$, i.e., $e(0) = 0$. Then $e(t) = 0$ for all $t \in \mathbb{R}$ implies

$$\begin{aligned}\dot{e}(0) &= Ke(0) - (VA - KV - LC)x(0) + (M - VB)u(0) \\ &= (VA - KV - LC)x(0) + (M - VB)u(0) \\ &= 0.\end{aligned}$$

Since $x(0)$ and $u(0)$ were arbitrary it follows $VA - KV - LC = M - VB = 0$, i.e., equations (15). $\square$

The next theorem provides the link to conditioned invariant subspaces.

Theorem 14. *There exists a tracking observer for the linear function Vx of the state of system* (13) *if and only if $\mathsf{V} = \operatorname{Ker} V$ is (C, A)-invariant.*

Proof. Let the system (14) be a tracking observer for Vx. According to Theorem 13 it follows $VA - KV = LC$. Let $x \in \operatorname{Ker} V \cap \operatorname{Ker} C$. Then $VAx = VAx - KVx = LCx = 0$ and $Ax \in \operatorname{Ker} V$. With $\mathsf{V} = \operatorname{Ker} V$ it follows $A(\mathsf{V} \cap \operatorname{Ker} C) \subset \mathsf{V}$ and V is (C, A)-invariant.
Conversely let $V \in \mathbb{F}^{k \times n}$ and let $\mathsf{V} = \operatorname{Ker} V$ be (C, A)-invariant. There exists $J \in \mathbb{F}^{n \times p}$ such that $(A - JC)\mathsf{V} \subset \mathsf{V}$. But then there exists a matrix $K \in \mathbb{F}^{k \times k}$ such that $V(A - JC) = KV$. Setting $L := VJ$ yields $VA - KV = LC$. Define $M := VB$. According to Theorem 13 the system $\dot{v} = Kv + Ly + Mu$ is a tracking observer for Vx. $\square$

If V is of full row rank k then the spectrum of a corestriction of A to $\operatorname{Ker} V$, i.e., of the map $(A - JC)|_{\mathbb{F}^n/\mathsf{V}}$ where J is a friend of V, is reflected in the matrix K of an appropriate tracking observer for Vx.

Theorem 15. *Let $V \in \mathbb{F}^{k \times n}$ be of full row rank k. For every friend $J \in \mathbb{F}^{n \times p}$ of $\mathsf{V} := \operatorname{Ker} V$ there exists a unique tracking observer for Vx such that K is similar to $(A - JC)|_{\mathbb{F}^n/\mathsf{V}}$. Conversely, for every tracking observer $\dot{v} = Kv + Ly + Mu$ for Vx there exists a friend J of V such that $(A - JC)|_{\mathbb{F}^n/\mathsf{V}}$ is similar to K.*

Proof. Let $(A - JC)\mathsf{V} \subset \mathsf{V}$ then there exists a matrix $K \in \mathbb{F}^{k \times k}$ such that $V(A - JC) = KV$, i.e. such that the following diagram commutes. Since V has full row rank, K is uniquely determined.

$$\begin{array}{ccc} \mathbb{F}^n & \xrightarrow{\;A-JC\;} & \mathbb{F}^n \\ V\big\downarrow & & \big\downarrow V \\ \mathbb{F}^k & \xrightarrow{\;K\;} & \mathbb{F}^k \end{array}$$

This induces a quotient diagram with the induced map $\bar{V}$ an isomorphism.

$$\begin{array}{ccc} \mathbb{F}^n/\mathsf{V} & \xrightarrow{\;(A-JC)_{\mathbb{F}^n/\mathsf{V}}\;} & \mathbb{F}^n/\mathsf{V} \\ \bar{V}\big\downarrow & & \big\downarrow \bar{V} \\ \mathbb{F}^k & \xrightarrow{\;K\;} & \mathbb{F}^k \end{array} \tag{16}$$

But then K is similar to $(A-JC)|_{\mathbb{F}^n/\mathsf{V}}$. Define $L := VJ$ then the first diagram yields $VA - LC = KV$. Define $M := VB$. It follows by Theorem 13 that $\dot{v} = Kv + Ly + Mu$ is a tracking observer for Vx.
Conversely let $\dot{v} = Kv + Ly + Mu$ be a tracking observer for Vx. It follows by Theorem 13 that $VA - KV = LC$. Since V is surjective there exists $J \in \mathbb{F}^{n\times p}$ such that $L = VJ$. But then $V(A - JC) = KV$ and hence $(A - JC)\mathsf{V} \subset \mathsf{V}$, i.e., J is a friend of V. Furthermore, Diagram (16) yields that $(A - JC)|_{\mathbb{F}^n/\mathsf{V}}$ is similar to K. □

If the system (13) is observable then the connection between (C, A)-invariant subspaces and tracking observers can be made very precise using the following manifold structure. First a technical lemma.

Lemma 2. *Let $A \in \mathbb{F}^{n\times n}$, $X \in \mathbb{F}^{n\times k}$ and $B \in \mathbb{F}^{k\times k}$. Then $AX - XB = 0$ implies $A^iX - XB^i = 0$ for all $i \in \mathbb{N}$. In particular, let $A, X \in \mathbb{F}^{n\times n}$ then $[A, X] = AX - XA = 0$ implies $[A^i, X] = A^iX - XA^i = 0$ for all $i \in \mathbb{N}$.*

Proof. The proof is by induction. Assume $A^iX - XB^i = 0$ then $A^iX = XB^i$ and $A^{i+1}X - XB^{i+1} = AA^iX - XBB^i = AXB^i - XBB^i = (AX - XB)B^i = 0$ where the last equality follows from the hypothesis $AX - XB = 0$. □

Theorem 16. *Let the system* (13) *be observable and let*

$$\mathrm{Obs}_k = \{(K, L, M, V) \in \mathbb{F}^{k\times(k+p+m+n)} \mid VA - KV = LC,\ M = VB\}$$

be the set of all order k tracking observer parameters for system (13). Obs_k *is a smooth submanifold of $\mathbb{F}^{k\times(k+p+m+n)}$ of dimension* $\dim \mathrm{Obs}_k = k^2 + kp$. *Its tangent space at the point $(K, L, M, V) \in \mathrm{Obs}_k$ is*

$$T_{(K,L,M,V)}\mathrm{Obs}_k = \{(\dot{K}, \dot{L}, \dot{M}, \dot{V}) \mid -\dot{K}V - \dot{L}C + \dot{V}A - K\dot{V} = \dot{M} - \dot{V}B = 0\}.$$

Proof. Consider the map

$$\begin{aligned} f : \mathbb{F}^{k\times(k+p+m+n)} &\longrightarrow \mathbb{F}^{k\times(n+m)}, \\ (K,L,M,V) &\mapsto (VA-KV-LC, M-VB). \end{aligned}$$

It will be shown that $(0,0)$ is a regular value of f, hence $\mathrm{Obs}_k = f^{-1}(0,0)$ is a smooth submanifold of $\mathbb{F}^{k\times(k+p+m+n)}$. The derivative of f at a point (K,L,M,V) is given by

$$Df: \ (\dot K, \dot L, \dot M, \dot V) \mapsto (-\dot K V - \dot L C + \dot V A - K\dot V, \dot M - \dot V B),$$

where $(\dot K, \dot L, \dot M, \dot V) \in T_{(K,L,M,V)}(\mathbb{F}^{k\times(k+p+m+n)})$.
An element $(\xi,\eta) \in T_{f(K,L,M,V)}(\mathbb{F}^{k\times(n+m)})$ is orthogonal to the image of Df if and only if

$$\operatorname{tr}\xi^*(-\dot K V - \dot L C + \dot V A - K\dot V) + \operatorname{tr}\eta^*(\dot M - \dot V B) = 0$$

for all $(\dot K, \dot L, \dot M, \dot V) \in T_{(K,L,M,V)}(\mathbb{F}^{k\times(k+p+m+n)})$. This is equivalent to

$$\begin{aligned} V\xi^* &= 0 \\ C\xi^* &= 0 \qquad (17)\\ \eta^* &= 0 \\ A\xi^* - \xi^* K &= 0. \qquad (18) \end{aligned}$$

From (18) it follows by Lemma 2 $A^i\xi^* - \xi^* K^i = 0$ for all $i \in \mathbb{N}$. Together with (17) this yields

$$\begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix} \xi^* = 0.$$

Since (C,A) is observable this implies $\xi^* = 0$. It follows that Df is surjective, f is a submersion and hence $(0,0)$ is a regular value of f.
The dimension of $\mathrm{Obs}_k = f^{-1}(0,0)$ is $k(k+p+m+n) - k(n+m) = k^2 + kp$.
From the fibre theorem it follows $T_{(K,L,M,V)}\mathrm{Obs}_k = (Df)^{-1}(0,0)$. □

Corollary 3. *Being an open subset of* Obs_k *the set*

$$\mathrm{Obs}_{k,k} = \{(K,L,M,V) \in \mathrm{Obs}_k \mid \operatorname{rk} V = k\}$$

is a smooth submanifold of $\mathbb{F}^{k\times(k+p+m+n)}$ *of dimension* $k^2 + kp$.

Now consider the *similarity action* on $\mathrm{Obs}_{k,k}$

$$\begin{aligned} \sigma : \mathrm{GL}_k(\mathbb{F}) \times \mathrm{Obs}_{k,k} &\longrightarrow \mathrm{Obs}_{k,k}, \\ (S,(K,L,M,V)) &\mapsto (SKS^{-1}, SL, SM, SV) \end{aligned}$$

and the induced *similarity classes*

$$[K, L, M, V]_\sigma = \{(SKS^{-1}, SL, SM, SV) \mid S \in \mathrm{GL}_k(\mathbb{F})\}.$$

Note that σ is well defined since $VA - KV = LC$ and $M = VB$ imply $SVA - SKS^{-1}SV = S(VA - KV) = SLC$ and $SM = SVB$.

Theorem 17. *The orbit space*

$$\mathrm{Obs}^\sigma_{k,k} = \{[K, L, M, V]_\sigma \mid (K, L, M, V) \in \mathrm{Obs}_{k,k}\}$$

of similarity classes of order k tracking observer parameters for system (13) *is a smooth manifold of dimension* $\dim \mathrm{Obs}^\sigma_{k,k} = kp$.

Proof. Since V has full row rank k for $(K, L, M, V) \in \mathrm{Obs}_{k,k}$, the similarity action is free and has a closed graph mapping (cf. Section 2): $SV = V$ implies $S = I$, furthermore $V_j \to V$ and $S_j V_j \to W$ imply $S_j \to S$ and $W = SV$. Hence the orbit space of σ is a smooth manifold of dimension $\dim \mathrm{Obs}^\sigma_{k,k} = \dim \mathrm{Obs}_{k,k} - \dim \mathrm{GL}_k(\mathbb{F}) = k^2 + kp - k^2 = kp$. $\square$

Finally, we are in the position to prove that InvJ_k is smooth.

Theorem 18. *Let the system* (13) *be observable. For each k the set*

$$\mathrm{InvJ}_k = \{(J, \mathsf{V}) \in \mathbb{F}^{n\times p} \times \mathrm{Grass}(k, \mathbb{F}^n) \mid (A - JC)\mathsf{V} \subset \mathsf{V}\}$$

is a smooth manifold of dimension $\dim \mathrm{InvJ}_k = np$. *The map*

$$\begin{aligned} \bar{f} : \mathrm{InvJ}_{n-k} &\longrightarrow \mathrm{Obs}^\sigma_{k,k}, \\ (J, \mathsf{V}) &\mapsto [K, L, M, V]_\sigma, \end{aligned}$$

defined by $\operatorname{Ker} V = \mathsf{V}$, $M = VB$, $L = VJ$ *and* $KV = VA - LC = V(A - JC)$ *is a smooth vector bundle with fiber* $\mathbb{F}^{(n-k)\times p}$.

Proof. Consider the set

$$\mathcal{M}_{n-k} = \{(J, V) \in \mathbb{F}^{n\times p} \times \mathrm{St}(k, n) \mid (A - JC) \operatorname{Ker} V \subset \operatorname{Ker} V\},$$

where $\mathrm{St}(k, n)$ denotes the set of full row rank $k\times n$ matrices (*Stiefel manifold*). Apparently, if $(J, V) \in \mathcal{M}_{n-k}$ then $\operatorname{Ker} V$ is a codimension k (C, A)-invariant subspace with friend J. Consider the map

$$\begin{aligned} f : \mathcal{M}_{n-k} &\longrightarrow \mathrm{Obs}_{k,k}, \\ (J, V) &\mapsto (K, L, M, V), \end{aligned}$$

where $L = VJ$, $M = VB$ and K is defined as the unique solution of the equation $KV = VA - LC = V(A - JC)$ (cf. Theorem 15, Part 1). By Theorem 15, Part 2, the map f is surjective. Since $K = V(A - JC)V^*(VV^*)^{-1}$, the map f is continuous. Moreover, it is the restriction of a smooth map defined

on $\mathbb{F}^{n\times p} \times \mathrm{St}(k,n)$, which is an open subset of $\mathbb{F}^{n\times p} \times \mathbb{F}^{k\times n}$. According to Corollary 3 the set $\mathrm{Obs}_{k,k}$ is a smooth submanifold of $\mathbb{F}^{k\times(k+p+m+n)}$.
Given V and $L = VJ$, the solution set of $VX = VJ$ is the affine space $V^*(VV^*)^{-1}(VJ) + \prod_{i=1}^{p} \mathrm{Ker}\, V$. Furthermore, $\dim \mathrm{Ker}\, V = n-k$. Therefore, for every $(K,L,M,V) \in \mathrm{Obs}_{k,k}$ the fiber $f^{-1}(K,L,M,V)$ is an affine space of dimension $(k-n)p$.
Let $V_0 \in \mathrm{St}(k,n)$. Since V_0 has full row rank there exists a permutation matrix P_0 such that $V_0P_0 = \begin{pmatrix} X_0 & Y_0 \end{pmatrix}$ with $X_0 \in \mathbb{F}^{k\times k}$ invertible. Then $W = \{\begin{pmatrix} X & Y \end{pmatrix} P_0^{-1} \,|\, X \text{ invertible}\}$ is an open neighborhood of V_0 in $\mathrm{St}(k,n)$ and $\mathrm{Ker}\, V = \{P_0([-X^{-1}Yy]^\top, y^\top)^\top \,|\, y \in \mathbb{F}^{n-k}\}$ for every $V = \begin{pmatrix} X & Y \end{pmatrix} P_0^{-1} \in W$. But then

$$\begin{aligned} \varphi_W : W \times \mathbb{F}^{n-k} &\longrightarrow \mathbb{F}^n \times W, \\ (V,y) &\mapsto (P_0[I_n - \begin{pmatrix} I_k \\ 0 \end{pmatrix} (VP_0 \begin{pmatrix} I_k \\ 0 \end{pmatrix})^{-1} VP_0] \begin{pmatrix} 0 \\ I_{n-k} \end{pmatrix} y, V) = \\ &\quad (P_0([-X^{-1}Yy]^\top, y^\top)^\top, V) \end{aligned}$$

is a smooth injective map mapping $(V, \mathbb{F}^{n-k})$ onto $(\mathrm{Ker}\, V, V)$ for every $V \in W$. Hence φ_W is a homeomorphism onto its image. The inverse map

$$\begin{aligned} \varphi_W^{-1} : \varphi_W(W \times \mathbb{F}^{n-k}) &\longrightarrow W \times \mathbb{F}^{n-k}, \\ (z,V) &\mapsto (V, \begin{pmatrix} 0 & I_{n-k} \end{pmatrix} P_0^{-1} z) \end{aligned}$$

is the restriction of a smooth map defined on all of $\mathbb{F}^n \times \mathbb{F}^{k\times n}$. If $V = \begin{pmatrix} X_1 & Y_1 \end{pmatrix} P_1^{-1} = \begin{pmatrix} X_2 & Y_2 \end{pmatrix} P_2^{-1} \in W_1 \cap W_2$ then the change of coordinate function $\varphi_{W_2}^{-1} \circ \varphi_{W_1}$ induces the invertible linear transformation

$$\vartheta_{W_2,W_1,V} : \; y \mapsto \begin{pmatrix} 0 & I_{n-k} \end{pmatrix} P_2^{-1} P_1 \begin{pmatrix} -X_1^{-1} Y_1 \\ I_{n-k} \end{pmatrix} y$$

on $\mathbb{F}^{n-k}$.
Using p-fold products of φ_W it is now easy to construct local trivializations of f. Given $(K_0, L_0, M_0, V_0) \in \mathrm{Obs}_{k,k}$ choose the neighborhood

$$U := (\mathbb{F}^{k\times k} \times \mathbb{F}^{k\times p} \times \mathbb{F}^{k\times m} \times W) \cap \mathrm{Obs}_{k,k},$$

which is open in $\mathrm{Obs}_{k,k}$. Let pr_1 denote the projection $(z,V) \mapsto z$, pr_2 the projection $(V,y) \mapsto y$ and consider the map

$$\begin{aligned} \phi_U : U \times \mathbb{F}^{(n-k)\times p} &\longrightarrow f^{-1}(U), \\ &(K,L,M,V, \begin{pmatrix} y_1 & \ldots & y_p \end{pmatrix}) \mapsto \\ &(V^*(VV^*)^{-1}L + \begin{pmatrix} \mathrm{pr}_1(\varphi_W(V,y_1)) & \ldots & \mathrm{pr}_1(\varphi_W(V,y_p)) \end{pmatrix}, V), \end{aligned}$$

where y_i, $i = 1,\ldots,p$, denotes the i-th column of the matrix $Y \in \mathbb{F}^{(n-k)\times p}$. Apparently, $f(\phi_U(K,L,M,V,Y)) = (K,L,M,V)$ for all $(K,L,M,V) \in U$

and all $Y \in \mathbb{F}^{(n-k)\times p}$. Furthermore, ϕ_U is bijective by construction. Since φ_W is smooth, so is ϕ_U. Let e_i, $i = 1, \ldots, p$, denote the i-th standard basis vector of $\mathbb{F}^p$. The inverse map

$$\begin{aligned}\phi_U^{-1} : f^{-1}(U) &\longrightarrow U \times \mathbb{F}^{(n-k)\times p},\\ (J,V) &\mapsto (f(J,V), \big(g_1(J,V) \ \ldots \ g_p(J,V)\big)),\end{aligned}$$

where

$$g_i(J,V) = \mathrm{pr}_2(\varphi_W^{-1}([J - V^*(VV^*)^{-1}(VJ)]\mathrm{e}_i, V)), \quad i = 1, \ldots, p,$$

is the restriction of a smooth map defined on $\mathbb{F}^{n\times p} \times \mathrm{St}(k,n)$, which is an open subset of $\mathbb{F}^{n\times p} \times \mathbb{F}^{k\times n}$. It follows that ϕ_U is a homeomorphism. Finally, if $(K,L,M,V) \in U_1 \cap U_2 = (\mathbb{F}^{k\times k} \times \mathbb{F}^{k\times p} \times \mathbb{F}^{k\times m} \times (W_1 \cap W_2)) \cap \mathrm{Obs}_{k,k}$ then the change of coordinate function $\phi_{U_2}^{-1} \circ \phi_{U_1}$ induces the invertible linear transformation

$$\theta_{U_2,U_1,(K,L,M,V)} : \ \big(y_1 \ \ldots \ y_p\big) \mapsto \big(\vartheta_{W_2,W_1,V}(y_1) \ \ldots \ \vartheta_{W_2,W_1,V}(y_p)\big)$$

on $\mathbb{F}^{(n-k)\times p}$.
According to Theorem 2 the set $\mathcal{M}_{n-k}$ is a smooth submanifold of $\mathbb{F}^{n\times p}\times\mathbb{F}^{k\times n}$ of dimension $\dim \mathrm{Obs}_{k,k} + (n-k)p = k^2 + kp + (n-k)p = k^2 + np$. Furthermore, the map f is a smooth vector bundle with fiber $\mathbb{F}^{(n-k)\times p}$.
As has been shown in the proof of Theorem 17, the similarity action σ on $\mathrm{Obs}_{k,k}$ is free and proper (cf. Theorem 1). By the same arguments this is also true for the similarity action on $\mathcal{M}_{n-k}$:

$$\begin{aligned}\sigma : \mathrm{GL}_k(\mathbb{F}) \times \mathcal{M}_{n-k} &\longrightarrow \mathcal{M}_{n-k},\\ (S,(J,V)) &\mapsto (J, SV).\end{aligned}$$

As is well known, the quotient space $\mathrm{St}(k,n)/\mathrm{GL}_k(\mathbb{F})$ is diffeomorphic to $\mathrm{Grass}(n-k, \mathbb{F}^n)$ via $[V]_\sigma \mapsto \mathrm{Ker}\, V$, hence the quotient $\mathcal{M}_{n-k}/\mathrm{GL}_k(\mathbb{F})$ is diffeomorphic to InvJ_{n-k} and the latter is a smooth manifold of dimension $\dim \mathcal{M}_{n-k} - \dim \mathrm{GL}_k(\mathbb{F}) = k^2 + np - k^2 = np$.
Apparently, $(K,L,M,V) \in U$ implies $(SKS^{-1}, SL, SM, SV) \in U$ since S being invertible and $V \in W$ implies $SV \in W$. Furthermore,

$$\begin{aligned}
&\phi_U(\sigma(S,(K,L,M,V)),Y) =\\
&\phi_U((SKS^{-1}, SL, SM, SV), Y) =\\
&((SV)^*(SV(SV)^*)^{-1}SL + \big(\mathrm{pr}_1(\varphi_W(SV,y_1)) \ \ldots \ \mathrm{pr}_1(\varphi_W(SV,y_p))\big), SV) =\\
&(V^*(VV^*)^{-1}L + \big(\mathrm{pr}_1(\varphi_W(V,y_1)) \ \ldots \ \mathrm{pr}_1(\varphi_W(V,y_p))\big), SV) =\\
&\sigma(S, (V^*(VV^*)^{-1}L + \big(\mathrm{pr}_1(\varphi_W(V,y_1)) \ \ldots \ \mathrm{pr}_1(\varphi_W(V,y_p))\big), V)) =\\
&\sigma(S, \phi_U((K,L,M,V),Y)).
\end{aligned}$$

But then Theorem 3 implies that $\bar{f}$ is a smooth vector bundle with fiber $\mathbb{F}^{(n-k)\times p}$, which completes the proof. □

6 A resolution of singularities for nilpotent matrices

In this section we will put our previous results together to construct a resolution of singularities for the variety $\mathcal{N}_n(\mathbb{F})$ of nilpotent $n \times n$ matrices. In order to do so, we first have to construct suitable transversal sections to nilpotent similarity orbits. We give two different ways to do so. The first one is standard and has been introduced by Arnol'd [2]. The second one is inspired by system theory and uses tangent spaces to output injection orbits.

Fix a nilpotent element $A \in \mathcal{N}_n(\mathbb{F})$. The tangent space to the similarity orbit $\mathcal{O} = \{SAS^{-1} \mid S \in \mathrm{GL}_n(\mathbb{F})\}$ at A then is

$$\mathrm{T}_A\,\mathcal{O} = \{[X, A] \mid X \in \mathfrak{gl}_n(\mathbb{F})\}.$$

Definition 6. *Let $\mathcal{S} \subset \mathfrak{gl}_n(\mathbb{F})$ be a linear subspace with*

$$\mathfrak{gl}_n(\mathbb{F}) = \mathcal{S} + T_A\mathcal{O}.$$

Then $A + \mathcal{S}$ is called an affine transverse section *of $\mathcal{O}$ at A.*

Our goal now is to construct such affine transverse sections for similarity orbits of nilpotent matrices. This is closely related to work of Arnol'd [2] on versal deformations of matrices. We briefly review his construction. Consider the positive definite inner product on $\mathfrak{gl}_n(\mathbb{F})$ defined as

$$(X, Y) := \mathrm{Re}\,\mathrm{tr}(X^*Y),$$

where $X^* := \bar{X}^\top$ denotes Hermitian transpose. Let $ad_A : \mathfrak{gl}_n(\mathbb{F}) \longrightarrow \mathfrak{gl}_n(\mathbb{F})$ denote the adjoint transformation by A, i.e., $ad_A(X) := [A, X] = AX - XA$. For any $A \in \mathfrak{gl}_n(\mathbb{F})$, a straightforward computation shows that $\mathrm{Im}(ad_A)^\perp = \mathrm{Ker}(ad_{A^*})$. Thus we have shown

Proposition 2. *For any $A \in \mathfrak{gl}_n(\mathbb{F})$, the affine subspace*

$$\mathcal{S}^{\mathrm{A}} := A + \mathrm{Ker}(ad_{A^*})$$

is an affine transverse section to the similarity orbit $\mathcal{O}$ at A. Here

$$\mathrm{Ker}(ad_{A^*}) = \{X \in \mathfrak{gl}_n(\mathbb{F}) \mid [A^*, X] = 0\}.$$

We refer to $\mathcal{S}^{\mathrm{A}}$ as the Arnol'd slice *to $\mathcal{O}$.*

The Arnol'd slice has the advantage that it is a transversal slice of smallest possible dimension. Thus we have the direct sum decomposition

$$\mathfrak{gl}_n(\mathbb{F}) = \mathrm{Ker}(ad_{A^*}) \oplus \mathrm{T}_A\,\mathcal{O}.$$

We illustrate this construction by an example. Let A denote the sub-regular nilpotent matrix in Jordan canonical form

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The tangent vectors $X \in \mathrm{T}_A \mathcal{O}$ are of the form

$$X = \begin{pmatrix} a & b & c & d \\ e & f & g & h \\ 0 & -e & i & 0 \\ 0 & j & k & 0 \end{pmatrix}$$

with $f = -a - i$. The kernel $\mathrm{Ker}(ad_{A^*})$ consists of all matrices X of the form

$$X = \begin{pmatrix} a & 0 & 0 & 0 \\ b & a & 0 & 0 \\ c & b & a & f \\ d & 0 & 0 & e \end{pmatrix}$$

and thus the Arnol'd slice consists of all matrices of the form

$$\begin{pmatrix} a & 1 & 0 & 0 \\ b & a & 1 & 0 \\ c & b & a & f \\ d & 0 & 0 & e \end{pmatrix}.$$

The nilpotent matrices in this transversal section are characterized by $ec - df - abe = 0$. Note that the block-Toeplitz matrices $X \in \mathrm{Ker}(ad_{A^*})$ can be interpreted as partial reachability matrices $X = (g_1, Fg_1, F^2 g_1, g_2)$, where $F := A^\top$ and $g_1 := (a, b, c, d)^\top$, $g_2 := (0, 0, f, e)^\top$. Indeed, this is no coincidence, as can be seen from [10].

Let us now develop a system theoretic approach to the construction of transversal sections. Choose an output matrix $C \in \mathbb{F}^{p \times n}$, $p \leq n$ suitable, such that (C, A) is in dual Brunovsky canonical form.

Definition 7. *The* Brunovsky slice *is the affine subspace* $\mathcal{S}^{\mathrm{B}} := A + \mathcal{L}_C$, *where* $\mathcal{L}_C$ *denotes the Lie subalgebra of* $\mathfrak{gl}_n(\mathbb{F})$

$$\mathcal{L}_C := \{JC \mid J \in \mathbb{F}^{n \times p}\}.$$

The first observation we make is that the Brunovsky slice is always an affine transversal section to the nilpotent similarity orbit. In fact, the following lemma is an immediate consequence of the Hermann-Martin Lemma 1.

Lemma 3. *Let* $(C, A) \in \mathbb{F}^{p \times n} \times \mathbb{F}^{n \times n}$ *be an observable pair in dual Brunovsky canonical form. Then the Brunovsky slice* $\mathcal{S}^{\mathrm{B}}$ *defines an affine transversal section for the nilpotent similarity orbit* $\mathcal{O}$ *at* A.

The linear subspace $\mathcal{L}_C$ has obvious invariance properties.

Proposition 3. *Let* $(C, A) \in \mathbb{F}^{p\times n} \times \mathbb{F}^{n\times n}$ *be an observable pair in dual Brunovsky canonical form and* $H := [A, A^T]$*. Then*

$$ad_{A^T}(\mathcal{L}_C) \subset \mathcal{L}_C,\ ad_H(\mathcal{L}_C) \subset \mathcal{L}_C,\ ad_A(\mathcal{L}_C) \subset \mathcal{L}_C + \mathcal{L}_{CA}.$$

Proof. The first formula is obvious, in view of the fact that $CA^T = 0$ holds for any for Brunovsky pair. The last formula is also straightforward to see. For the middle one we observe that H is a diagonal matrix. Note, that C has exactly one nonzero entry (= 1) in each row. But then $CH = DC$ for a diagonal matrix D. Thus $ad_H(JC) = HJC - JCH = (HJ - JD)C$ and the result follows. □

While the Brunovsky slice intersects the similarity orbit transversally at A, it is not one of smallest dimension. It is therefore of interest to see, if one can reduce the number of parameters in the transversal subspace $\mathcal{L}_C$ to obtain a transversal section of minimal dimension. The idea here is to replace $\mathcal{L}_C$ by the intersection

$$\mathcal{L}_C^* := \mathcal{L}_C \cap (\mathcal{L}_C \cap T_A\mathcal{O})^\perp$$

of dimension $\dim \mathcal{L}_C^* = \dim \mathcal{L}_C - \dim(\mathcal{L}_C \cap T_A\mathcal{O}) = n^2 - \dim T_A\mathcal{O}$. We refer to

$$\mathcal{S}_{\min}^{\mathrm{B}} := A + \mathcal{L}_C^*$$

as the *minimal Brunovsky slice*. Before presenting an explicit description of $\mathcal{L}_C^*$ in the general case, let us return to the previous example. With C chosen as

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

the Brunovsky slice $\mathcal{L}_C$ consists of matrices of the form

$$X = \begin{pmatrix} a & 0 & 0 & e \\ b & 0 & 0 & f \\ c & 0 & 0 & g \\ d & 0 & 0 & h \end{pmatrix}.$$

Comparing this with the above formula for the tangent space elements we conclude that $\mathcal{L}_C^*$ consists of matrices

$$X = \begin{pmatrix} a & 0 & 0 & 0 \\ b & 0 & 0 & 0 \\ c & 0 & 0 & g \\ d & 0 & 0 & h \end{pmatrix}.$$

Note that this has exactly the same number of parameters as in the Arnol'd slice, as it should be.

We will now show that our Conjecture 1 holds for this specific pair (C, A). In the terminology of Section 4 we have $L = \{J \in \mathbb{F}^{n\times p} \mid JC \in \mathcal{L}_C \cap T_A\mathcal{O}\}$ and

$S^* = \{J \in \mathbb{F}^{n\times p} \,|\, JC \in \mathcal{L}_C^*\}$. We have to show that $\mathcal{L}_C^*$ is transversal to all the spaces $T_J := \{[X, A - JC] \,|\, X \in \mathfrak{gl}_n(\mathbb{F})\}$ for *all* $J \in S^*$, not just for $J = 0$ where $T_0 = \mathrm{T}_A\,\mathcal{O}$. Note that T_J is the tangent space to the similarity orbit of $A - JC$ at $A - JC$. For

$$X = \begin{pmatrix} a & e & i & m \\ b & f & j & n \\ c & g & k & p \\ d & h & l & q \end{pmatrix} \quad \text{compute} \quad [X, A] = \begin{pmatrix} -b & a-f & e-j & -n \\ -c & b-g & f-k & -p \\ 0 & c & g & 0 \\ 0 & d & h & 0 \end{pmatrix}$$

and note that the matrix elements in $[X, A]$ that correspond to the zeros in the elements of $\mathcal{L}_C^*$ can be arbitrarily assigned by choosing values for a, b, c, d (second column), e, f, g, h (third column) and n and p (fourth column) and hence $\mathcal{L}_C^*$ is transversal to T_0. We will focus on these elements and see how they are affected by introducing $J \in S^*$. We have

$$JC = \begin{pmatrix} \alpha & 0 & 0 & 0 \\ \beta & 0 & 0 & 0 \\ \gamma & 0 & 0 & \epsilon \\ \delta & 0 & 0 & \phi \end{pmatrix}$$

and hence looking at the second column of $[X, A]$ the $a - f$ entry is modified by adding $-\alpha e$, the $b - g$ entry by $-\beta e$, the c entry by $-\gamma e - \epsilon h$ and the d entry by $-\delta e - \epsilon h$. The decisive thing is that the modifications do not depend on a, b, c, d, and hence we can compensate for them by choosing those variables accordingly. The analogous argument applies to the third column. In the fourth column, $-n$ is modified by adding $i\epsilon + m\phi - \alpha m$ which is again independent of n, and the analogous argument applies to the remaining entry. Together we get that $\mathcal{L}_C^*$ is transversal to all T_J, $J \in S^*$, and Conjecture 1 follows for this pair (C, A).

A second example is given by a Brunovsky pair (C, A) with generic observability indices $(n, \cdots, n)$. Thus consider the nilpotent $pn \times pn$ block matrix

$$A = \begin{pmatrix} 0 & I_p & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & I_p \\ 0 & \dots & \dots & 0 \end{pmatrix}$$

and

$$C = \begin{pmatrix} I_p & 0 & \dots & 0 \end{pmatrix}.$$

In this case the Brunovsky slice $\mathcal{S}^{\mathrm{B}}$ coincides with the minimal Brunovsky slice $\mathcal{S}^{\mathrm{B}}_{\min}$ and their elements are of the form

$$X = \begin{pmatrix} X_{11} & I_p & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ X_{n-1,1} & 0 & \dots & I_p \\ X_{n1} & 0 & \dots & 0 \end{pmatrix}$$

for suitable $p \times p$ matrices X_{i1}, $i = 1, \dots, n$.

Note, that in this case $\mathcal{L}_C^* = \mathbb{F}^{n\times p} C$ and hence $S^* = \mathbb{F}^{n\times p}$ in the terminology of Section 4. But in this case Conjecture 1 is a direct consequence of the Hermann-Martin Lemma 1, cf. Theorem 9, and is hence also true for the case of generic observability indices.

For a general pair $(C, A) \in \mathbb{F}^{p\times n} \times \mathbb{F}^{n\times n}$ in dual Brunovsky form with observability indices $\mu_1 \geq \cdots \geq \mu_p$, i.e.,

$$C = \left(\begin{array}{c|c|c} 0 \dots 0\,1 & & \\ \hline & \ddots & \\ \hline & & 0 \dots 0\,1 \end{array}\right)_{\underbrace{}_{\mu_1} \cdots \underbrace{}_{\mu_p}} \quad \text{and} \quad A = \left(\begin{array}{c|c|c} \begin{matrix} 0 & & \\ 1 & \ddots & \\ & \ddots & \ddots \\ & & 1 \;\; 0 \end{matrix} & & \\ \hline & \ddots & \\ \hline & & \begin{matrix} 0 & & \\ 1 & \ddots & \\ & \ddots & \ddots \\ & & 1 \;\; 0 \end{matrix} \end{array}\right)_{\underbrace{}_{\mu_1} \cdots \underbrace{}_{\mu_p}},$$

an explicit formula for $\mathcal{L}_C \cap T_A\mathcal{O}$ has been given in [30, Lemma 1]. A straightforward calculation then shows $\mathcal{L}_C^* = \{JC \mid J = (J_{ij})_{i,j=1}^p\}$, where

$$J_{ij} \in \mathbb{F}^{\mu_i \times 1}, \; J_{ij} = \begin{cases} \begin{pmatrix} y_{ij}^1 & \dots & y_{ij}^{\mu_i} \end{pmatrix}^\top & \text{if } \mu_i \leq \mu_j, \\ \begin{pmatrix} y_{ij}^1 & \dots & y_{ij}^{\mu_j} & 0 & \dots & 0 \end{pmatrix}^\top & \text{if } \mu_i > \mu_j. \end{cases}$$

Moreover, we get the dimension formula

$$\dim \mathcal{S}_{\min}^{\mathrm{B}} = \sum_{i,j=1}^{p} \min\{\mu_i, \mu_j\}.$$

In our first example ($\mu_1 = 3$, $\mu_2 = 1$) we would get

$$\mathcal{L}_C^* = \left\{ \begin{pmatrix} 0 & 0 & a & e \\ 0 & 0 & b & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & d & f \end{pmatrix} \middle| \; a, b, c, d, e, f \in \mathbb{F} \right\}$$

which coincides with our previous result after accounting for the permutation of variables that relates the two (C, A) pairs. Note that a change of variables does not affect transversality of $\mathcal{L}_C^*$ and T_J, since $[X, S(A - JC)S^{-1}] - YCS^{-1} = S([S^{-1}XS, A - JC] - S^{-1}YC)S^{-1}$ for all $S \in \mathrm{GL}_n(\mathbb{F})$.

Algebraic geometers have early found interest in the deformation analysis of similarity orbits of semisimple groups. In his address at the International

Congress of Mathematicians in Nice 1970, Brieskorn outlined a program how the singularities of nilpotent matrices may contribute to a deeper understanding of classical geometric problems, such as isolated singularities for complex surfaces. We briefly recall the most important results in this direction, specializing to the simplest case of the general linear group $\mathrm{GL}_n(\mathbb{F})$. Consider the complete flag manifold $\mathrm{Flag}(\mathbb{F}^n)$, given as in Section 3. We have already shown in Theorem 6 and Corollary 2 that

$$\begin{aligned}&\mathrm{T}^*\mathrm{Flag}(\mathbb{F}^n) \simeq N(\mathbb{F}^n) = \\ &\quad \{(A,(V_1,\dots,V_n)) \in \mathfrak{gl}_n(\mathbb{F}) \times \mathrm{Flag}(\mathbb{F}^n) \mid AV_i \subset V_{i-1}, i=1,\dots,n\} = \\ &\qquad \{(A,(V_1,\dots,V_n)) \in \mathcal{N}_n(\mathbb{F}) \times \mathrm{Flag}(\mathbb{F}^n) \mid AV_i \subset V_i, i=1,\dots,n-1\}\end{aligned}$$

is a smooth manifold. Its dimension is $n(n-1)$, i.e. twice the dimension of $\mathrm{Flag}(\mathbb{F}^n)$. The dualized moment map for the natural $\mathrm{GL}_n(\mathbb{F})$-action on the cotangent bundle $\mathrm{T}^*\mathrm{Flag}(\mathbb{F}^n)$ coincides with the projection on the first factor

$$\mathrm{pr}_1 : \mathrm{T}^*\mathrm{Flag}(\mathbb{F}^n) \longrightarrow \mathfrak{gl}_n(\mathbb{F}),\ (A,(V_1,\dots,V_r)) \mapsto A.$$

Moreover, its image set is equal to the singular algebraic variety $\mathcal{N}_n(\mathbb{F})$ of nilpotent matrices. Note, that the dimension of $\mathcal{N}_n(\mathbb{F})$ is also equal to $n(n-1)$ and therefore the two sets have equal dimension. Now suppose that A is a nilpotent matrix with a single Jordan block. Thus we assume that A is cyclic. Then A has a unique A-invariant flag. This shows

Proposition 4 (Steinberg [27]). *Let*

$$\mathrm{T}^*\mathrm{Flag}^{\mathrm{reg}}(\mathbb{F}^n) := \{(A,(V_1,\dots,V_n)) \in N(\mathbb{F}^n) \mid A \text{ cyclic}\}$$

and let $\mathcal{N}_n^{\mathrm{reg}}(\mathbb{F})$ denote the set of cyclic nilpotent $n \times n$ matrices. Then $\mathrm{T}^*\mathrm{Flag}^{\mathrm{reg}}(\mathbb{F}^n)$ and $\mathcal{N}_n^{\mathrm{reg}}(\mathbb{F})$ are smooth manifolds that are open and dense in $\mathrm{T}^*\mathrm{Flag}(\mathbb{F}^n)$ and $\mathcal{N}_n(\mathbb{F})$, respectively. Moreover, the dualized moment map restricts to a diffeomeorphism

$$\mathrm{pr}_1|_{\mathrm{T}^*\mathrm{Flag}^{\mathrm{reg}}(\mathbb{F}^n)} : \mathrm{T}^*\mathrm{Flag}^{\mathrm{reg}}(\mathbb{F}^n) \longrightarrow \mathcal{N}_n^{\mathrm{reg}}(\mathbb{F}).$$

The result shows that, indeed, the dualized moment map defines a desingularization of the nilpotent variety $\mathcal{N}_n(\mathbb{F})$. Next, consider the subregular case of nilpotent matrices with a subgeneric Jordan structure, i.e., one nilpotent block of size $(n-1) \times (n-1)$ and a second (zero block) of size 1×1. These matrices thus constitute a single similarity orbit of nilpotent ones. In contrast to the regular case, the fibres $\mathrm{pr}_1^{-1}(A)$ of a subregular nilpotent matrix are not single points, but form a two-dimensional variety of A-invariant flags. We quote the following result that answers a conjecture of Grothendieck.

Theorem 19 (Brieskorn [4]). *Let A be a subregular nilpotent matrix and $\mathcal{S}$ an $n+2$-dimensional transversal section to the similarity orbit $\mathcal{O}$. Then*

1. *The intersection* $\mathcal{S} \cap \mathcal{N}_n(\mathbb{C})$ *is a two-dimensional complex surface with an isolated singularity at the point A. The singularity is Kleinian and in fact isomorpic to the surface singularity* $\mathbb{C}^2/G$*, where G is the cyclic subgroup of* $\mathrm{SU}(2)$ *of order n.*
2. *If* $\mathcal{S}$ *is chosen sufficiently small then*

$$\mathrm{pr}_1^{-1}(\mathcal{S} \cap \mathcal{N}_n(\mathbb{C}))$$

is a smooth two-dimensional manifold and the projection map pr_1 *restricts to a resolution of singularities of* $\mathcal{S} \cap \mathcal{N}_n(\mathbb{C})$.

As mentioned in the introduction, our goal is to develop a system theoretic approach to such results. This is motivated by the attempt to obtain a better understanding of the transversal slices for subregular nilpotents, constructed in [28]. Note that, although all minimal dimensional transversal sections at a point are conjugate, the right choice of a transversal section still becomes an issue. The situation here is similar to the search for good normal forms in linear algebra. Thus, instead of using the above desingularization via the dualized moment map of the flag manifold, we construct an alternative one where the manifold consists of pairs of conditioned invariant flags together with their friends. The motivation behind this is that it might lead to easier constructions of resolutions for nilpotent orbits of arbitrary co-dimension. So far we have however not achieved that purpose and therefore only explain two partial results.
The first result is an analogon of Proposition 4 for certain unions of nilpotent similarity orbits that arise out of Rosenbrock's theorem, see [20]. Recall from Theorem 11 that

$$\begin{aligned}
&\mathrm{NilTJ}(\mathbb{F}^n) = \\
&\{(T, J, (V_1, \ldots, V_n)) \in P(\mathbb{F}^n) \,|\, T(A - JC)T^{-1}V_i \subset V_{i-1}, i = 1, \ldots, n\} = \\
&\quad \{(T, J, \mathsf{V}) \in P(\mathbb{F}^n) \,|\, T(A - JC)T^{-1}\mathsf{V} \subset \mathsf{V},\, T(A - JC)T^{-1} \in \mathcal{N}_n(\mathbb{F})\}
\end{aligned}$$

is a smooth manifold of dimension $n^2 + (p-1)n$.

Theorem 20. *Let* $(C, A) \in \mathbb{F}^{p\times n} \times \mathbb{F}^{n\times n}$ *denote an observable pair in dual Brunovsky canonical form with observability indices* $\mu_1 \geq \cdots \geq \mu_p$*. Let* $\mathcal{N}^{\mu_1,\ldots,\mu_p}(\mathbb{F}) \subset \mathcal{N}_n(\mathbb{F})$ *denote the set of nilpotent matrices whose nilpotency indices* $n_1 \geq \cdots \geq n_p$ *fullfill the* Rosenbrock conditions

$$\begin{aligned}
&\sum_{i=1}^{j} n_i \geq \sum_{i=1}^{j} \mu_i \quad \text{for } j = 1, \ldots, p-1 \text{ and} \\
&\sum_{i=1}^{p} n_i = \sum_{i=1}^{p} \mu_i .
\end{aligned} \tag{19}$$

Then $\mathcal{N}^{\mu_1,\dots,\mu_p}(\mathbb{F})$ *is a disjoint union of nilpotent similarity orbits which contains* $\mathcal{N}_n^{\mathrm{reg}}$*. Furthermore, we have the surjective smooth map*

$$f: \ \mathrm{NilTJ}(\mathbb{F}^n) \longrightarrow \mathcal{N}^{\mu_1,\dots,\mu_p}(\mathbb{F}), \ (T, J, \mathsf{V}) \mapsto T(A - JC)T^{-1}.$$

Proof. The nilpotency indices form a complete set of invariants for similarity on $\mathcal{N}_n(\mathbb{F})$, hence $\mathcal{N}^{\mu_1,\dots,\mu_p}(\mathbb{F})$ is a disjoint union of nilpotent similarity orbits. Since $n_1 = n$ and $n_2 = \cdots = n_p = 0$ fullfills (19), it contains $\mathcal{N}_n^{\mathrm{reg}}$. According to Rosenbrock's theorem, the elements of the Brunovsky orbit $\Gamma_{(C,A)} = \{T(A - JC)T^{-1} \in \mathfrak{gl}_n(\mathbb{F}) \,|\, T \in \mathrm{GL}_n(\mathbb{F}), J \in \mathbb{F}^{n\times p}\}$ are precisely characterized by (19), where the n_i have to be interpreted as the degrees of the invariant factors of $sI - A$. For nilpotent A, the latter are equal to the nilpotency indices of A, thus f maps indeed into $\mathcal{N}^{\mu_1,\dots,\mu_p}(\mathbb{F})$. Since for every nilpotent map there exists an invariant flag, the map f is surjective. It is clearly smooth as a map into $\mathfrak{gl}_n(\mathbb{F})$. □

The second result generalizes Theorem 19. Recall from Theorem 11 that

$$\begin{aligned}
&\mathrm{NilJ}(\mathbb{F}^n) = \\
&\{(J, (V_1, \dots, V_n)) \in \mathbb{F}^{n\times p} \times \mathrm{Flag}(\mathbb{F}^n) \,|\, (A - JC)V_i \subset V_{i-1}, i = 1, \dots, n\} = \\
&\qquad\qquad \{(J, \mathsf{V}) \in \mathbb{F}^{n\times p} \times \mathrm{Flag}(\mathbb{F}^n) \,|\, (A - JC)\mathsf{V} \subset \mathsf{V}, \ A - JC \in \mathcal{N}_n(\mathbb{F})\}
\end{aligned}$$

is a smooth manifold of dimension $(p - 1)n$.

Theorem 21. *Let* $(C, A) \in \mathbb{F}^{p\times n} \times \mathbb{F}^{n\times n}$ *denote an observable pair in dual Brunovsky canonical form. Let* $\mathcal{S}^{\mathrm{B}}$ *be the Brunovsky slice for the nilpotent similarity orbit* $\mathcal{O}$ *through* A*. Then the surjective smooth map*

$$g: \ \mathrm{NilJ}(\mathbb{F}^n) \longrightarrow \mathcal{S}^{\mathrm{B}} \cap \mathcal{N}_n(\mathbb{F}), \ (J, \mathsf{V}) \mapsto A - JC$$

restricts to a surjective map

$$\mathrm{NilJ}^*(\mathbb{F}^n) = g^{-1}(\mathcal{S}^{\mathrm{B}}_{\min} \cap \mathcal{N}_n(\mathbb{F})) \longrightarrow \mathcal{S}^{\mathrm{B}}_{\min} \cap \mathcal{N}_n(\mathbb{F}).$$

Proof. By definition, g maps into $\mathcal{S}^{\mathrm{B}} \cap \mathcal{N}_n(\mathbb{F})$ and is clearly smooth as a map into $\mathfrak{gl}_n(\mathbb{F})$. Since for every nilpotent map there exists an invariant flag, the map g is surjective. From the definitions of $\mathrm{NilJ}^*(\mathbb{F}^n)$ and $\mathcal{S}^{\mathrm{B}}_{\min}$ it follows that $\mathrm{NilJ}^*(\mathbb{F}^n) = g^{-1}(\mathcal{S}^{\mathrm{B}}_{\min} \cap \mathcal{N}_n(\mathbb{F}))$. □

A consequence of Conjecture 1 would be that $\mathrm{NilJ}^*(\mathbb{F}^n)$ is a smooth submanifold of $\mathrm{NilJ}(\mathbb{F}^n)$, cf. Theorem 12. Hence it is reasonable to formulate the following second conjecture.

Conjecture 2. $\mathrm{NilJ}^*(\mathbb{F}^n)$ *is a smooth submanifold of* $\mathrm{NilJ}(\mathbb{F}^n)$ *and*

$$g: \ \mathrm{NilJ}^*(\mathbb{F}^n) \longrightarrow \mathcal{S}^{\mathrm{B}}_{\min} \cap \mathcal{N}_n(\mathbb{F}), \ (J, \mathsf{V}) \mapsto A - JC$$

defines a resolution of singularities of $\mathcal{S}^{\mathrm{B}}_{\min} \cap \mathcal{N}_n(\mathbb{F})$*.*

In our first example ($\mu_1 = 3$, $\mu_2 = 1$) from above, which is a subregular case, we end up with the same transversal slice as Steinberg in [28]. Note, that in this case we have a proof of Conjecture 1. However, as has been pointed out before, the manifold appearing in the above desingularization differs from that used by Steinberg.

References

1. G. S. Ammar and C. F. Martin. The geometry of matrix eigenvalue methods. *Acta Applicandae Math.* 5: 239–278, 1986.
2. V. I. Arnol'd. On matrices depending on parameters. *Uspekhi Mat. Nauk.* 26: 101–114, 1971.
3. W. Borho and R. MacPherson. Partial resolutions of nilpotent varieties. *Analysis and Topology on Singular Spaces, II, III, Asterisque*, Vol. 101-102, Soc. Math. France, Paris, 1983, pp. 23-74.
4. E. Brieskorn. Singular elements of semi-simple algebraic groups. *Intern. Congress Math. Nice* 2: 279-284, 1970.
5. J. Dieudonné. *Treatise on analysis. Vol. III.* Academic Press, New York, 1972. Translated from the French by I. G. MacDonald, Pure and Applied Mathematics.
6. J. Ferer, F. Puerta and X. Puerta. Differentiable structure of the set of controllable $(A, B)^t$-invariant subspaces. *Linear Algebra and its Appl.* 275-276: 161–177, 1998.
7. P. A. Fuhrmann. *Linear Systems and Operators in Hilbert Space.* McGraw–Hill, New York, 1981.
8. P. A. Fuhrmann. *A Polynomial Approach to Linear Algebra.* Springer Verlag, New York, 1996.
9. P. A. Fuhrmann and U. Helmke. A homeomorphism between observable pairs and conditioned invariant subspaces. *Systems and Control Letters*, 30: 217–223, 1997.
10. P. A. Fuhrmann and U. Helmke. On the parametrization of conditioned invariant subspaces and observer theory. *Linear Algebra and its Applications*, 332–334: 265–353, 2001.
11. I. Gohberg, P. Lancaster and L. Rodman. *Invariant Subspaces of Matrices with Applications.* Wiley, New York, 1986.
12. M. Hazewinkel and C. F. Martin. Representations of the symmetric groups, the specialization order, Schubert cells and systems. *Enseignement Mat.* 29: 53–87, 1983.
13. U. Helmke. *The cohomology of moduli spaces of linear dynamical systems.* Regensburger Math. Schriften, Vol. 24, 1992.
14. U. Helmke. Linear dynamical systems and instantons in Yang-Mills theory. *IMA J. Math. Control Inf.* 3: 151–166, 1986.
15. R. Hermann and C. F. Martin. Applications of algebraic geometry to systems theory. I. *IEEE Trans. Automatic Control*, AC-22(1):19–25, 1977.
16. R. Hermann and C. F Martin. Lie and Morse theory for periodic orbits of vector fields and matrix Riccati equations. I. General Lie-theoretic methods. *Math. Systems Theory* 15: 277–284, 1982.

17. D. Hinrichsen, H. F. Münzner and D. Prätzel-Wolters. Parametrizations of (C, A)–invariant subspaces. *Systems and Control Letters* 1: 192–199, 1981.
18. C. F. Martin. Grassmannian manifolds, Riccati equations and feedback invariants of linear systems. *Geometrical methods for the study of linear systems*, NATO Adv. Study Inst. Ser. C: Math. Phys. Sci., Vol.62, Reidel, Dordrecht-Boston, Mass., pp. 195–211, 1980.
19. X. Puerta and U. Helmke. The topology of the set of conditioned invariant subspaces. *Systems and Control Letters* 40: 97–105, 2000.
20. H. H. Rosenbrock. *State-space and multivariable theory.* John Wiley & Sons, Inc. [Wiley Interscience Division], New York, 1970.
21. M. A. Shayman. On the variety of invariant subspaces of a finite-dimensional linear operator. *Trans. of the Amer. Math. Soc.*, 274: 721–747, 1982.
22. M. A. Shayman. Geometry of the algebraic Riccati equation. I, II. *SIAM J. Control Optim.* 21: 375–394, 395–409, 1983.
23. E. D. Sontag. A remark on bilinear systems and moduli spaces of instantons. *Systems and Control Letters* 9: 361-368, 1987.
24. N. Spaltenstein. On the fixed point set of a unipotent transformation on the flag manifold. *Nederl. Akad. Wetensch. Indag. Math.* 38: 452–458, 1976.
25. T. A. Springer. The unipotent variety of a semisimple group. *Proc. Bombay Colloq. on Alg. Geom.* Oxford Press, London, pp. 452–458, 1969.
26. T. A. Springer. A construction of representations of Weyl groups. *Invent. Math.* 44: 279–293, 1978.
27. R. Steinberg. Desingularization of the unipotent variety. *Invent. Math.* 36: 209–224, 1976.
28. R. Steinberg. Kleinian singularities and unipotent elements. *Proc. of Symposia in Pure Math.* 37: 265–270, 1980.
29. J. Trumpf. *On the geometry and parametrization of almost invariant subspaces and observer theory.* Ph.D. Thesis, Universität Würzburg, 2002.
30. J. Trumpf, U. Helmke, and P. A. Fuhrmann. Towards a compactification of the set of conditioned invariant subspaces. *Systems and Control Letters*, 48: 101–111, 2003.
31. W. M. Wonham. *Linear Multivariable Control: A Geometric Approach.* Springer, New York, 1979.