

Towards a compactification of the set of conditioned invariant subspaces[★]

J. Trumpf, U. Helmke

*Mathematisches Institut
Universität Würzburg
Würzburg, Germany*

P.A. Fuhrmann¹

*Department of Mathematics
Ben-Gurion University of the Negev
Beer Sheva, Israel*

Abstract

A compactification of the set of conditioned invariant subspaces of fixed dimension for an observable pair (C, A) is proposed. It contains the almost conditioned invariant subspaces of the same dimension. In certain cases the compactification is shown to be smooth and a complete geometric description is given in the case of a single output system.

Key words: conditioned invariant subspace; geometric control; Grassmann manifold; parametrization problems; compactification

1 Introduction

Conditioned invariant subspaces were introduced early with the advent of geometric control theory, starting with the work of Basile and Marro and Wonham and Morse, see Wonham [20]. The concept of a conditioned invariant subspace generalizes that of an invariant subspace of a linear operator in a very natural way. Conditioned invariant subspaces have played an important role in

[★] Partially supported by a grant from the German-Israeli Foundation GIF I-526-034.06/97.

¹ Earl Katz Family Chair in Algebraic System Theory

various system theoretic problems, most prominently in observer theory. For example, it is well known (see e.g. Willems [18]) that a d -dimensional observer that tracks a linear function of the state exists if and only if the kernel of the function is a conditioned invariant subspace of codimension d . This result shows that the parametrization problems for functional observers and conditioned invariant subspaces of fixed dimension are closely related. Similar to the dual issues in controller design, the task of parametrizing observers of a given dimension is an important, but essentially open, research problem. The goal of this paper is to contribute to the parametrization of observers and conditioned invariant subspaces. Motivated by the above mentioned connection we focus on the parametrization of conditioned invariant subspaces.

The starting point for work in this area is due to Hinrichsen, Münzner and Prätzel-Wolters [9] who developed a module theoretic approach to the parametrization of conditioned invariant subspaces. In this important paper a cell decomposition has been implicitly described via an echelon type canonical form. More recently, Ferrer, Puerta and Puerta [3], Puerta and Helmke [12] and Fuhrmann and Helmke [5; 7] have obtained further results on the parametrization of the set $\text{Inv}_d(C, A)$ of all d -dimensional conditioned invariant subspaces of a fixed observable pair (C, A) . In particular, kernel as well as image representations were constructed. The link to observer theory has been strengthened as well. Despite all of this prior work the understanding of the structure of the set of all conditioned invariant subspaces is still incomplete.

An important issue that has not received sufficient attention in the literature is that of compactifying the set of conditioned invariant subspaces. It is easy to see that the set $\text{Inv}_d(C, A)$ is never compact and therefore there exist sequences of conditioned invariant subspaces without a limit point in $\text{Inv}_d(C, A)$. Such problems arise in e.g. high gain observer design and the question emerges whether or not one can describe explicitly the set of subspaces that occur as limit points of sequences of conditioned invariant subspaces. Moreover, a system theoretic interpretation of the boundary points of $\text{Inv}_d(C, A)$ is desirable, too. The construction of compactifications of linear systems plays an important role in e.g. the solution of pole placement problems, see Byrnes [2] or Rosenthal [13], but are useful also for analysing the convergence behavior of numerical algorithms for controller design. The same remark holds true for the dual problem of observer design and provides motivation for studying compactifications of conditioned invariant subspaces. However these applications require a deeper investigation of the topology of the compactification and are beyond the scope of this paper.

In the present paper we propose a compactification $\mathcal{C}_d(C, A)$ of $\text{Inv}_d(C, A)$ and study its geometric properties. A first attempt at obtaining a compactification might be to extend the set by including all almost conditioned invariant subspaces, objects introduced in a series of papers by Willems [16; 17; 18].

An example due to Özveren, Verghese and Willsky [11] however shows that this attempt fails by constructing a subspace in the closure of all conditioned invariant subspaces which is not almost conditioned invariant. The compactification we propose is therefore constructed in a different way and depends on an explicit embedding of the set of tight conditioned invariant subspaces in a product of Grassmannians. If the number of system outputs is greater or equal than the subspace dimension then the compactification $\mathcal{C}_d(C, A)$ turns out to be the full Grassmann manifold. For generic observability indices of (C, A) and large subspace dimensions the proposed compactification $\mathcal{C}_d(C, A)$ is shown to be smooth. Moreover, in the single output case it carries the structure of a projective space. In other cases the geometry is considerably more complicated and not yet fully understood.

A different angle to approach this problem is by using and extending polynomial methods that proved to be effective in earlier work. The key to the work of Hinrichsen, Münzner and Prätzel-Wolters [9] as well as that of Fuhrmann and Helmke [5; 7] is a description of a conditioned invariant subspace as the intersection of a polynomial model with a module over the ring of polynomials derived in Fuhrmann [4]. There is evidence to the possibility that the subspaces appearing in the compactification may have an analogous description.

2 Preliminaries

We consider a fixed observable pair $(C, A) \in \mathcal{F}^{p \times n} \times \mathcal{F}^{n \times n}$ ($\mathcal{F} = \mathbb{R}, \mathbb{C}$) in dual Brunovsky canonical form with observability indices $\mu_1 \geq \dots \geq \mu_p \geq 1$, which implies that C has full row rank p . Then $\mu_1 + \dots + \mu_p = n$ and (C, A) is of the form

$$C = \left(\begin{array}{ccc|ccc} 0 & \dots & 0 & 1 & & \\ \hline & & & & \ddots & \\ \hline & & & & & 0 & \dots & 0 & 1 \\ \hline \underbrace{\hspace{1.5cm}}_{\mu_1} & \dots & \underbrace{\hspace{1.5cm}}_{\mu_p} \end{array} \right) \quad \text{and} \quad A = \left(\begin{array}{ccc|ccc} 0 & & & & & \\ \hline 1 & \dots & & & & \\ \hline & & & 1 & 0 & \\ \hline & & & & \ddots & \\ \hline & & & & & 0 & \dots & 0 & 1 \\ \hline \underbrace{\hspace{1.5cm}}_{\mu_1} & \dots & \underbrace{\hspace{1.5cm}}_{\mu_p} \end{array} \right).$$

The following definitions are standard in geometric control.

Definition 1 *A linear subspace $\mathcal{V} \subset \mathcal{F}^n$ is called (C, A) -invariant if there exists an output injection $J \in \mathcal{F}^{n \times p}$ such that $(A - JC)\mathcal{V} \subset \mathcal{V}$. Such a J is called a friend of \mathcal{V} . An equivalent condition is $A(\mathcal{V} \cap \text{Ker } C) \subset \mathcal{V}$. Let $\text{Inv}_d(C, A)$ denote the set of all d -dimensional (C, A) -invariant subspaces.*

A linear subspace $\mathcal{V} \subset \mathcal{F}^n$ is called almost (C, A) -invariant if for every $\varepsilon > 0$ there exists $J \in \mathcal{F}^{n \times p}$ such that $\text{dist}(e^{t(A-JC)} x_0, \mathcal{V}) < \varepsilon$ for all $t \geq 0$ and $x_0 \in \mathcal{V}$. Here 'dist' denotes the euclidean distance.

It has been shown by Willems [16; 17; 18] that every d -dimensional almost (C, A) -invariant subspace is the limit of a sequence of d -dimensional (C, A) -invariant subspaces in the topology of the corresponding Grassmannian. Nevertheless the set of all d -dimensional almost (C, A) -invariant subspaces is not closed in the Grassmannian as the following example due to Özveren, Verghese and Willsky [11] shows.

Example 2 Let $p = 3, n = 6$ and $\mu_1 = \mu_2 = \mu_3 = 2$. Then for every $\varepsilon > 0$ the 3-dimensional subspace

$$\mathcal{V}_\varepsilon = \text{Im} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ -\frac{1}{\varepsilon} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \text{Ker} \begin{pmatrix} 0 & 1 & \varepsilon & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

is (C, A) -invariant while the 3-dimensional limit subspace

$$\mathcal{V}_0 = \text{Im} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \text{Ker} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

is not almost (C, A) -invariant.

The following concept is the dual notion to *coasting subspace* introduced by Willems [16; 17]. It plays a crucial role in our approach to compactifying $\text{Inv}_d(C, A)$.

Definition 3 A (C, A) -invariant subspace \mathcal{V} is called tight if $\mathcal{V} + \text{Ker } C = \mathcal{F}^n$.

If a subspace $\mathcal{V} \in \text{Inv}_d(C, A)$ is tight, then the dimension of $\mathcal{V} \cap \text{Ker } C$ is $d - p$. An immediate consequence is that \mathcal{V} can not be tight unless $d \geq p$. There is the following characterization of tightness in terms of the *restriction indices*.

Let $\mathcal{V} \subset \mathcal{F}^n$ be (C, A) -invariant, then there exists an output injection J such that $(A - JC)\mathcal{V} \subset \mathcal{V}$, i.e. \mathcal{V} is an invariant subspace of the linear map $A - JC$. But then the commutative diagram of linear maps

$$\begin{array}{ccccc} \mathcal{F}^n & \xrightarrow{A - JC} & \mathcal{F}^n & \xrightarrow{C} & \mathcal{F}^p \\ \uparrow & & \uparrow & & \uparrow \\ \mathcal{V} & \xrightarrow{\bar{A}} & \mathcal{V} & \xrightarrow{\bar{C}} & C(\mathcal{V}) \end{array}$$

defines a *restriction* (\bar{C}, \bar{A}) of the pair (C, A) to \mathcal{V} . Since (C, A) is observable, so is (\bar{C}, \bar{A}) . The observability indices of the restriction do not depend on the choice of J . They are called the *restriction indices* of (C, A) with respect to \mathcal{V} . Note that there are at most p nonzero restriction indices and that their sum equals $\dim \mathcal{V}$. They are usually ordered decreasingly. A proof of the following result can be found in Fuhrmann and Helmke [6].

Proposition 4 *A (C, A) -invariant subspace \mathcal{V} is tight if and only if $\text{rk } \bar{C} = \dim C(\mathcal{V}) = p$, i.e. if and only if the smallest restriction index is $\lambda_p \geq 1$. In particular, in the single output case, i.e. for $p = 1$, every (C, A) -invariant subspace is tight.*

3 Construction of the compactification

Let $\mathcal{G}_d(\mathcal{F}^n)$ denote the Grassmann manifold of all d -dimensional subspaces of \mathcal{F}^n , see e.g. Griffiths and Harris [8]. It is well known that $\mathcal{G}_d(\mathcal{F}^n)$ is a compact smooth manifold of dimension $(n - d)d$. For $d \leq p$ we show that this compactifies $\text{Inv}_d(C, A)$.

Theorem 5 *Let $d \leq p$. Then $\text{Inv}_d(C, A)$ is dense in $\mathcal{G}_d(\mathcal{F}^n)$. For $d = p$ actually the subset of tight (C, A) -invariant subspaces is dense in $\mathcal{G}_d(\mathcal{F}^n)$.*

PROOF. With respect to the standard basis any subspace in $\mathcal{G}_d(\mathcal{F}^n)$ is given as the image of a rank d matrix $X \in \mathcal{F}^{n \times d}$. Using results from Ferrer, Puerta and Puerta [3, Theorem 3.1 and Proposition 3.4] it follows that the subspaces in $\text{Inv}_d(C, A)$ with restriction indices $\lambda_1 = \dots = \lambda_d = 1$ are the images of the matrices

$$X = \begin{pmatrix} x_{11}^{(1)} & \dots & x_{1d}^{(1)} \\ \vdots & & \vdots \\ x_{11}^{(\mu_1)} & \dots & x_{1d}^{(\mu_1)} \\ \vdots & & \vdots \\ x_{p1}^{(1)} & \dots & x_{pd}^{(1)} \\ \vdots & & \vdots \\ x_{p1}^{(\mu_p)} & \dots & x_{pd}^{(\mu_p)} \end{pmatrix} \in \mathcal{F}^{n \times d}$$

for which the rows with numbers $\mu_1, \mu_1 + \mu_2, \dots, \mu_1 + \dots + \mu_p$ form a submatrix of full rank d . The set of these matrices is dense in the set of all rank d matrices in $\mathcal{F}^{n \times d}$. Hence already the subset $\text{Inv}_\lambda(C, A)$ of $\text{Inv}_d(C, A)$ formed by all (C, A) -invariant subspaces with restriction indices $\lambda_1 = \dots = \lambda_d = 1$ is dense in $\mathcal{G}_d(\mathcal{F}^n)$. If $d = p$ then $\lambda_p = 1$ and Proposition 4 implies that $\text{Inv}_\lambda(C, A)$ consists of all d -dimensional tight (C, A) -invariant subspaces. \square

For $d > p$ we use a construction closely related to the (more general) notion of compensator couples, which served as a tool in the construction of low order dynamic output feedback compensators, see Schumacher [14].

Definition 6 *Let $d \geq p$ and consider the set*

$$\mathcal{S}_d(C, A) := \{(\mathcal{W}, \mathcal{V}) \in \mathcal{G}_{d-p}(\mathcal{F}^n) \times \mathcal{G}_d(\mathcal{F}^n) \mid \mathcal{W} \subset \mathcal{V} \cap \text{Ker } C, A\mathcal{W} \subset \mathcal{V}\} .$$

The proposed compactification of $\text{Inv}_d(C, A)$ is

$$\mathcal{C}_d(C, A) := \{\mathcal{V} \in \mathcal{G}_d(\mathcal{F}^n) \mid \exists \mathcal{W} \in \mathcal{G}_{d-p}(\mathcal{F}^n) : (\mathcal{W}, \mathcal{V}) \in \mathcal{S}_d(C, A)\} .$$

For $\mathcal{V} \in \text{Inv}_d(C, A)$ we have $\dim(\mathcal{V} \cap \text{Ker } C) \geq d - p$ and $A(\mathcal{V} \cap \text{Ker } C) \subset \mathcal{V}$. Hence choosing any $(d - p)$ -dimensional subspace \mathcal{W} of $\mathcal{V} \cap \text{Ker } C$ yields $(\mathcal{W}, \mathcal{V}) \in \mathcal{S}_d(C, A)$. Therefore the projection $\pi_2(\mathcal{S}_d(C, A)) = \mathcal{C}_d(C, A)$ with

$$\pi_2 : \mathcal{G}_{d-p}(\mathcal{F}^n) \times \mathcal{G}_d(\mathcal{F}^n) \longrightarrow \mathcal{G}_d(\mathcal{F}^n), (\mathcal{W}, \mathcal{V}) \mapsto \mathcal{V}$$

contains $\text{Inv}_d(C, A)$.

For $d = p$ we get $\mathcal{S}_d(C, A) = \{0\} \times \mathcal{G}_d(\mathcal{F}^n)$, hence in this case $\mathcal{C}_d(C, A)$ is the full Grassmannian $\mathcal{G}_d(\mathcal{F}^n)$. According to Theorem 5 this is the right choice.

Proposition 7 $\mathcal{S}_d(C, A)$ is compact and so is $\mathcal{C}_d(C, A)$.

PROOF. Consider the compact space

$$\mathcal{P}_d(\mathcal{F}^n) := \{P \in \mathcal{F}^{n \times n} \mid P^* = P, P^2 = P, \text{rk } P = d\}$$

of rank d hermitian projection operators, which is diffeomorphic to $\mathcal{G}_d(\mathcal{F}^n)$ via $P \mapsto \text{Im } P$ and to $\mathcal{G}_{n-d}(\mathcal{F}^n)$ via $P \mapsto \text{Ker } P$. Then the set

$$\mathcal{P}_d(C, A) := \{(W, V) \in \mathcal{P}_{d-p}(\mathcal{F}^n) \times \mathcal{P}_{n-d}(\mathcal{F}^n) \mid VW = 0, CW = 0, VAW = 0\}$$

is compact being a closed subspace of a compact space, and homeomorphic to $\mathcal{S}_d(C, A)$ via $(W, V) \mapsto (\text{Im } W, \text{Ker } V)$. Since π_2 is continuous, $\mathcal{C}_d(C, A) = \pi_2(\mathcal{S}_d(C, A))$ is compact. \square

An immediate consequence of Proposition 7 is that $\mathcal{C}_d(C, A)$ contains all d -dimensional almost (C, A) -invariant subspaces.

If we could show that $\text{Inv}_d(C, A)$ is dense in $\mathcal{C}_d(C, A)$ then the latter would really be a compactification of $\text{Inv}_d(C, A)$. Unfortunately, we can not show this, except in special cases given below. Other properties of interest of the spaces $\mathcal{C}_d(C, A)$ or $\mathcal{S}_d(C, A)$ are smoothness, connectivity or refined geometric

information. We proceed by showing that generically $\mathcal{S}_d(C, A)$ is a smooth manifold of dimension $(n - d)p$ (which we already know for $d = p$). We need the following lemma on Sylvester equations.

Lemma 8 *Let (C, A) be in dual Brunovsky canonical form. The solutions $X \in \mathcal{F}^{n \times n}$ and $Y \in \mathcal{F}^{n \times p}$ of*

$$AX - XA = YC \quad (\text{S})$$

are given by $X = (X_{ij})_{i,j=1}^p$, $X_{ij} \in \mathcal{F}^{\mu_i \times \mu_j}$, and $Y = (Y_{ij})_{i,j=1}^p$, $Y_{ij} \in \mathcal{F}^{\mu_i \times 1}$, where

$$\begin{aligned} \bullet X_{ij} &= \begin{pmatrix} x_{ij}^{(1)} & 0 & \dots & \dots & \dots & \dots & 0 \\ x_{ij}^{(2)} & x_{ij}^{(1)} & 0 & \dots & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ x_{ij}^{(\mu_i)} & \dots & x_{ij}^{(2)} & x_{ij}^{(1)} & 0 & \dots & 0 \end{pmatrix}, Y_{ij} = 0 \text{ if } \mu_i \leq \mu_j \text{ and} \\ \bullet X_{ij} &= \begin{pmatrix} y_{ij}^{(\mu_i - \mu_j - 1)} & 0 & \dots & 0 \\ \vdots & y_{ij}^{(\mu_i - \mu_j - 1)} & \ddots & \vdots \\ y_{ij}^{(0)} & \vdots & \ddots & 0 \\ x_{ij}^{(1)} & y_{ij}^{(0)} & y_{ij}^{(\mu_i - \mu_j - 1)} & \\ \vdots & x_{ij}^{(1)} & \ddots & \vdots \\ x_{ij}^{(\mu_j - 1)} & \vdots & \ddots & y_{ij}^{(0)} \\ x_{ij}^{(\mu_j)} & x_{ij}^{(\mu_j - 1)} & \dots & x_{ij}^{(1)} \end{pmatrix}, Y_{ij} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ y_{ij}^{(\mu_i - \mu_j - 1)} \\ \vdots \\ y_{ij}^{(0)} \end{pmatrix} \text{ if } \mu_i > \mu_j. \end{aligned}$$

The solutions X, Y of equation (S) with $CX = 0$ and $CY = 0$ are given by

$$\begin{aligned} \bullet X_{ij} &= 0, Y_{ij} = 0 \text{ if } \mu_i \leq \mu_j + 1 \text{ and} \\ \bullet X_{ij} &= \begin{pmatrix} y_{ij}^{(\mu_i - \mu_j - 1)} & 0 & \dots & 0 \\ \vdots & y_{ij}^{(\mu_i - \mu_j - 1)} & \ddots & \vdots \\ y_{ij}^{(1)} & \vdots & \ddots & 0 \\ 0 & y_{ij}^{(1)} & y_{ij}^{(\mu_i - \mu_j - 1)} & \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \ddots & y_{ij}^{(1)} \\ \vdots & \vdots & \ddots & 0 \\ 0 & \vdots & \ddots & 0 \\ 0 & 0 & \dots & 0 \end{pmatrix}, Y_{ij} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ y_{ij}^{(\mu_i - \mu_j - 1)} \\ \vdots \\ y_{ij}^{(1)} \\ 0 \end{pmatrix} \text{ if } \mu_i > \mu_j + 1. \end{aligned}$$

PROOF. It is well known that the dimension of the kernel of the linear map $X \mapsto AX - XA$ is $\sum_{i,j=1}^p \min\{\mu_i, \mu_j\}$ (see e.g. Horn and Johnson [10, Corollary 4.4.15]). Since the stated X and Y fulfill equation (S) and the number of free parameters in X is equal to that dimension, it remains to show that if equation (S) has a solution then Y is of the stated form.

Let i, j be fixed and let $X_{ij} = (x_{ij}^{(kl)})$, where $k = 1, \dots, \mu_i$ and $l = 1, \dots, \mu_j$. Let $Y_{ij} = (y_{ij}^{(k)})$, where the row numbering $k = 0, \dots, \mu_i - 1$ runs from bottom to top. Multiplying X from the left by A moves every row of X_{ij} down by one, dropping out the last row and filling in a zero row on top. Multiplying X from the right by A moves every column of X_{ij} to the left by one, dropping out the left most column and filling in a zero column on the right. On the other hand all columns of $(YC)_{ij}$ are zero except for the right most column which is equal to Y_{ij} . Hence $AX - XA = YC$ implies (reading the first row of the ij -block) $x_{ij}^{(12)} = \dots = x_{ij}^{(1\mu_j)} = 0$ and $y_{ij}^{(\mu_i-1)} = 0$. Assume that $1 \leq k < \min\{\mu_j, \mu_i\}$ and $x_{ij}^{(k[k+1])} = \dots = x_{ij}^{(k\mu_j)} = 0$. Then $AX - XA = YC$ implies (reading the $(k+1)$ th row of the ij -block) $-x_{ij}^{([k+1][k+2])} = \dots = -x_{ij}^{([k+1]\mu_j)} = 0$ and $y_{ij}^{(\mu_i-k-1)} = 0$. By induction we have $y_{ij}^{(\mu_i-1)} = \dots = y_{ij}^{(\mu_i-\min\{\mu_j, \mu_i\})} = 0$ and Y is of the stated form.

The second statement of the lemma follows immediately from the special forms of C, X and Y . \square

Now we are in the position to prove the smoothness of $\mathcal{S}_d(C, A)$, at least in the generic case $\mu_1 = \dots = \mu_\nu = k + 1$ and $\mu_{\nu+1} = \dots = \mu_p = k$, where $n = kp + \nu$ (cf. Wonham [20, Corollary 5.4]). Note that this includes the single output case ($p = 1$).

Theorem 9 *Let $\max\{|\mu_i - \mu_j| : 1 \leq i, j \leq p\} \leq 1$ then $\mathcal{S}_d(C, A)$ is a smooth submanifold of $\mathcal{G}_{d-p}(\mathcal{F}^n) \times \mathcal{G}_d(\mathcal{F}^n)$ of dimension $(n-d)p$. In particular, in the single output case $\mathcal{S}_d(C, A)$ is a smooth submanifold of $\mathcal{G}_{d-1}(\mathcal{F}^n) \times \mathcal{G}_d(\mathcal{F}^n)$ of dimension $(n-d)$.*

PROOF. For $k \leq n$ let $V_k(\mathcal{F}^n) := \{X \in \mathcal{F}^{n \times k} \mid \text{rk } X = k\}$ and $V_k^\top(\mathcal{F}^n) := \{X \in \mathcal{F}^{k \times n} \mid \text{rk } X = k\}$ denote the Stiefel manifolds of rank k matrices. Consider the set of matrices

$$\mathcal{M}_d(C, A) := \{(W, V) \in V_{d-p}(\mathcal{F}^n) \times V_{n-d}^\top(\mathcal{F}^n) \mid VW = 0, CW = 0, VAW = 0\}$$

and the map

$$f : V_{d-p}(\mathcal{F}^n) \times V_{n-d}^\top(\mathcal{F}^n) \longrightarrow \mathcal{F}^{(n-d) \times (d-p)} \times \mathcal{F}^{p \times (d-p)} \times \mathcal{F}^{(n-d) \times (d-p)}, \\ (W, V) \mapsto (VW, CW, VAW).$$

We will show that the derivative $Df(W, V)$

$$(\dot{W}, \dot{V}) \mapsto (\dot{V}W + V\dot{W}, C\dot{W}, \dot{V}AW + V\dot{A}W)$$

of f at a point $(W, V) \in \mathcal{M}_d(C, A)$ is surjective, hence $\mathcal{M}_d(C, A) = f^{-1}(0, 0, 0)$ is a smooth submanifold of $V_{d-p}(\mathcal{F}^n) \times V_{n-d}^\top(\mathcal{F}^n)$.

Let $(W, V) \in \mathcal{M}_d(C, A)$ and let $(Z_1, Z_2, Z_3) \in T_{f(W, V)}(\mathcal{F}^{[(n-d)+p+(n-d)] \times (d-p)})$ be orthogonal to the image of $Df(W, V)$. Then

$$\text{tr} \left[Z_1^*(\dot{V}W + V\dot{W}) + Z_2^*C\dot{W} + Z_3^*(\dot{V}AW + VAW\dot{W}) \right] = 0$$

for all $(\dot{W}, \dot{V}) \in T_{(W, V)}(\mathbb{V}_{d-p}(\mathcal{F}^n) \times \mathbb{V}_{n-d}^\top(\mathcal{F}^n)) \simeq \mathcal{F}^{n \times (d-p)} \times \mathcal{F}^{(n-d) \times n}$. It follows

$$\begin{aligned} Z_1^*V + Z_2^*C + Z_3^*VA &= 0 \quad \text{and} \\ WZ_1^* + AWZ_3^* &= 0 . \end{aligned}$$

Multiplying the first equation by W from the left and the second one by V from the right and setting $Y_1 := WZ_1^*V$, $Y_2 := WZ_2^*$ and $Y_3 := WZ_3^*V$ we have

$$\begin{aligned} Y_1 + Y_2C + Y_3A &= 0 \quad \text{and} \\ Y_1 + AY_3 &= 0 . \end{aligned}$$

Plugging the second equation into the first yields the Sylvester equation $AY_3 - Y_3A = Y_2C$. Since $(W, V) \in \mathcal{M}_d(C, A)$ it is $CW = 0$ hence $CY_2 = 0$ and $CY_3 = 0$. By Lemma 8 we get $Y_2 = 0$ and $Y_3 = 0$ hence also $Y_1 = 0$. Since W and V have full column rank and row rank, respectively, this implies $Z_1^* = 0$, $Z_2^* = 0$ and $Z_3^* = 0$. It follows that $Df(W, V)$ is surjective.

We calculate

$$\begin{aligned} \dim \mathcal{M}_d(C, A) &= \dim \mathbb{V}_{d-p}(\mathcal{F}^n) + \dim \mathbb{V}_{n-d}^\top(\mathcal{F}^n) - \text{rk } Df|_{\mathcal{M}_d(C, A)} \\ &= (d-p)n + (n-d)n - [(n-d) + p + (n-d)](d-p) \\ &= (d-p)(2d-n-p) + (n-d)n . \end{aligned}$$

Now consider the similarity action

$$\sigma : ((S, T), (W, V)) \mapsto (WS, TV)$$

of $\text{Gl}_{d-p}(\mathcal{F}^n) \times \text{Gl}_{n-d}(\mathcal{F}^n)$ on $\mathbb{V}_{d-p}(\mathcal{F}^n) \times \mathbb{V}_{n-d}^\top(\mathcal{F}^n)$. It yields a diffeomorphism $\bar{\sigma}$ of the orbit space $\mathbb{V}_{d-p}(\mathcal{F}^n) \times \mathbb{V}_{n-d}^\top(\mathcal{F}^n) / \text{Gl}_{d-p}(\mathcal{F}^n) \times \text{Gl}_{n-d}(\mathcal{F}^n)$ onto the product of Grassmannians $\mathcal{G}_{d-p}(\mathcal{F}^n) \times \mathcal{G}_d(\mathcal{F}^n)$ given by $\bar{\sigma}(W, V) = (\text{Im } W, \text{Ker } V)$. Apparently our manifold $\mathcal{M}_d(C, A)$ is invariant under σ and $\bar{\sigma}$ maps $\mathcal{M}_d(C, A) / \text{Gl}_{d-p}(\mathcal{F}^n) \times \text{Gl}_{n-d}(\mathcal{F}^n)$ diffeomorphically onto $\mathcal{S}_d(C, A)$. Hence the later is a smooth submanifold of $\mathcal{G}_{d-p}(\mathcal{F}^n) \times \mathcal{G}_d(\mathcal{F}^n)$ of dimension

$$\begin{aligned} \dim \mathcal{S}_d(C, A) &= \dim \mathcal{M}_d(C, A) - (\dim \text{Gl}_{d-p}(\mathcal{F}^n) + \dim \text{Gl}_{n-d}(\mathcal{F}^n)) \\ &= (d-p)(2d-n-p) + (n-d)n - (d-p)^2 - (n-d)^2 \\ &= (d-p)(d-n) + (n-d)d \\ &= (n-d)p , \end{aligned}$$

which completes the proof. \square

The next example shows that the method we used in the last proof does not work in general. Of course, this does not disprove smoothness.

Example 10 Let $p = 2$, $n = 5$ and $\mu_1 = 4$, $\mu_2 = 1$. Let $d = 1$. Consider

$$W_1 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad V_1 = (1 \ 0 \ 0 \ 0 \ 0) .$$

It is $(W_1, V_1) \in \mathcal{M}_d(C, A)$ and $Z_1^*V_1 + Z_2^*C + Z_3^*V_1A = 0$ implies $Z_1^* = 0$ and $Z_2^* = 0$. Hence $W_1Z_1^* + AW_1Z_3^* = 0$ implies $Z_3^* = 0$ and $Df(W_1, V_1)$ is surjective. Now consider

$$W_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad V_2 = (0 \ 0 \ 0 \ 0 \ 1) .$$

Again it is $(W_2, V_2) \in \mathcal{M}_d(C, A)$ but choosing

$$Z_1^* = \begin{pmatrix} 0 \\ 1 \end{pmatrix} , \quad Z_2^* = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad Z_3^* = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

yields $Z_1^*V_2 + Z_2^*C + Z_3^*V_2A = 0$ and $W_2Z_1^* + AW_2Z_3^* = 0$, hence $Df(W_2, V_2)$ is not surjective.

In the single output case we can show that $\text{Inv}_d(C, A)$ is dense in $\mathcal{C}_d(C, A)$. In fact, we have the following result.

Theorem 11 If $p = 1$ then the set $\mathcal{C}_d(C, A)$ is equal to the set of all d -dimensional almost (C, A) -invariant subspaces. It has the structure of the projective space \mathbb{P}^{n-d} with the 'big' $(n - d)$ -dimensional Bruhat cell (which is dense in \mathbb{P}^{n-d}) formed by $\text{Inv}_d(C, A)$.

PROOF. Consider $(\mathcal{W}, \mathcal{V}) \in \mathcal{S}_d(C, A)$. Choose a basis of \mathcal{W} in column echelon form and extend it (by adding one more vector) to a basis of \mathcal{V} . Since

$\mathcal{W} \subset \text{Ker } C$, the basis we get has the following form:

$$\mathcal{V} = \text{Im} \begin{pmatrix} x_1 & * & * & \dots & * & * & * \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ x_j & * & * & \dots & * & * & * \\ 1 & 0 & 0 & \dots & 0 & 0 & * \\ 0 & * & * & \dots & * & * & * \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & * & * & \dots & * & * & * \\ 0 & 1 & 0 & \dots & 0 & 0 & * \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & * & * \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & * & * \\ 0 & 0 & 0 & \dots & 0 & 1 & * \\ 0 & 0 & 0 & \dots & 0 & 0 & * \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & * \end{pmatrix},$$

where at least one of the entries in the last block of the last column (below the 1 in the last but one column) is nonzero, since otherwise the last but one column is not mapped into \mathcal{V} by A , a contradiction to $A\mathcal{W} \subset \mathcal{V}$. But then $A\mathcal{W} \subset \mathcal{V}$ implies that the first column is mapped not only into \mathcal{V} but into \mathcal{W} by A , since, due to the nonzero entry we just mentioned, this column can not show up in a basis representation of the A -image of the first column. But then we can replace the second basis vector of \mathcal{W} by the A -image of the first. Iterating this argument we end up with a basis of the form

$$\mathcal{V} = \text{Im} \begin{pmatrix} x_1 & 0 & \dots & 0 & * \\ \vdots & x_1 & \ddots & \vdots & \vdots \\ x_j & \vdots & \ddots & 0 & * \\ 1 & x_{j-1} & & x_1 & * \\ 0 & 1 & \ddots & \vdots & \vdots \\ 0 & 0 & \ddots & x_{j-1} & * \\ \vdots & \vdots & \ddots & 1 & * \\ 0 & 0 & \dots & 0 & * \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & * \end{pmatrix}.$$

But then $A\mathcal{W} \subset \mathcal{V}$ implies that the last column is the A -image of the last but one column, i.e. has the same entries but shifted down by one. From the format of our basis matrix it follows $j \leq n - d$. Since the columns of every matrix of the described form span a pair $(\mathcal{W}, \mathcal{V}) \in \mathcal{S}_d(C, A)$, it follows that $\mathcal{C}_d(C, A)$ has the structure of \mathbb{P}^{n-d} . It remains to show that $j = n - d$ yields a (C, A) -invariant subspace, while $j < n - d$ yields an almost (C, A) -invariant subspace which is *not* (C, A) -invariant.

If $j = n - d$ then the 1 in the last column is in the last row, i.e. $\mathcal{W} = \mathcal{V} \cap \text{Ker } C$ is spanned by all but the last column, and $A\mathcal{W} \subset \mathcal{V}$ implies $A(\mathcal{V} \cap \text{Ker } C) \subset \mathcal{V}$, i.e. \mathcal{V} is (C, A) -invariant. If $j < n - d$ then the whole last row is zero, i.e. $\mathcal{V} \cap \text{Ker } C = \mathcal{V}$ and $A(\mathcal{V} \cap \text{Ker } C) \not\subset \mathcal{V}$. Hence \mathcal{V} is *not* (C, A) -invariant. On

the other hand we can write

$$\mathcal{V} = \text{Ker} \left(\begin{array}{c} R_n(A_1, b_1) \\ \overline{R}_n(N, b_2) \end{array} \right) = \text{Ker} \left(\begin{array}{cccc} b_1 & A_1 b_1 & \dots & A_1^{n-1} b_1 \\ N^{n-1} b_2 & N^{n-2} b_2 & \dots & b_2 \end{array} \right),$$

where

$$A_1 = \begin{pmatrix} 0 & & -x_1 \\ 1 & 0 & -x_2 \\ & \ddots & \vdots \\ & & 1 & -x_j \end{pmatrix}, \quad b_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad N = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \end{pmatrix} \quad \text{and} \quad b_2 = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

with $b_2 \in \mathcal{F}^{n-d-j}$. Since (A_1, b_1) and (N, b_2) are both reachable and N is nilpotent it follows that \mathcal{V} is almost (C, A) -invariant (cf. Trumpf [15, Proposition 5.13]). \square

References

- [1] G. Basile and G. Marro. Controlled and conditioned invariant subspaces in linear system theory. *J. Optimiz. Th. Applic.*, 3(5):305–315, 1969.
- [2] C.I. Byrnes. Pole assignment by output feedback. In H. Nijmeijer and J.M. Schumacher, editors, *Three decades of mathematical system theory. A collection of surveys at the occasion of the 50th birthday of Jan C. Willems*, Lecture Notes in Control and Information Sciences, 135:31–78, 1989.
- [3] J. Ferrer, F. Puerta and X. Puerta. Differentiable structure of the set of controllable $(A, B)^t$ -invariant subspaces. *Linear Algebra and its Applications*, 275–276:161–177, 1998.
- [4] P.A. Fuhrmann. Duality in polynomial models with some applications to geometric control theory. *IEEE Transactions on Automatic Control*, AC-26:284–295, 1981.
- [5] P.A. Fuhrmann and U. Helmke. A homeomorphism between observable pairs and conditioned invariant subspaces. *Systems and Control Letters*, 30:217–223, 1997.
- [6] P.A. Fuhrmann and U. Helmke. On the parametrization of conditioned invariant subspaces. In F. Colonius, U. Helmke, D. Prätzel-Wolters, and F. Wirth, editors, *Advances in mathematical systems theory. A volume in honor of Diederich Hinrichsen on the occasion of his 60th birthday*, Systems and Control: Foundations and Applications. Birkhäuser, Boston, 2000.
- [7] 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.
- [8] P. Griffiths and J. Harris. *Principles of Algebraic Geometry*. Wiley-Interscience, 1978.

- [9] D. Hinrichsen, H.F. Münzner and D. Prätzel-Wolters. Parametrization of (C,A) -invariant subspaces. *Systems and Control Letters*, 1:192–199, 1981.
- [10] R.A. Horn and C.R. Johnson. *Topics in matrix analysis*. Cambridge University Press, 1991.
- [11] C.M. Özveren, G.C. Verghese and A.S. Willsky. Asymptotic orders of reachability in perturbed linear systems. *IEEE Transactions on Automatic Control*, AC-33(10):915–923, 1988.
- [12] X. Puerta and U. Helmke. The topology of the set of conditioned invariant subspaces. *Systems and Control Letters*, 40:97–105, 2000.
- [13] J. Rosenthal. On dynamic feedback compensation and compactification of systems. *SIAM J. Control Optimization*, 32(1):279–296, 1994.
- [14] J.M. Schumacher. *Dynamic feedback in finite and infinite dimensional linear systems*. Number 143 in Math. Centre Tracts. Amsterdam, 1981.
- [15] J. Trumpf. *On the geometry and parametrization of almost invariant subspaces and observer theory*. PhD thesis, Universität Würzburg, 2002.
- [16] J.C. Willems. Almost $A(\text{mod } B)$ -invariant subspaces. *Astérisque*, 75–76:239–248, 1980.
- [17] J.C. Willems. Almost invariant subspaces: An approach to high gain feedback design — Part I: Almost controlled invariant subspaces. *IEEE Transactions on Automatic Control*, AC-26(1):235–252, 1981.
- [18] J.C. Willems. Almost invariant subspaces: An approach to high gain feedback design — Part II: Almost conditionally invariant subspaces. *IEEE Transactions on Automatic Control*, AC-27(5):1071–1084, 1982.
- [19] W.M. Wonham and A.S. Morse. Decoupling and pole assignment in linear multivariable systems: a geometric approach. *SIAM J. Control*, 8(1):1–18, 1970.
- [20] W.M. Wonham. *Linear multivariable control. A geometric approach*. Springer, 1974.