ELSEVIER Linear Algebra and its Applications 351–352 (2002) 585–599
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

On the geometry of the set of controllability subspaces of a pair (A, B) ☆

F. Puerta*, X. Puerta

ETS Enginyers Industrials UPC, EUPB, Department de Matemàtica Aplicada I, Universitat Politecnica de Catalunya, Diagonal 647, 08028 Barcelona, Spain

Received 20 February 2001; accepted 9 October 2001

Submitted by D. Hinrichsen

Abstract

Given a controllable system defined by a pair of matrices (A, B), we investigate the geometry of the set of controllability subspaces. This set is a subset of the set of (A, B)-invariant subspaces. We prove that, in fact, it is a stratified submanifold and we compute its dimension.

AMS classification: 93 B27; 93 B29

Keywords: (A, B)-invariant; Controllability subspace; Manifold; Orbit space

Introduction

Given a controllable time-invariant multivariable system

$$\dot{x} = Ax + Bu$$

with $A \in K^{n\times n}$ and $B \in K^{m\times n}$ (K denotes the field of real or complex numbers), a subspace $\mathscr{S}$ of K^n is called a *controllability subspace* if $\mathscr{S}$ has the form (following the notation of [11])

$$\mathscr{S} = \langle A + BF|\mathrm{Im}\, BG\rangle$$

☆ Partially supported by the CAICYT, Proyecto de Investigación PB97-0599-CO3-01.
* Corresponding author.
E-mail addresses: puerta@mal.upc.es (F. Puerta), coll@mal.upc.es (X. Puerta).

0024-3795/02/$ - see front matter
PII: S0024-3795(01)00531-6

with $F \in K^{m\times n}$ and $G \in K^{m\times l}$. We recall that $\langle A+BF|\operatorname{Im} BG\rangle = \operatorname{Im} BG + \operatorname{Im}(A+BF)BG + \cdots + \operatorname{Im}(A+BF)^{n-1}BG$. Since $\mathscr{S}$ is $(A+BF)$-invariant, $\mathscr{S}$ is an (A, B)-invariant subspace. Therefore, the set of controllability subspaces is a subset of the set of (A, B)-invariant subspaces.

Let $G_d(K^n)$ denote the Grassmann manifold of d-dimensional linear subspaces of K^n. We define

$$\begin{aligned}\operatorname{Inv}_d(A, B) &= \left\{\mathscr{S} \in G_d(K^n) | \mathscr{S} \text{ is } (A, B)\text{-invariant}\right\},\\ \operatorname{Ctr}_d(A, B) &= \left\{\mathscr{S} \in G_d(K^n) | \mathscr{S} \text{ is a controllability subspace}\right\}.\end{aligned}$$

We have that $\operatorname{Ctr}_d(A, B) \subset \operatorname{Inv}_d(A, B)$.

Controllability and (A, B)-invariant subspaces play an important role in geometric control theory (significant references are [6,10,11]). The geometry of the set $\operatorname{Inv}_d(A, B)$ has been a subject of interest in the last few years (see [2–5,7–9]). However, most of the above references deal with the dual case, that is to say, the set of (C, A)-invariant subspaces. Since the map $\mathscr{S} \mapsto \mathscr{S}^{\perp}$ is a bijection between the set of (A, B) and $(B^{\mathrm{t}}, A^{\mathrm{t}})$-invariant subspaces, the properties of the set of $(B^{\mathrm{t}}, A^{\mathrm{t}})$-invariant subspaces can be transferred in a natural way to the set of (A, B)-invariant subspaces. In particular, from [4,5], where the set of (C, A)-invariant subspaces is stratified by fixing the Brunovsky indices of the restriction of (C, A), one can obtain a stratification of $\operatorname{Inv}_d(A, B)$.

Nevertheless, this stratification has no relation with controllability subspaces, which are the object of our study. Here we introduce a new stratification of $\operatorname{Inv}_d(A, B)$ according to the Brunovsky form of a restriction $(\overline{A}, \overline{B})$ defined directly from the pair (A, B). Since this restriction need not be controllable, the corresponding stratification is not finite, in general. However, it is the suitable restriction when we deal with controllability subspaces. In fact, we prove that $\operatorname{Ctr}_d(A, B)$ is the set of (A, B)-invariant subspaces of $\operatorname{Inv}_d(A, B)$ such that the restriction of (A, B) to each one of them is controllable. Therefore, the introduced stratification of $\operatorname{Inv}_d(A, B)$ induces a finite stratification of $\operatorname{Ctr}_d(A, B)$ defined by

$$\operatorname{Ctr}_d(A, B) = \bigcup_{(\overline{A}, \overline{B}) \text{ controllable}} \operatorname{Inv}_{(\overline{A}, \overline{B})}(A, B) \quad \text{(see Section 2)}.$$

We prove that each stratum $\operatorname{Inv}_{(\overline{A}, \overline{B})}(A, B)$ is a smooth manifold by describing it as an orbit space $\mathscr{M}/\mathscr{G}$, $\mathscr{M}$ being a matrix space and $\mathscr{G}$ a Lie group acting on $\mathscr{M}$. The dimension of $\mathscr{M}/\mathscr{G}$ is obtained by describing the elements of $\mathscr{M}$ and $\mathscr{G}$.

For convenience, we denote a basis $(u_1, \ldots, u_p)$ simply by u if no confussion is possible. Then if f is a linear map, $f(u)$ means the family $(f(u_1), \ldots, f(u_p))$. If u is a set of vectors, $[u]$ means the subspace spanned by u.

We will assume throughout the paper that (A, B) is a controllable pair and B a full rank matrix.

1. On the restriction of (A, B) to an (A, B)-invariant subspace

The concept of restriction of a pair (A, B) to an (A, B)-invariant subspace has been studied in [1]. We introduce in this section a different approach which is more convenient for the applications in Sections 2 and 3. In order to define this restriction we associate to the pair (A, B) the pair formed by the linear map $f: K^{n+m} \to K^n$ defined by $f(x, y) = Ax + By$ and the natural projection map $\pi: K^{n+m} \to K^n$ defined by $\pi(x, y) = x$. Conversely, if for each pair of linear maps (f, π) from an $n + m$-dimensional vector space $\mathscr{K}$ to an n-dimensional vector space $\mathscr{F}$ with π surjective we take a basis of $\mathscr{K}$ of the form (u, v) where v is a basis of the kernel of π and $\pi(u)$ is a basis of $\mathscr{F}$, then the matrix of f with regard to these bases, is a two block matrix (AB) with $A \in K^{n\times n}$ and $B \in K^{n\times m}$. Notice that the matrix of π with regard to the above bases is $(I0)$. We call (A, B) a *matrix representation* of (f, π). In other words, (A, B) is a *matrix representation* of (f, π) if and only if there exist isomorphisms $\phi: \mathscr{K} \to K^{n+m}$ and $\psi: \mathscr{F} \to K^n$ such that $f = \psi^{-1}(AB)\phi$ and $\pi = \psi^{-1}(I0)\phi$.

We have the following proposition:

Proposition 1.1. *Let (f, π) be as above and (A, B) a matrix representation of (f, π). Then a pair (A', B') is a matrix representation of (f, π) if and only if (A', B') is feedback equivalent to (A, B). In particular, there exists a matrix representation of (f, π) in the Brunovsky canonical form.*

Proof. Let (A, B) be the matrix representation of (f, π) with regard to the bases (u, v) and $\pi(u)$ (v is a basis of ker π). Let (u', v') be a basis of $\mathscr{K}$ with v' a basis of ker π. The components of the vectors of (u', v') with regard to (u, v) arranged by columns form a matrix of the form

$$\begin{pmatrix} S & 0 \\ F & T \end{pmatrix}.$$

Notice that the columns of S are the components of the vectors of $\pi(u')$ with regard to $\pi(u)$.

Then (A', B') is the matrix representation of (f, π) with regard to (u', v') and $\pi(u')$ if and only if

$$(AB)\begin{pmatrix} S & 0 \\ F & T \end{pmatrix} = S(A'B'),$$

which is equivalent to $A' = S^{-1}AS + S^{-1}BF$ and $B' = S^{-1}BT$, as we wanted to prove. $\square$

Since a subspace $\mathscr{S} \in K^n$ is (A, B)-invariant if and only if it is (A', B')-invariant for any pair (A', B') feedback equivalent to (A, B), we can define this notion in terms of the pair of linear maps (f, π) associated to (A, B). More precisely, we have the following proposition.

Proposition 1.2. *Let (f, π) be defined by $f(x, y) = Ax + By$ and $\pi(x, y) = x$. Then a subspace $\mathscr{S}$ of K^n is (A, B)-invariant if and only if $\mathscr{S} \subset \pi(f^{-1}(\mathscr{S}))$.*

Proof. $f^{-1}(\mathscr{S}) = \{(x, y) | Ax + By \in \mathscr{S}\}$. Hence, $\pi(f^{-1}(\mathscr{S})) = \{x \in K^n$ such that there exists $y \in K^m$ with $Ax + By \in \mathscr{S}\}$. Therefore, $\mathscr{S} \subset \pi(f^{-1}(\mathscr{S}))$ if and only if $A(\mathscr{S}) \subset \mathscr{S} + \operatorname{Im} B$, that is to say, if and only if $\mathscr{S}$ is an (A, B)-invariant subspace. $\square$

Given a pair of linear maps (f, π) as above, every subspace $\mathscr{S}$ of K^n defines two pairs of linear maps $(\overline{f}, \overline{\pi})$ and $(\tilde{f}, \tilde{\pi})$ which render the following diagram commutative

$$\begin{array}{llcl}
\overline{f}, \overline{\pi}: & \pi^{-1}(\mathscr{S}) \cap f^{-1}(\mathscr{S}) & \to & \mathscr{S} \\
& \cap & & \cap \\
f, \pi: & K^{n+m} & \to & K^n \\
& \downarrow & & \downarrow \\
\tilde{f}, \tilde{\pi}: & K^{n+m}/(\pi^{-1}(\mathscr{S}) \cap f^{-1}(\mathscr{S})) & \to & K^n/\mathscr{S},
\end{array}$$

where the vertical arrows are the natural projections, $\overline{f}$ and $\overline{\pi}$ are the restrictions of f and π to $\pi^{-1}(\mathscr{S}) \cap f^{-1}(\mathscr{S})$ and $\tilde{f}, \tilde{\pi}$, the corresponding maps induced on the quotients. Remark that, while $\tilde{\pi}$ is always surjective, $\overline{\pi}$ does not need to be surjective.

Applying Proposition 1.2, we have that if $\mathscr{S}$ is (A, B)-invariant, then π is surjective (and conversely). Therefore, each (A, B)-invariant subspace defines two pairs of linear maps $(\overline{f}, \overline{\pi})$ and $(\tilde{f}, \tilde{\pi})$ of the same type as (f, π). We call $(\overline{f}, \overline{\pi})$ the *restriction* of (f, π) and $(\tilde{f}, \tilde{\pi})$ the *quotient induced map*. Since the matrix representations of $(\overline{f}, \overline{\pi})$ and $(\tilde{f}, \tilde{\pi})$ are feedback equivalent, respectively, it makes sense to define the Brunovsky indices of $(\overline{f}, \overline{\pi})$ and the Brunovsky indices of $(\overline{f}, \overline{\pi})$ as the Brunovsky indices of any matrix representation (A, B) of these pairs.

Remark 1.3. Let (u, v, w, y) be a basis of K^{n+m} with w a basis of $\ker \pi \cap \pi^{-1}(\mathscr{S}) \cap f^{-1}(\mathscr{S}) = \ker \pi \cap f^{-1}(\mathscr{S})$, (w, y) a basis of $\ker \pi$ and (u, w) a basis of $\pi^{-1}(\mathscr{S}) \cap f^{-1}(\mathscr{S})$. Then $\pi(u)$ is a basis of $\mathscr{S}$, $(\pi(u), \pi(v))$ is a basis of K^n and the matrix representation of (f, π) with respect to these bases is

$$\left(\begin{array}{cc|cc} \overline{A} & X & \overline{B} & Y \\ 0 & \tilde{A} & 0 & \tilde{B} \end{array}\right).$$

It can be easily seen that $(\overline{A}, \overline{B})$ and $(\tilde{A}, \tilde{B})$ are matrix representations of $(\overline{f}, \overline{\pi})$ and $(\tilde{f}, \tilde{\pi})$, respectively.

We now define a decomposition of $\mathrm{Inv}_d(A, B)$ according to the Brunovsky indices of $(\overline{f}, \overline{\pi})$.

Definition 1.4. Let (f, π) be the pair of linear maps defined by $f(x, y) = Ax + By$ and $\pi(x, y) = x$. Let $(\overline{A}, \overline{B})$ with $\overline{A} \in K^{d\times d}$ and $\overline{B} \in K^{d\times r}$ be a pair in Brunovsky canonical form. We define

$$\mathrm{Inv}_{(\overline{A},\overline{B})}(A, B) = \big\{\mathscr{S} \in \mathrm{Inv}_d(A, B) \,|\, (\overline{f}, \overline{\pi}) \text{ has } (\overline{A}, \overline{B}) \text{ as matrix representation}\big\}.$$

We have that

$$\mathrm{Inv}_d(A, B) = \bigcup_{(\overline{A},\overline{B})} \mathrm{Inv}_{(\overline{A},\overline{B})}(A, B).$$

In [1] conditions in order to ensure that $\mathrm{Inv}_{(\overline{A},\overline{B})}(A, B) \neq \emptyset$ are given (see Section 3). We remark that, in contrast with the restriction defined in the dual case, even when (A, B) is controllable, the pair $(\overline{A}, \overline{B})$ need not be so. Consider, for example

$$(A, B) = \left(\begin{array}{cc|cc} \lambda & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right), \quad (\overline{A}, \overline{B}) = (\lambda, 0), \quad (\tilde{A}, \tilde{B}) = (0, 1).$$

(A, B) and $(\tilde{A}, \tilde{B})$ are controllable, but $(\overline{A}, \overline{B})$ is not. Therefore the above union is infinite in general. In the next section, we prove that the controllability of $(\overline{A}, \overline{B})$ characterizes the controllability subspaces.

2. The set of controllability subspaces

Let (f, π) be, as in the previous section, the pair of linear maps defined by $f(x, y) = Ax + By$ and $\pi(x, y) = x$, $\mathscr{S} \in \mathrm{Inv}_d(A, B)$ and $(\overline{f}, \overline{\pi})$ the restriction of (f, π),

$$\overline{f}, \overline{\pi}: \pi^{-1}(\mathscr{S}) \cap f^{-1}(\mathscr{S}) \to \mathscr{S}.$$

Notice that from Proposition 1.1 if a matrix representation of $(\overline{f}, \overline{\pi})$ is a controllable pair of matries, any other matrix representation of $(\overline{f}, \overline{\pi})$ is also controllable. We say then that $(\overline{f}, \overline{\pi})$ is *controllable*. We prove the following theorem.

Theorem 2.1. *With the above notation, $\mathscr{S}$ is a controllability subspace with regard to (A, B) if and only if $(\overline{f}, \overline{\pi})$ is controllable.*

Proof. We recall the notation of Remark 1.3: (u, v, w, y) is a basis of K^{n+m} with w a basis of $\ker \pi \cap f^{-1}(\mathscr{S})$, (w, y) a basis of $\ker \pi = K^m$ and (u, w) a basis of $\pi^{-1}(\mathscr{S}) \cap f^{-1}(\mathscr{S})$. Then $(\pi(u), \pi(v))$ is a basis of K^n and $\pi(u)$ is a basis of $\mathscr{S}$. Let P be the $n \times n$-matrix having the vectors $(\pi(u), \pi(v))$ as columns and Q the

$m \times m$-matrix whose columns are the vectors (w, y). We remark that $P^{-1}(\mathscr{S})$ is the subspace spanned by the first d vectors of the standard basis of K^n.

Assume first that $\mathscr{S}$ is a controllability subspace, that is to say, $\mathscr{S} = \langle A + BF | \mathrm{Im}\, BG \rangle$. Taking into account that $\mathscr{S}$ is $(A + BF)$-invariant, the matrix of $A + BF$ with regard to the basis $(\pi(u), \pi(v))$ has the form

$$\widehat{A} = P^{-1}(A + BF)P = \begin{pmatrix} \overline{A} & X \\ 0 & \tilde{A} \end{pmatrix},$$

where $\overline{A}$ has size $d \times d$. Likewise, the matrix of B with regard to the basis (w, y) and $(\pi(u), \pi(v))$ has the form

$$\widehat{B} = P^{-1}BQ = \begin{pmatrix} \overline{B} & Y \\ 0 & \tilde{B} \end{pmatrix},$$

where $\overline{B}$ has size $d \times l$ with $l = \dim B^{-1}(\mathscr{S})$. We have that

$$(\widehat{A}\widehat{B}) = P^{-1}(AB) \begin{pmatrix} P & 0 \\ FP & Q \end{pmatrix}$$

so that, according to Proposition 1.1 and Remark 1.3, $(\widehat{A}, \widehat{B})$ and $(\overline{A}, \overline{B})$ are matrix representations of (f, π) and $(\overline{f}, \overline{\pi})$, respectively.

We tackle now the proof that $(\overline{A}, \overline{B})$ is a controllable pair. We have that

$$\begin{aligned} P^{-1}(\mathscr{S}) &= P^{-1}\langle A + BF \mid \mathrm{Im}\, BG \rangle \\ &= \langle \widehat{A} \mid \mathrm{Im}\, \widehat{B}G \rangle \\ &= \mathrm{Im}\, (\widehat{B}G | \widehat{A}\widehat{B}G | \cdots | \widehat{A}^{n-1}\widehat{B}G). \end{aligned}$$

Let $G = \binom{\overline{G}}{\tilde{G}}$. We have that

$$\widehat{B}G = \begin{pmatrix} \overline{B}\overline{G} + Y\tilde{G} \\ \tilde{B}\tilde{G} \end{pmatrix}.$$

Since $\mathrm{Im}\, \widehat{B}G \subset P^{-1}\mathscr{S}$, we have that $\tilde{B}\tilde{G} = 0$. But $\widehat{B}$ has full rank and $n - d \geqslant m - l$; hence $\tilde{B}$ is injective and $\tilde{G} = 0$. Therefore,

$$(\widehat{B}G | \widehat{A}\widehat{B}G | \cdots | \widehat{A}^{n-1}\widehat{B}G) = \begin{pmatrix} \overline{B} & \overline{A}\,\overline{B} & \cdots & \overline{A}^{n-1}\overline{B} \\ 0 & 0 & 0 & 0 \end{pmatrix} \mathrm{diag}\,(\overline{G}, \ldots, \overline{G}).$$

Then, $\dim \mathscr{S} = d$ implies that $\mathrm{rank}\, (\overline{B} | \overline{A}\,\overline{B} | \ldots | \overline{A}^{d-1}\overline{B}) = d$, that is to say, the pair $(\overline{A}, \overline{B})$, which is a matrix representation of $(\overline{f}, \overline{\pi})$, is controllable.

Conversely, assume that $(\overline{f}, \overline{\pi})$ is controllable. Since $\mathscr{S}$ is an (A, B)-invariant subspace (π is surjective) there exists a feedback F such that $\mathscr{S}$ is $(A + BF)$-invariant. We consider the matrix representations of (f, π) and $(\overline{f}, \overline{\pi})$ by the matrices $(\widehat{A}, \widehat{B})$ and $(\overline{A}, \overline{B})$ as before. Since $P^{-1}(\mathscr{S})$ is the subspace spanned by the first d elements of the standard basis of K^n, the controllability of $(\overline{A}, \overline{B})$ implies that

$$P^{-1}(\mathscr{S}) = \mathrm{Im} \begin{pmatrix} \overline{B} & \overline{A}\,\overline{B} & \cdots & \overline{A}^{n-1}\overline{B} \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

On the other hand, it is easily checked that if

$$T = \operatorname{diag}\left(\begin{pmatrix} I_l \\ 0 \end{pmatrix}, \overset{n}{\ldots}, \begin{pmatrix} I_l \\ 0 \end{pmatrix}\right)$$

one has that

$$\begin{pmatrix} \overline{B} & \overline{A}\,\overline{B} & \cdots & \overline{A}^{n-1}\overline{B} \\ 0 & 0 & 0 & 0 \end{pmatrix} = (\widehat{B}|\widehat{A}\widehat{B}|\ldots|\widehat{A}^{n-1}\widehat{B})T.$$

Then, if we denote $H = \operatorname{diag}(Q, \overset{n}{\ldots}, Q)$,

$$\mathscr{S} = \operatorname{Im} P(\widehat{B}Q^{-1}|\widehat{A}P^{-1}P\widehat{B}Q^{-1}|\ldots|\widehat{A}^{n-1}P^{-1}P\widehat{B}Q^{-1})HT,$$

and defining $G = Q\binom{I_l}{0}$ we have that $\mathscr{S} = \langle A + BF|\operatorname{Im} BG\rangle$; thus the proof of the theorem is completed. $\square$

Remark 2.2. Theorem 2.1 implies that

$$\operatorname{Ctr}_d(A, B) = \bigcup_{(\overline{A},\overline{B}) \text{ controllable}} \operatorname{Inv}_{(\overline{A},\overline{B})}(A, B),$$

which is a finite union. In the following section, we show that each set $\operatorname{Inv}_{(\overline{A},\overline{B})}(A, B)$ is a smooth manifold.

3. Orbit space structure of the strata

In this section we prove that the decomposition (1) is a finite stratification of $\operatorname{Ctr}_d(A, B)$.

Let (A, B) and $(\overline{A}, \overline{B})$ be controllable pairs where $A \in K^{n\times n}$, $B \in K^{n\times m}$, $\overline{A} \in K^{d\times d}$ and $\overline{B} \in K^{d\times l}$. We assume without loss of generality that (A, B) and $(\overline{A}, \overline{B})$ are in the Brunovsky canonical form. We denote by $k = (k_1, \ldots, k_r)$, $k_1 \geqslant \cdots \geqslant k_r \geqslant 0$ with $k_1 + \cdots + k_r = n$ and $h = (h_1, \ldots, h_s)$, $h_1 \geqslant \cdots \geqslant h_s \geqslant 0$ with $h_1 + \cdots + h_s = d$ the controllability indices of (A, B) and $(\overline{A}, \overline{B})$, respectively. That is to say, $A = \operatorname{diag}\{N_{k_1}, \ldots, N_{k_r}\}$ being N_i the standard upper nilpotent $i \times i$ matrix, $B = \operatorname{diag}\{E_{k_1}, \ldots, E_{k_r}\}$ being $E_i = (0, \ldots, 0, 1)^{\mathrm{t}} \in K^{k_i}$ and analogously, $\overline{A} = \operatorname{diag}\{N_{h_1}, \ldots, N_{h_s}\}$ and $\overline{E} = \operatorname{diag}\{E_{h_1}, \ldots, E_{h_s}\}$.

We also assume that k and h satisfy the following compatibility conditions given in [1] (in order to ensure that $\operatorname{Inv}_{(\overline{A},\overline{B})}(A, B)$ is not empty). Let $r = (r_1, \ldots, r_k)$ and $s = (s_1, \ldots, s_h)$ be the conjugate (or dual) partitions of k and h, respectively. Then the conditions are:

1. $r_i - s_i \leqslant r_1 - s_1$, $i = 1, \ldots, k$, where $s_i = 0$ for $i \geqslant h + 1$.
2. $\sum_{j=1}^{h_p}(r_j - s_j - p) \geqslant 0$, where $h_p = \max\{i|r_i - s_i \geqslant p\}$, $0 \leqslant p \leqslant r_1 - s_1$.

The next results depend on the following lemmas.

Lemma 3.1. *Let $(A, B) \in K^{n\times n} \times K^{n\times m}$ and $(\overline{A}, \overline{B}) \in K^{d\times d} \times K^{d\times l}$ Brunovsky pairs of matrices and $X \in K^{n\times d}$. Then the following conditions are equivalent*:

1. $\begin{cases} (I_n - BB^{\mathrm{t}})X\overline{B} &= 0, \\ (I_n - BB^{\mathrm{t}})X\overline{A} &= AX. \end{cases}$
2. *There exist matrices Y and Z such that*

$$(AB)\begin{pmatrix} X & 0 \\ Y & Z \end{pmatrix} = X(\overline{A}\,\overline{B}).$$

Besides, if the above conditions hold, we have that $Y = B^{\mathrm{t}}X\overline{A}$ and $Z = B^{\mathrm{t}}X\overline{B}$.

Proof. Condition 2 implies that

$$\begin{cases} BZ &= X\overline{B}, \\ BY &= X\overline{A} - AX. \end{cases}$$

Now notice that $B^{\mathrm{t}}B = I$ and $B^{\mathrm{t}}A = 0$ because (A, B) is in the Brunovsky canonical form. Then multiplying the above equalities by B^{t} on the left we obtain

$$Z = B^{\mathrm{t}}X\overline{B} \quad \text{and} \quad Y = B^{\mathrm{t}}X\overline{A}.$$

Replacing Y and Z in the above equalities, condition 1 follows immediately.

Conversely, taking $Z = B^{\mathrm{t}}X\overline{B}$ and $Y = B^{\mathrm{t}}X\overline{A}$ it is clear that (1) implies (2). □

Lemma 3.2. *With the notation of the previous lemma, a matrix X is a solution of Lemma* 3.1(1) *if and only if $X = (X_{ij})_{1\leqslant i\leqslant r, 1\leqslant j\leqslant s}$ with*

$$X_{i,j} = \begin{pmatrix} x_{i,j}^1 & \cdots & x_{i,j}^{h_j-k_i+1} & 0 & 0 & \cdots & 0 \\ 0 & x_{i,j}^1 & \cdots & x_{i,j}^{h_j-k_i+1} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & x_{i,j}^1 & \cdots & x_{i,j}^{h_j-k_i+1} \end{pmatrix}$$

if $k_i \leqslant h_j$ or 0 otherwise.

Proof. We partition X into blocks $X_{i,j}$ of size $k_i \times h_j$, $1 \leqslant i \leqslant r$, $1 \leqslant j \leqslant s$. An easy computation shows that $(I_n - BB^{\mathrm{t}})X\overline{B} = 0$ is equivalent to

$$J_i X_{i,j} \overline{E}_j = 0,$$

$1 \leqslant i \leqslant r$, $1 \leqslant j \leqslant s$, where

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

So, if $X_{i,j} = (x_{p.q})_{1\leqslant p\leqslant k_i, 1\leqslant q\leqslant h_j}$ we conclude from the above equality that $x_{1,h_j} = \cdots = x_{k_i-1,h_j} = 0$, that is to say, $X_{i,j}$ has the form

$$X_{i,j} = \begin{pmatrix} * & \cdots & * & 0 \\ \cdots & \cdots & \cdots & \cdots \\ * & \cdots & * & 0 \\ * & \cdots & * & * \end{pmatrix}.$$

Proceeding analogously, we see that $(I_n - BB^{\mathrm{t}})X\overline{A} = AX$ is equivalent to

$$J_i X_{i,j} \overline{N}_j = N_i X_{i,j},$$

$1 \leqslant i \leqslant r$, $1 \leqslant j \leqslant s$, so that $X_{i,j}$ satisfies

$$\begin{pmatrix} x_{2,1} & \cdots & x_{2,h_j} \\ \cdots & \cdots & \cdots \\ x_{k_i,1} & \cdots & x_{k_i,h_j} \\ 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} 0 & x_{1,1} & \cdots & x_{1,h_j-1} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & x_{k_i-1,1} & \cdots & x_{k_i-1,h_j-1} \\ 0 & \cdots & \cdots & 0 \end{pmatrix},$$

where, according to (1), $x_{2,h_j} = \cdots = x_{k_i,h_j} = x_{1,h_j-1} = \cdots = x_{k_i-1,h_j-1} = 0$. Now the lemma follows easily. □

Corollary 3.3. *With the above notation, if X has full rank, then $B^{\mathrm{t}}X\overline{B}$ has full rank, too.*

Proof. Note that $B^{\mathrm{t}}X\overline{B}$ is the matrix whose entries are the right-bottom corner of the blocks $X_{i,j}$ of X. Since the entries of the last column of $X_{i,j}$ are all 0 with the possible exception of the last one, we have that if X has full rank, then $B^{\mathrm{t}}X\overline{B}$ has also full rank. □

The following theorem characterizes the elements of $\mathrm{Inv}_{(\overline{A},\overline{B})}(A, B)$.

Theorem 3.4. *$\mathscr{S} \in \mathrm{Inv}_{(\overline{A},\overline{B})}(A, B)$ if and only if $\mathscr{S} = \mathrm{Im}\, X$, where X is a full rank matrix satisfying*

$$\begin{cases} (I_n - BB^{\mathrm{t}})X\overline{B} = 0, \\ (I_n - BB^{\mathrm{t}})X\overline{A} = AX. \end{cases}$$

Proof. We have that $\mathscr{S} \in \mathrm{Inv}_{(\overline{A},\overline{B})}(A, B)$ if and only if there exist matrices X, Y, Z with $\mathscr{S} = \mathrm{Im}\, X$ making the following diagram commutative:

$$\begin{array}{ccc} K^{d+l} & \overset{(\overline{A},\overline{B})}{\rightarrow} & K^{d} \\ \downarrow \begin{pmatrix} X & 0 \\ Y & Z \end{pmatrix} & & X \downarrow, \\ K^{n+m} & \overset{(\overline{A},\overline{B})}{\rightarrow} & K^{n} \end{array}$$

where

$$X \quad \text{and} \quad \begin{pmatrix} X & 0 \\ Y & Z \end{pmatrix}$$

have full rank. Notice that

$$\operatorname{Im}\begin{pmatrix} X & 0 \\ Y & Z \end{pmatrix} = \pi^{-1}(\mathscr{S}) \cap f^{-1}(\mathscr{S}).$$

Then, applying Lemma 3.1 and Corollary 3.3 the theorem follows. □

The above theorem leads to the following definition.

Definition 3.5. Given $(A, B) \in K^{n\times n} \times K^{n\times m}$ and $(\overline{A}, \overline{B}) \in K^{d\times d} \times K^{d\times l}$ Brunovsky pairs of matrices, we define

$$\mathscr{M}_{(\overline{A},\overline{B})}(A, B) = \big\{X \in K^{n\times d} | (I_n - BB^{\mathrm{t}})X\overline{B} = 0 \text{ and } (I_n - BB^{\mathrm{t}})X\overline{A} = AX \\ \text{where } X \text{ has full rank}\big\}$$
$$\mathscr{G}_{(\overline{A},\overline{B})} = \mathscr{M}_{(\overline{A},\overline{B})}(\overline{A}, \overline{B})$$

If no confusion is possible, we denote $\mathscr{M}_{(\overline{A},\overline{B})}(A, B)$ and $\mathscr{G}_{(\overline{A},\overline{B})}$ by $\mathscr{M}$ and $\mathscr{G}$, respectively.

We remark that $\mathscr{M}$ is a submanifold of $K^{n\times m}$. In fact, it is an open subset of a linear subvariety of $K^{n\times m}$. We have the following proposition.

Proposition 3.6. *With the above notation we have*:

1. $\mathscr{G}$ *is a Lie subgroup of* $Gl(d)$ *acting freely on* $\mathscr{M}$ *on the right by matrix multiplication.*
2. *The orbit space* $\mathscr{M}/\mathscr{G}$ *has a differentiable structure such that the projection* $\pi : \mathscr{M} \to \mathscr{M}/\mathscr{G}$ *is a submersion.*
3. $\dim \mathscr{M}/\mathscr{G} = \dim \mathscr{M} - \dim \mathscr{G}$

Proof. Let $X, X' \in \mathscr{G}$. According to Lemma 3.1 this is equivalent to the existence of Y, Z and Y', Z' such that

$$(\overline{A}\,\overline{B})\begin{pmatrix} X & 0 \\ Y & Z \end{pmatrix} = X(\overline{A}\,\overline{B}), \qquad (\overline{A}\,\overline{B})\begin{pmatrix} X' & 0 \\ Y' & Z' \end{pmatrix} = X'(\overline{A}\,\overline{B})$$

and $X, X', \overline{B}^{\mathrm{t}}X\overline{B} = Z, \overline{B}^{\mathrm{t}}X'\overline{B} = Z'$ have full rank.

Since

$$\begin{pmatrix} X & 0 \\ Y & Z \end{pmatrix}\begin{pmatrix} X' & 0 \\ Y' & Z' \end{pmatrix} = \begin{pmatrix} XX' & 0 \\ YX' + ZY' & ZZ' \end{pmatrix}$$

has full rank and verifies

$$(\overline{A}\,\overline{B})\begin{pmatrix} XX' & 0 \\ YX'+ZY' & ZZ' \end{pmatrix} = XX'(\overline{A}\,\overline{B}),$$

it follows that $XX' \in \mathscr{G}$. By a similar reasoning we conclude that if $X \in \mathscr{G}$, then $X^{-1} \in \mathscr{G}$. Obviously, $I_d \in \mathscr{G}$. Hence, $\mathscr{G}$ is a Lie subgroup of $Gl(d)$.

Let $T \in \mathscr{G}$ and $X \in \mathscr{M}$. The proof that $XT \in \mathscr{M}$ follows the same pattern as above and is left to the reader. Finally if $XT = X$ with $X \in \mathscr{M}$ and $T \in \mathscr{G}$, then since X has full rank, $T = I_d$. So, (1) is true.

The proofs of (2) and (3) are the same as those given in Theorem 4.5 of [4]. □

We now state the main result of this section.

Theorem 3.7. *The map* $\phi\colon \mathscr{M} \to Gr_d(K^n)$ *defined by* $X \mapsto \operatorname{Im} X$ *induces a submersion* $\tilde{\phi} : \mathscr{M}/\mathscr{G} \to Gr_d(K^n)$ *having image* $\operatorname{Inv}_{(\overline{A},\overline{B})}(A, B)$. *Moreover, with the differentiable structure induced by* ϕ, $\operatorname{Inv}_{(\overline{A},\overline{B})}(A, B)$ *is a submanifold of* $Gr_d(K^n)$ *of dimension* $\dim \mathscr{M} - \dim \mathscr{G}$.

Proof. The image of ϕ is $\operatorname{Inv}_{(\overline{A},\overline{B})}(A, B)$ according to Theorem 3.4. Moreover, $\tilde{\phi}$ is well defined because $\operatorname{Im} X = \operatorname{Im} XG$ for any matrix G of $\mathscr{G}$. For the injectivity of $\tilde{\phi}$ applying Lemma 3.1, we have that if $X \in \mathscr{M}$ and $XG \in \mathscr{M}$ with $G \in Gl(d)$, then $G \in \mathscr{G}$. The rest of the proof is similar to that of [5]. □

Notice that from the above theorem we have that

$$\operatorname{Inv}_{(\overline{A},\overline{B})}(A, B) \cong \mathscr{M}_{(\overline{A},\overline{B})}(A, B)/\mathscr{G}_{(\overline{A},\overline{B})}$$

In Lemma 3.2 the elements of $\mathscr{M}_{(\overline{A},\overline{B})}(A, B)$ and $\mathscr{G}_{(\overline{A},\overline{B})}$ are described in terms of h and k (with the additional full rank condition). Hence, taking into account that the number of free parameters in $X_{i,j}$ is $h_j - k_i + 1$ if $k_i \leqslant h_j$ and 0 otherwise, we have the following proposition.

Proposition 3.8.

$$\begin{aligned}\dim \operatorname{Inv}_{(\overline{A},\overline{B})}(A, B) = &\sum_{1\leqslant j\leqslant s, 1\leqslant i\leqslant r} \sup\{h_j - k_i + 1, 0\} \\ &- \sum_{1\leqslant j,i\leqslant s} \sup\{h_j - h_i + 1, 0\}\end{aligned}$$

Example 3.9. Let $k = (4, 3, 3, 1, 1)$ and $h = (3, 3, 1)$. Then the strata of controllability subspaces $\operatorname{Inv}_{(\overline{A},\overline{B})}(A, B)$ are represented by the matrices

$$\left(\begin{array}{ccc|ccc|c}
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
x_1 & 0 & 0 & x_9 & 0 & 0 & 0 \\
0 & x_1 & 0 & 0 & x_9 & 0 & 0 \\
0 & 0 & x_1 & 0 & 0 & x_9 & 0 \\
\hline
x_2 & 0 & 0 & x_{10} & 0 & 0 & 0 \\
0 & x_2 & 0 & 0 & x_{10} & 0 & 0 \\
0 & 0 & x_2 & 0 & 0 & x_{10} & 0 \\
\hline
x_3 & x_4 & x_5 & x_{11} & x_{12} & x_{13} & x_{17} \\
\hline
x_6 & x_7 & x_8 & x_{14} & x_{15} & x_{16} & x_{18}
\end{array}\right)$$

and $\dim \operatorname{Inv}_{(\overline{A},\overline{B})}(A, B) = 7$.

In a similar way to [9] one has another matrix representation of a controllability subspace in terms of the conjugate partitions of k and h obtained by reordering the bases of K^n and $\mathscr{S}$. As in [9], this representation is more suitable to obtain a canonical representative of each controllability subspace (see Example 3.11). Besides, a more compact formula for the dimension of $\operatorname{Inv}_{(\overline{A},\overline{B})}(A, B)$ can be obtained.

We index the usual basis v of K^n by $(v_{1,1}, \dots, v_{1,k_1}, \dots, v_{r,1}, \dots, v_{r,k_r})$ (we recall that (A, B) is in the Brunovsky canonical from and k_i are the size of the nilpotent blocks). Given a matrix $X \in \mathscr{M}_{(\overline{A},\overline{B})}(A, B)$, its columns form a basis of the subspace $\mathscr{S} = \operatorname{Im} X \in \operatorname{Inv}_{(\overline{A},\overline{B})}(A, B)$. Let $u = (u_{1,1}, \dots, u_{1,h_1}, \dots, u_{s,1}, \dots u_{s,h_s})$ be this basis. We reorder the bases of K^n and $\mathscr{S}$ in the following way (see Example 3.11). If r and s are the dual partitions of k and h, respectively,

$$v \mapsto \overline{v} = (v_{1,k_1}, \dots, v_{r_1,k_r}, \dots, v_{1,k_1-i}, \dots, v_{r_{i+1},k_r-i}, \dots, v_{1,1}, \dots, v_{r_k,1})$$
$$u \mapsto \overline{u} = (u_{1,h_1}, \dots, u_{s_1,h_s}, \dots, u_{1,h_1-i}, \dots, u_{s_{i+1},h_s-i}, \dots, u_{1,1}, \dots, u_{s_h,1}).$$

We denote by Y the matrix whose columns are the components of $\overline{u}$ with regard to $\overline{v}$. We have the following proposition, whose proof is similar to that of (2.5) [9].

Proposition 3.10. *Y can be partioned into blocks $Y_{i,j} \in K^{r_i \times s_j}$, $1 \leqslant i \leqslant k$, $1 \leqslant j \leqslant h$ with*
1. *$Y_{i,j} = 0$ if $i > j$*
2. *$Y_{1,j} = \left(Z^j_\alpha\right)_{1 \leqslant \alpha \leqslant h-j+1}$, where Z^j_α is a $r_1 \times (s_{h-\alpha+1} - s_{h-\alpha+2})$-matrix with the first $r_{h-j-\alpha+3}$ rows zero $(1 \leqslant j \leqslant h)$.*
3. *$Y_{i+1,j+1}$ is obtained from $Y_{i,j}$ by removing the last $r_i - r_{i+1}$ rows and the last $s_j - s_{j+1}$ columns.*

We denote by $\mathscr{M}^*(r, s)$ the matrices Y described in Proposition 3.10 and $\mathscr{G}^*(s) = \mathscr{M}^*(s, s)$. We have that the map $Y \mapsto VYU^{-1}$ with U and V being the permutation matrices corresponding to the rearrangement of the bases that we have considered is a bijection between $\mathscr{M}^*(r, s)$ and $\mathscr{M}(k, h)$. Moreover, we can easily see that $\mathscr{G}^*(s)$ acts on $\mathscr{M}^*(r, s)$, and we have a similar proposition to Proposition 3.6 for $\mathscr{M}^*(r, s)/\mathscr{G}^*(s)$ so that the map $Y \mapsto VYU^{-1}$ induces a diffeomorphism

$$\mathscr{M}^*(r, s)/\mathscr{G}^*(s) \cong \mathscr{M}(k, h)/\mathscr{G}(h).$$

Example 3.11. As in Example 3.9, consider $k = (4, 3, 3, 1, 1)$ and $h = (3, 3, 1)$, so that $r = (5, 3, 3, 1)$ and $s = (3, 2, 2)$. We arrange the basis of K^{12} and $\mathscr{S}(\cong K^7)$ in the following way:

$$\begin{array}{l} v_{1,1}, v_{1,2}, v_{1,3}, v_{1,4} \\ v_{2,1}, v_{2,2}, v_{2,3}, \\ v_{3,1}, v_{3,2}, v_{3,3}, \\ v_{4,1}, \\ v_{5,1} \end{array} \quad \mapsto \quad \begin{array}{l} v_{1,4}, v_{2,3}, v_{3,3}, v_{4,4}, v_{5,1}, \\ v_{1,3}, v_{2,2}, v_{3,2}, \\ v_{1,2}, v_{2,1}, v_{3,1}, \\ v_{1,1} \end{array}$$

$$\begin{array}{l} u_{1,1}, u_{1,2}, u_{1,3}, \\ u_{2,1}, u_{2,2}, u_{2,3}, \\ v_{3,1} \end{array} \quad \mapsto \quad \begin{array}{l} u_{1,3}, u_{2,3}, u_{3,1}, \\ u_{1,2}, u_{2,2}, \\ u_{1,1}, u_{2,1} \end{array}$$

Then, with regard to the new bases, the matrix of Example 3.9 is

$$Y = \left(\begin{array}{ccc|cc|cc} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ x_1 & x_9 & 0 & 0 & 0 & 0 & 0 \\ x_2 & x_{10} & 0 & 0 & 0 & 0 & 0 \\ x_5 & x_{13} & x_{17} & x_4 & x_{12} & x_3 & x_{11} \\ x_8 & x_{16} & x_{18} & x_7 & x_{15} & x_6 & x_{14} \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & x_9 & 0 & 0 \\ 0 & 0 & 0 & x_2 & x_{10} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & x_1 & x_9 \\ 0 & 0 & 0 & 0 & 0 & x_2 & x_{10} \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Remark 3.12. Notice that in this representation, all the parameters are in the upper blocks of Y and the remaining blocks are obtained from the previous ones by removing rows and columns according to Proposition 3.10(3). Moreover, $Y_{1,1} = B^{\mathrm{t}} X \overline{B}$, so $Y_{1,1}$ has full rank (see Proposition 3.2 and Corollary 3.3). Due to this and to the form of the blocks of Y, one can eliminate parameters by making left elementary transformations in a similar way to [9]. The next example shows how a canoncial representative of the controllability subspace representated by Y can be obtained, as in [9].

Example 3.13. We consider the matrix representation of the controllability subspaces of the previous example. Then making linear combinations of its columns one can easily check that we can reduce this matrix to the following matrix:

$$\left(\begin{array}{ccc|cc|cc}
0 & 0 & 0 & 0 & 0 & 0 & 0\\
1 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 1 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 1 & 0 & 0 & 0 & 0\\
x & y & z & t & u & v & w\\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 1 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 1 & 0 & 0\\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 1 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 1\\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right),$$

which is of the form given in Proposition 3.10. We remark that the number of parameters of the above matrix coincide with the dimension of $\mathrm{Inv}_{(\overline{A},\overline{B})}(A, B)$ and, in fact, we can show that it parametrizes an open dense set of $\mathrm{Inv}_{(\overline{A},\overline{B})}(A, B)$ in a similar way of [9].

We end this section with a more compact formula for the dimension of $\mathrm{Inv}_{(\overline{A},\overline{B})}(A, B)$ in terms of the conjugate partitions of k and h, r and s. This formula is obtained by counting the parameters of $\mathscr{M}^*(r, s)$ and $\mathscr{G}^*(s)$.

Corollary 3.14. *With the above notation, we have*

$$\dim \mathrm{Inv}_{(\overline{A},\overline{B})}(A, B) = \sum_{i=1}^{h} s_i((r_1 - s_1) - (r_{i+1} - s_{i+1})).$$

Acknowledgement

We are very grateful to the referees for their suggestions and comments.

References

[1] I. Baragaña, I. Zaballa, Block similarity of restrictions to (A, B)-invariant, Subspaces Linear Algebra Appl. 220 (1995) 31–62.

[2] P.A. Fuhrmann, U. Helmke, A homeomorphism between observable pairs and (A, B)-invariant subspaces, Systems Control Lett. 30 (1997) 217–223.

[3] P.A. Fuhrmann, U. Helmke, On the parametrization of (A, B)-invariant subspaces, Linear Algebra Appl. (to appear).

[4] J. Ferrer, F. Puerta, X. Puerta, Differentiable structure of the set of controllable $(A, B)^{\mathrm{t}}$-invariant subspaces, Linear Algebra Appl. 275–276 (1998) 161–177.

[5] J. Ferrer, F. Puerta, X. Puerta, Stratification of the set of general (A, B)-invariant subspaces (preprint).
[6] M. Heymann, Controllability subspaces and feedback simulation, SIAM J. Control Optim. 14 (1976) 769–789.
[7] D. Hinrichsen, H.F. Münzer, D. Prätzel-Wolters, Parametrization of (C, A)-invariant subspaces, Systems Control Lett. 1 (1981) 192–199.
[8] X. Puerta, U. Helmke, The topology of the set of conditioned-invariant subspaces, Systems Control Lett. 40 (2000) 97–105.
[9] F. Puerta, X. Puerta, I. Zaballa, A coordinate atlas of the manifold of observable conditioned-invariant subspaces, Internat. J. Control (to appear).
[10] M.E. Warren, A.E. Eckberg, On the dimension of controllability subspaces: a characterization via polynomial matrices and Kronecker invariants, SIAM J. Control Optim. 13 (1975) 434–445.
[11] W.M. Wohham, Linear Multivariable Control Theory: A Geometric Approach, Prentice-Hall, New Jercy, 1974.