

A GEOMETRIC APPROACH TO THE CARLSON PROBLEM*

ALBERT COMPTA[†] AND JOSEP FERRER[†]

Abstract. The possible observability indices of an observable pair of matrices, when supplementary subpairs are prescribed, are characterized when the “quotient” one is nilpotent. The geometric techniques used are also valid in the classical Carlson problem for square matrices.

Key words. Carlson problem, supplementary pairs of matrices, (C, A) -invariant subspaces, block similarity, Littlewood–Richardson sequences, Brunovsky–Kronecker reduced form

AMS subject classifications. 15A04, 15A21, 93B10, 93B27

PII. S0895479898349148

1. Introduction. This work is a contribution to the problem analogous to the Carlson problem, but involves pairs of matrices instead of single square matrices. In addition, it should be emphasized that the geometric techniques used can also be applied to construct explicit solutions (see section 7) and to study the classical Carlson problem (section 2).

Because of our geometric approach, it is convenient to deal with vertical pairs of matrices, corresponding to linear maps defined on a subspace (see [5]). The dual case of horizontal pairs of matrices, corresponding to maps defined modulo a subspace, is more appropriate to matricial techniques (as in [3]).

So pairs of matrices $P = \begin{pmatrix} A \\ C \end{pmatrix}$, where $A : \mathbb{C}^n \rightarrow \mathbb{C}^n$, $C : \mathbb{C}^n \rightarrow \mathbb{C}^m$ ($m \leq n$), are considered with the following equivalence relation (named “block-similarity” in [8] or “equivalence” in [11]), which generalizes the usual similarity between square matrices: P and P' are block-similar if

$$P' \equiv \begin{pmatrix} A' \\ C' \end{pmatrix} = \begin{pmatrix} Q & S \\ 0 & T \end{pmatrix} \begin{pmatrix} A \\ C \end{pmatrix} Q^{-1}$$

or, equivalently,

$$A' = Q(A + FC)Q^{-1}, \quad C' = TCQ^{-1},$$

where Q and T are nonsingular, and $F = Q^{-1}S$. Throughout the paper, the letters BK (from Brunovsky–Kronecker) will denote the invariants, reduced canonical form, etc., relative to this equivalence relation (see, for example, [8, pp. 96–209] or [5, p. 52]).

With this notation, the general Carlson problem for pairs of matrices can be formulated as follows: characterization of the possible BK-invariants of the pair

$$P = \begin{pmatrix} A \\ C \end{pmatrix} = \begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \\ C_1 & C_3 \\ 0 & C_2 \end{pmatrix}$$

*Received by the editors December 11, 1998; accepted for publication (in revised form) by D. Boley January 20, 2000; published electronically June 20, 2000. This work was partially supported by DGICYT grant PB94-1365-C03-03. This work was partially presented in the ILAS Conference, Madison, WI, 1998 and in the SSC'98 Conference, Nantes, France, 1998.

<http://www.siam.org/journals/simax/22-1/34914.html>

[†]Departament de Matemàtica Aplicada I, Universitat Politècnica de Catalunya, Diagonal 647, 08028 Barcelona, Spain (compta@ma1.upc.es, ferrer@ma1.upc.es).

when A_3, C_3 vary, if the pairs (or equivalently its BK-invariants) $P_1 = \begin{pmatrix} A_1 \\ C_1 \end{pmatrix}$ and $P_2 = \begin{pmatrix} A_2 \\ C_2 \end{pmatrix}$ are fixed.

In system theory, this problem arises in a natural way, for example, when two systems are composed in a “simple cascade” (see [8], [1]).

Baragaña and Zaballa [2] characterize these possible BK-invariants for the particular case when P_2 is observable. Here, the “supplementary” (see Remark 4.7) particular case is considered, when P_2 is an endomorphism (i.e., $C_2 = 0$). When it has a single eigenvalue, Theorem 3.1 gives implicit and explicit characterizations of the possible BK-indices of P , the former being in some sense analogous to the existence of Littlewood–Richardson sequences for the classical Carlson problem. The proof of these implicit characterizations is constructive, so that some examples of explicit solutions P are included in the last section.

Section 2 contains a geometric approach to the classical Carlson problem, which is taken as a motivation of the techniques used in this paper. Section 3 contains the precise definitions and statement of the main theorem (Theorem 3.1), whose proof is delayed until section 5 (necessity) and section 6 (sufficiency), after a geometric reformulation of the problem (Corollary 4.5) in section 4. Some examples are presented in section 7.

In this paper, X will be a finite-dimensional vector space over the complex numbers \mathbb{C} , and Y, W, \dots will denote vector subspaces of X . If $B \subset X$ is a subset, $[B]$ will be the subspace spanned by the vectors in B . A basis B of X will be called *adapted* to the subspaces Y, W, \dots if $B \cap Y, B \cap W, \dots$ are respective bases.

$\mathbb{C}^{p \times q}$ means the set of complex matrices having p rows and q columns. $\mathbb{C}^{p \times q} \times \mathbb{C}^{p' \times q'}$ means the set of vertical pairs of matrices, the one at the top being of $\mathbb{C}^{p \times q}$ and the one at the bottom of $\mathbb{C}^{p' \times q'}$.

In the paper, a *partition*

$$a = (a_1, a_2, \dots, a_{\ell(a)}, 0, \dots, 0)$$

will be a finite nonincreasing sequence of nonnegative integers

$$a_1 \geq a_2 \geq \dots \geq a_{\ell(a)} > 0,$$

where $\ell(a)$ is called its *length*. We note $|a| = a_1 + a_2 + \dots + a_{\ell(a)}$ (named its *weight*).

Its *conjugate* partition (see [7, p. 54]) $a^* = (a_1^*, a_2^*, \dots)$ is defined by means of

$$a_j^* = \#\{1 \leq i \leq \ell(a) : a_i \geq j\},$$

where the symbol $\#$ means “cardinal.” Notice that $a_1^* = \ell(a)$, $\ell(a^*) = a_1$, $|a^*| = |a|$, $(a^*)^* = a$.

Given two partitions a and b , symbol $a \prec b$ means $|a| = |b|$ and

$$a_1 + \dots + a_i \leq b_1 + \dots + b_i \quad (i \geq 1).$$

The Segre characteristic relative to any square matrix eigenvalue is the partition of the sizes of his Jordan blocks.

2. A geometric approach to the classical Carlson problem. Let us see how the geometric tools used in this paper arise in a natural way in the classical Carlson problem concerning square matrices. We recall the key theorem is due to Klein [9], relating the decomposition of p -modules with the existence of so-called LR-sequences. On the other hand, [6] proves the equivalence between the Carlson problem

and the one of invariant factors of the product of polynomial matrices, which in turn is related by [10] with the decomposition of p -modules. To summarize, we have the following well known result which reduces the Carlson problem to the existence of LR-sequences.

THEOREM 2.1. *Let there be three partitions*

$$\begin{aligned} \omega &= (\omega_1, \omega_2, \dots), & |\omega| &= n, \\ w &= (w_1, w_2, \dots), & |w| &= d, \\ b &= (b_1, b_2, \dots), & |b| &= n - d. \end{aligned}$$

The following conditions are equivalent:

(I) *For any nilpotent matrices $A_1 \in \mathbb{C}^{d \times d}$ and $A_2 \in \mathbb{C}^{(n-d) \times (n-d)}$ having Segre characteristic w^* and b^* , respectively, there is a matrix $Z \in \mathbb{C}^{d \times (n-d)}$ such that the matrix*

$$(2.1) \quad A = \begin{pmatrix} A_1 & Z \\ 0 & A_2 \end{pmatrix} \in \mathbb{C}^{n \times n}$$

has Segre characteristic ω^* .

(II) *There is a finite sequence of partitions (named after Littlewood–Richardson) w^0, w^1, \dots, w^s ($s = \ell(b)$) such that $w^0 = w$, $w^s = \omega$, and, for all $i, j \geq 1$,*

- (a) $|w^j| - |w^{j-1}| = b_j$,
- (b) $w_1^{j+1} = w_1^j$; $w_i^j \geq w_i^{j-1} \geq w_{i+1}^j$,
- (c) $\sum_{\ell \leq i+1} (w_\ell^{j+1} - w_\ell^j) \leq \sum_{\ell \leq i} (w_\ell^j - w_\ell^{j-1})$.

From a geometric point of view, let us consider an endomorphism $f : X \rightarrow X$, and $W \subset X$ an invariant subspace (i.e., $f(W) \subset W$). Then, in any basis B of X adapted to W , the matrix of f has the form (2.1) above, where A_1 and A_2 are the matrices of the natural endomorphisms

$$\begin{aligned} \widehat{f} &: W \rightarrow W, \\ \widetilde{f} &: \frac{X}{W} \rightarrow \frac{X}{W}, \end{aligned}$$

respectively, in the bases induced by B in a natural way.

If condition (I) holds, let us consider the subspaces

$$W_i^j = \text{Ker } f^i \cap f^{-j}(W) \quad (i, j \geq 0),$$

which can be organized in the following diagram:

$$\begin{array}{ccccccccc} W & \subset & f^{-1}(W) & \subset & f^{-2}(W) & \subset \cdots \subset & f^{-s}(W) & = & X \\ \parallel & & \parallel & & \parallel & & \parallel & & \parallel \\ \text{Ker } \widehat{f}^n & \subset & W_n^1 & \subset & W_n^2 & \subset \cdots \subset & W_n^s & = & \text{Ker } f^n \\ \cup & & \cup & & \cup & & \cup & & \cup \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \cup & & \cup & & \cup & & \cup & & \cup \\ \text{Ker } \widehat{f}^i & \subset & W_i^1 & \subset & W_i^2 & \subset \cdots \subset & W_i^s & = & \text{Ker } f^i \\ \cup & & \cup & & \cup & & \cup & & \cup \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \cup & & \cup & & \cup & & \cup & & \cup \\ \text{Ker } \widehat{f} & \subset & W_1^1 & = & W_1^2 & = \cdots = & W_1^s & = & \text{Ker } f \end{array},$$

where $s = \ell(b)$.

Notice that $W_0^j = 0$ for all $j \geq 0$ and

$$\begin{aligned} \omega_i &= \dim \text{Ker } f^i - \dim \text{Ker } f^{i-1}, \\ w_i &= \dim \text{Ker } \widehat{f}^i - \dim \text{Ker } \widehat{f}^{i-1}, \\ b_j &= \dim f^{-j}(W) - \dim f^{-j+1}(W). \end{aligned}$$

Then, it can be proved that condition (II) holds by taking

$$w_i^j = \dim W_i^j - \dim W_{i-1}^j.$$

In fact, condition (II)(a) is trivial, and the other ones are equivalent to the injectivity of the maps

$$\frac{W_{i+1}^j}{W_i^j} \longrightarrow \frac{W_i^{j-1}}{W_{i-1}^{j-1}}, \quad \frac{W_i^{j+1}}{W_i^j} \longrightarrow \frac{W_{i-1}^j}{W_{i-1}^{j-1}}$$

induced by f .

3. Precise definitions and statement of the main theorem. As a natural generalization of the Carlson problem, let us consider pairs of matrices of the form

$$\begin{aligned} P &= \begin{pmatrix} A \\ C \end{pmatrix} = \begin{pmatrix} A_1 & Z \\ 0 & A_2 \\ C_1 & C_3 \\ 0 & C_2 \end{pmatrix}, \\ P_1 &\equiv \begin{pmatrix} A_1 \\ C_1 \end{pmatrix}, \quad P_2 \equiv \begin{pmatrix} A_2 \\ C_2 \end{pmatrix}, \end{aligned}$$

where A_1 and A_2 are square matrices. One wonders about the existence of, the way to obtain, etc., matrices Z when $P_1, P_2,$ and C_3 as well as the block-similarity class of P are prescribed. Obviously, in the classical Carlson problem one assumes that $C_1 = 0, C_2 = 0,$ and $C_3 = 0,$ that is to say, $P, P_1,$ and P_2 are endomorphisms.

In this paper, we consider the case when P is observable (and therefore so is P_1) with prescribed observability indices, and P_2 is an endomorphism (i.e., $C_2 = 0;$ see Remark 4.7) having only an eigenvalue $\lambda.$ We recall that the observability indices of P form the dual partition of

$$r_i = \text{rang} \begin{pmatrix} C \\ CA \\ \vdots \\ CA^i \end{pmatrix} - \text{rang} \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{i-1} \end{pmatrix}$$

and P is observable if

$$\text{rang} \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix} = n.$$

In fact, one can assume that P_1 is a BK -matrix, $C_3 = 0,$ and A_2 is a nilpotent Jordan matrix. This remark and the main results in this paper are summarized in the following theorem.

THEOREM 3.1. *Let there be three partitions:*

$$\begin{aligned} R &= (R_1, R_2, \dots), & |R| &= n, \\ r &= (r_1, r_2, \dots), & |r| &= d, \\ b &= (b_1, b_2, \dots), & |b| &= n - d. \end{aligned}$$

The following conditions are equivalent:

(I) *For any observable pair $P_1 \in \mathbb{C}^{d \times d} \times \mathbb{C}^{r_1 \times d}$ having observability indices r^* , any square matrix $A_2 \in \mathbb{C}^{(n-d) \times (n-d)}$ having only one eigenvalue λ , with Segre characteristic b^* , and any matrix $C_3 \in \mathbb{C}^{r_1 \times (n-d)}$ there is a matrix $Z \in \mathbb{C}^{d \times (n-d)}$ such that the pair*

$$P = \begin{pmatrix} A_1 & Z \\ 0 & A_2 \\ C_1 & C_3 \end{pmatrix}$$

is observable having observability indices R^ .*

(I') *Condition (I) holds in the particular case when P_1 is a BK-matrix, A_2 is a nilpotent Jordan matrix, and $C_3 = 0$.*

(II) *There is a finite sequence of partitions r^0, r^1, \dots, r^s ($s = \ell(b)$) such that $r^0 = r$, $r^s = R$, and for all $i, j \geq 1$*

- (a) $|r^j| - |r^{j-1}| = b_j$,
- (b) $r_1^j = r_1^{j-1}$, $r_i^{j-1} \geq r_{i+1}^j \geq r_{i+1}^{j-1}$,
- (c) $\sum_{\ell \geq i+1} (r_\ell^{j+1} - r_\ell^j) \leq \sum_{\ell \geq i} (r_\ell^j - r_\ell^{j-1})$.

(II') *There is a finite sequence of partitions c^0, c^1, \dots, c^s ($s = \ell(b)$) such that $c^0 = r^*$, $c^s = R^*$, and for all $\nu, j, i \geq 1$*

- (a) $|c^j| - |c^{j-1}| = b_j$,
- (b) $\ell(c^j) = r_1$, $c_\nu^{j-1} \leq c_\nu^j \leq c_\nu^{j-1} + 1$,
- (c) $\sum_{\eta \in I(i+1, j+1)} (c_\eta^{j+1} - c_\eta^j) \leq \sum_{\eta \in I(i, j)} (c_\eta^j - c_\eta^{j-1})$, where $I(i, j) = \{\eta : c_\eta^j \geq i\}$.

(III) (see [3]) $b_1 \leq r_1 = R_1$, $(R^*)_\nu \geq (r^*)_\nu$ ($\nu = 1, 2, \dots$), and $R^* - r^* \prec b^*$, where $R^* - r^*$ is assumed to be reordered to become nonincreasing.

Remark 3.2. Notice that conditions (II)(a)–(II)(c) are similar to the Littlewood–Richardson ones which appear in the classical Carlson problem (see Theorem 2.1). In fact, (II)(a) and (II)(b) are almost the same, whereas (II)(c) is in some sense “opposite.”

Remark 3.3. Condition (III) has been suggested by [3] by means of fully different methods. In fact, it is an immediate consequence of (II')(a) and (II')(b). Likewise (see the preceding remark) condition (III), except where $b_1 \leq r_1 = R_1$ (which holds only if $\text{Ker } f \subset W$), is a necessary condition in the classical Carlson problem, but in that case it is not sufficient.

Remark 3.4. When (III) holds, explicit solutions Z can be computed by means of (II), thus being nonequivalent for different sequences r^0, r^1, \dots, r^s (see section 7).

Remark 3.5. Conditions (II)–(II') can be sketched by means of the usual diagrams representing partitions in a way similar to the Littlewood–Richardson sequences.

In order to do that, let us take $c^j = (r^j)^*$ and represent them by a diagram D^j formed by $r_1^j (= r_1)$ towers having heights c_1^j, c_2^j, \dots , or, equivalently, each floor being r_1^j, r_2^j, \dots large. Then, D^j should be obtained by adding b_j blocks to D^{j-1} (condition (II)(a) or (II')(a)), in such a way that the rules (II)(b) and (II)(c), or equivalently, (II')(b) and (II')(c), are respected.

Condition (II)(b) says that the $(i + 1)$ -flat can increase up to the length of the i -one in D^{j-1} . That is to say, each tower can increase one block maximum (condition (II')(b)). As for the rule (II)(c), let us represent partition b by an analogous diagram and label the blocks on the j th floor b_j blocks. Recall that the rule (II)(b), or equivalently, (II')(b), means that to obtain D^j , the blocks labeled “ j ” should be assigned to different towers of D^{j-1} . Then, condition (II)(c), or (II')(c), means that the number of $(j + 1)$ -blocks installed at levels greater than $(i + 1)$ are at most the number of j -blocks at levels greater than i (for all i).

For instance, if $b = (3, 2)$ and $r = (4, 3, 2, 1)$, then the sequences

$$\begin{array}{cccc}
 & & & 2 \\
 1 & 2 & & 1 \\
 * & 1 & & * & 1 \\
 * & * & 1 & * & * & 1 \\
 * & * & * & 2 & * & * & * & 2 \\
 * & * & * & * & * & * & * & *
 \end{array}$$

are allowed, whereas

$$\begin{array}{cccc}
 2 & 2 & & \\
 * & 1 & & \\
 * & * & 1 & \\
 * & * & * & 1 \\
 * & * & * & *
 \end{array}$$

is not.

Proof of the equivalence (I) and (I'), and (II) and (II'). To see that (I') implies (I), notice that

$$\begin{aligned}
 & \left(\begin{array}{cc|c} Q_1 & Q_{12} & S_1 \\ 0 & Q_2 & 0 \\ \hline 0 & 0 & T_1 \end{array} \right) \left(\begin{array}{cc} A_1 & Z \\ 0 & A_2 \\ \hline C_1 & C_3 \end{array} \right) \left(\begin{array}{cc} Q_1 & Q_{12} \\ 0 & Q_2 \end{array} \right)^{-1} \\
 &= \left(\begin{array}{cc|cc} Q_1(A_1 + Q_1^{-1}S_1C_1)Q_1^{-1} & & \dots & \\ & 0 & & Q_2A_2Q_2^{-1} \\ \hline T_1C_1Q_1^{-1} & & T_1(-C_1Q_1^{-1}Q_{12} + C_3)Q_2^{-1} & \end{array} \right).
 \end{aligned}$$

Hence, P_1 can be reduced to a BK -matrix, A_2 to a Jordan matrix, and C_3 to 0 (since C_1 has the maximal rank, after eliminating, if necessary, its null rows and the corresponding ones in C_3). Furthermore, one can assume $\lambda = 0$ because

$$P - \lambda \begin{pmatrix} I_n \\ 0 \end{pmatrix}, \quad P_1 - \lambda \begin{pmatrix} I_d \\ 0 \end{pmatrix}$$

are block-similar to P and P_1 , respectively.

The equivalence (II) \Leftrightarrow (II') is a straightforward computation, taking $c^j = (r^j)^*$ (see Remark 3.5).

The next three sections are devoted to prove that conditions (I)–(I′) are equivalent to the (II)–(II′) ones. In fact, in section 4 we introduce a geometric version (I′′) of (I)–(I′), and in sections 5 and 6 we prove that (II) is a necessary and sufficient condition for (I′′), respectively.

Proof of the equivalence (II′) and (III). As it has been already remarked, (III) follows immediately from (II′)(a)–(II′)(b) (use Remark 3.5). Conversely, if (III) holds, the following strategy allows us to construct recurrently a sequence c^0, c^1, \dots, c^s which verifies condition (II′): for each $j = 1, 2, \dots$, let c^j be a maximal (with regard to the partial ordering \prec) partition such that

- (a) $|c^j| - |c^{j-1}| = b_j$,
- (b) $\ell(c^j) = r_1, c_\nu^{j-1} \leq c_\nu^j \leq c_\nu^{j-1} + 1$,
- (c′) $c^s - c^j \prec (b^j)^*$,

where $b^j = (b_{j+1}, b_{j+2}, \dots, b_s)$.

Notice that the set of partitions verifying (a)–(c′) is not empty because, in general, if $(\alpha_1, \alpha_2, \dots) \prec (\delta_1, \delta_2, \dots, \delta_\ell)$, then $(\alpha_1 - 1, \alpha_2 - 1, \dots, \alpha_\ell - 1, \alpha_{\ell+1}, \dots) \prec (\delta_1 - 1, \delta_2 - 1, \dots, \delta_\ell - 1)$, where the left member is assumed reordered to become nonincreasing. The proposed strategy takes c^j as a maximal element in this nonempty set.

By construction, conditions (II′)(a) and (II′)(b) are verified. Finally, let us see that if (II′)(c) does not hold, then c^{j-1} is not in fact maximal among the partitions verifying (a)–(c′) in the previous step. Broadly speaking, the following lemma shows that if the sequence $c^0, \dots, c^{j-1}, c^j, c^{j+1}, \dots$ verifies (a)–(c′) and

$$\begin{aligned} c_\mu^j &= c_\mu^{j-1}, & c_\eta^j &= c_\eta^{j-1} + 1, \\ c_\mu^{j+1} &= c_\mu^j + 1, & c_\eta^{j+1} &= c_\eta^j, \end{aligned}$$

then conditions (a)–(c′) are verified too if the partition c^j is replaced by \bar{c}^j , which differs from c^j only in

$$\bar{c}_\mu^j = c_\mu^{j-1} + 1, \quad \bar{c}_\eta^j = c_\eta^{j-1}$$

(that is to say, we have permuted the increasing order of the “towers” μ, η (see Remark 3.5)). In particular, if $c_\mu^{j+1} > c_\eta^{j+1}$, then $c^j \prec \bar{c}^j$, so that c^j was not in fact maximal among the partitions verifying (a)–(c′). \square

LEMMA 3.6. *Let α, δ be partitions such that*

$$\alpha \equiv (\alpha_1, \alpha_2, \dots) \prec \delta \equiv (\delta_1, \dots, \delta_\ell).$$

Assume that there is β such that

$$\begin{aligned} (\beta_1, \beta_2, \dots) &\prec (\delta_1 - 1, \dots, \delta_\ell - 1), \\ |\beta| &= |\alpha| - \ell, \\ \beta_\mu &= \alpha_\mu, & \beta_\eta &= \alpha_\eta - 1, \\ \alpha_\nu - 1 &\leq \beta_\nu \leq \alpha_\nu & \text{for all } \nu. \end{aligned}$$

Then partition $\bar{\alpha}$ defined by

$$\begin{aligned} \bar{\alpha}_\nu &= \alpha_\nu & \text{if } \nu \neq \mu, \eta, \\ \bar{\alpha}_\mu &= \alpha_\mu + 1, & \bar{\alpha}_\eta &= \alpha_\eta - 1 \end{aligned}$$

verifies $\bar{\alpha} \prec \delta$.

Proof. If $\alpha_\eta = \alpha_\mu + 1$, then $\bar{\alpha} = \alpha$ and there is nothing to prove.

If $\alpha_\eta > \alpha_\mu + 1$, then $\bar{\alpha} \prec \alpha \prec \delta$.

If $\alpha_\eta \leq \alpha_\mu$, we can assume (reordering, if it is necessary) $\alpha_{\mu-1} > \alpha_\mu \geq \alpha_\eta > \alpha_{\eta+1}$ (or $\mu = 1$ and $\alpha_1 \geq \alpha_\eta > \alpha_{\eta+1}$). Then the order in α and $\bar{\alpha}$ is the same and we have

$$\sum_{\nu \leq \mu_0} \bar{\alpha}_\nu = \sum_{\nu \leq \mu_0} \alpha_\nu \quad \text{if } \mu_0 < \mu \text{ or } \mu_0 \geq \eta,$$

and

$$\begin{aligned} \sum_{\nu \leq \mu_0} \bar{\alpha}_\nu &= 1 + \sum_{\nu \leq \mu_0} \alpha_\nu \leq 1 + \min(\mu_0 - 1, \ell - 1) + \sum_{\nu \leq \mu_0} \beta_\nu \\ &\leq 1 + \min(\mu_0 - 1, \ell - 1) + \sum_{\eta \leq \min(\mu_0, \ell)} (\delta_\nu - 1) = \sum_{\nu \leq \mu_0} \delta_\nu \quad \text{if } \mu \leq \mu_0 < \eta. \quad \square \end{aligned}$$

4. Geometric formulation. Let us consider a geometric approach, analogous to the one in section 2 for square matrices. The study of pairs of matrices (see [5])

$$P = \begin{pmatrix} A \\ C \end{pmatrix} \in \mathbb{C}^{(n+m) \times n}$$

under the block-similarity is equivalent to the one of linear maps defined on a subspace $f : Y \rightarrow X$, $Y \subset X$ ($\dim Y = n$, $\dim X = n + m$) under the following natural equivalence relation: $f \sim f'$ if and only if there is an automorphism φ of X , such that $\varphi(Y) = Y$ and $\varphi \circ f = f' \circ \hat{\varphi}$ (where $\hat{\varphi}$ means the restriction of φ to the subspace Y).

In fact, it is sufficient to consider P to be the matrix of f in any basis of X adapted to Y . In particular, the condition of C having the maximal rank is equivalent to $X = Y + f(Y)$. (This equality will hold throughout the paper.)

In these conditions, the observability indices of P (or f) can be computed as the conjugate partition of $\dim Y_0 - \dim Y_1, \dim Y_1 - \dim Y_2, \dots$, where

$$Y_i = f^{-i}(Y), \quad Y = Y_0 \supset Y_1 \supset \dots \supset Y_k = Y_{k+1} \equiv Y_\infty.$$

In particular, P (or f) is observable if and only if $Y_\infty = \{0\}$. Notice that then f is injective.

A subspace $W \subset Y$ is f -invariant (or (C, A) -invariant) if and only if $f(W) \cap Y \subset W$ (see [1], [4]). Let us see that the special form of P in section 3 appears in a natural way when invariant subspaces are considered.

DEFINITION 4.1. Let $f : Y \rightarrow X$ be a linear map defined on a subspace $Y \subset X$, and $W \subset Y$ an f -invariant subspace. Then

$$\begin{aligned} \hat{f} : W &\rightarrow W + f(W), \\ \tilde{f} : \frac{Y}{W} &\rightarrow \frac{X}{W + f(W)} \end{aligned}$$

will be the maps induced in a natural way by f .

Remark 4.2. It is clear that \hat{f} is a linear map defined on a subspace. Moreover, if W is f -invariant, \tilde{f} can also be considered to be of this kind by means of the following identification:

$$\begin{aligned} \frac{Y}{W} &= \frac{Y}{W + (f(W) \cap Y)} \cong \frac{Y + f(W)}{W + f(W)} \\ &\subset \frac{Y + f(Y)}{W + f(W)} = \frac{X}{W + f(W)}. \end{aligned}$$

PROPOSITION 4.3 (see [1], [4]). Let $f : Y \rightarrow X$ be as above, and $W \subset Y$ an f -invariant subspace. If $W_i = \widehat{f}^{-i}(W)$, then $W_i = Y_i \cap W$.

In particular, if f is observable, then \widehat{f} is observable too. Moreover, if their observability indices are $(R_1, R_2, \dots)^*$ and $(r_1, r_2, \dots)^*$, respectively, then $r_i \leq R_i$ for all $i = 1, 2, \dots$.

Let us characterize geometrically the special form of the matrices involved in this problem.

PROPOSITION 4.4. Let $f : Y \rightarrow X$ as above, and a subspace $W \subset Y$.

- (1) W is f -invariant if and only if the matrix of f in any basis adapted to $W \subset Y \subset Y + f(W) \subset X$ has the form

$$\begin{pmatrix} A \\ C \end{pmatrix} = \begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \\ \hline C_1 & C_3 \\ 0 & C_2 \end{pmatrix},$$

where $\begin{pmatrix} A_1 \\ C_1 \end{pmatrix}$ is the matrix of \widehat{f} in the same basis.

- (2) In the conditions of (1), the pair $\begin{pmatrix} A_2 \\ C_2 \end{pmatrix}$ is the matrix of \widetilde{f} in the basis induced in a natural way by the one considered in X .
- (3) In the above conditions, if f is observable, then \widetilde{f} is an endomorphism if and only if there is a basis of X adapted to $W \subset Y \subset X$ such that the matrix of f has the form

$$\begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \\ \hline C_1 & 0 \end{pmatrix}.$$

Proof.

- (1) It is a direct consequence of the inclusion $f(W) \cap Y \subset W$ which characterizes the f -invariant subspaces.
- (2) It is straightforward.
- (3) Because of Remark 4.2, \widetilde{f} is an endomorphism if and only if $f(Y) \subset Y + f(W)$. Obviously, this relation is verified for the matrices of the form considered. Conversely, taking anti-images in this inclusion, $Y = Y_1 + W$. Hence, there is a subspace V such that

$$Y = V \oplus W, \quad Y_1 = V \oplus W_1.$$

The latter implies $f(V) \subset Y$, so that in any basis adapted to $W, V \subset Y \subset X$ the matrix of f has the desired form. \square

Therefore, conditions (I)–(I') in Theorem 3.1 can be translated in the following geometric way, which will be used in the proof of the main theorem.

COROLLARY 4.5. Within the context of Theorem 3.1, conditions (I)–(I') are equivalent to the following condition:

(I'') There is a linear map defined on a subspace $f : Y \rightarrow X$, $Y \subset X$, and a f -invariant subspace $W \subset Y$ such that

- (1) f is observable, having observability indices R^* ;
- (2) \widehat{f} is observable, having observability indices r^* ;
- (3) \widetilde{f} is a nilpotent endomorphism, having Segre characteristic b^* .

Remark 4.6. From the proof of (3) in Proposition 4.4, it follows that if f is observable, then \tilde{f} is an endomorphism if and only if

$$\dim Y - \dim W = \dim Y_1 - \dim W_1.$$

Then a necessary condition for (I'') above is $R_1 = r_1$.

Remark 4.7. The assumption that \tilde{f} is an endomorphism is not a significant restriction. In general, if one considers the decreasing stationary chain of subspaces

$$\begin{aligned} Y &= Y^0 \supset Y^1 \supset \dots \supset Y^h = Y^{h+1} \equiv Y^\infty (\supset W), \\ Y^j &= f^{-1}(Y^{j-1}) + W, \end{aligned}$$

then for the restriction $f^\infty : Y^\infty \rightarrow X$ one has that $\widetilde{f^\infty}$ is an endomorphism, whereas the map induced by f in Y/Y^∞ is observable.

5. Proof of the necessity. Now let $f : Y \rightarrow X$ and $W \subset X$ be as in (I'') of Corollary 4.5. Following the pattern in section 2 in order to prove that condition (II) in Theorem 3.1 is verified, a double family of subspaces will be introduced: $W_i^j = Y_i \cap W^j$, where W^j is defined in such a way that $\text{Ker } \tilde{f}^j = W^j/W$, as will be seen below.

Notice that, as it has been recalled in section 4 (see [5]),

$$\begin{aligned} R_i &= \dim Y_{i-1} - \dim Y_i, \\ r_i &= \dim W_{i-1} - \dim W_i, \end{aligned}$$

where $Y_i = f^{-i}(Y)$, $W_i = \widehat{f}^{-i}(W) = f^{-i}(W) \cap W = Y_i \cap W$.

DEFINITION 5.1. Let $f : Y \rightarrow X$ and $W \subset Y$ be as in (I'') of Corollary 4.5. Then

$$\begin{aligned} W^0 &= W, \\ W^j &= f^{-1}(W^{j-1} + f(W)) = f^{-1}(W^{j-1}) + W, \quad j \geq 1. \end{aligned}$$

PROPOSITION 5.2. With the notation in the above definition,

- (1) $\text{Ker } \tilde{f}^j = \frac{W^j}{W}$ for all $j \geq 1$;
- (2) $W = W^0 \subset W^1 \subset \dots \subset W^{\ell(b)} = W^{\ell(b)+1} = \dots = Y$, $b_j = \dim W^j - \dim W^{j-1}$ for all $j \geq 1$;
- (3) the subspaces W^j are f -invariant. In fact, they verify $f(W^j) \cap Y \subset W^{j-1}$ for all $j \geq 1$.

Proof. (1) We proceed by induction, using the identification in Remark 4.2. It is obvious for $j = 0$. Assume that

$$\text{Ker } \tilde{f}^j = \frac{W^j}{W} \cong \frac{W^j + f(W)}{W + f(W)} \subset \frac{X}{W + f(W)}.$$

Then

$$\begin{aligned} \text{Ker } \tilde{f}^{j+1} &= \tilde{f}^{-1}(\text{Ker } \tilde{f}^j) \\ &= \frac{f^{-1}(W^j + f(W)) + W}{W} = \frac{W^{j+1}}{W}. \end{aligned}$$

- (2) It follows immediately from (1).

(3) The proof also follows by induction. For $j = 1$, we have

$$\begin{aligned} f(W^1) \cap Y &= f(f^{-1}(W) + W) \cap Y \\ &\subset (W + f(W)) \cap Y \subset W. \end{aligned}$$

If the property is verified by W^j , then

$$\begin{aligned} f(W^{j+1}) \cap Y &= f(f^{-1}(W^j) + W) \cap Y \\ &\subset (W^j + f(W)) \cap Y \\ &\subset (W^j + f(W^j)) \cap Y \subset W^j + W^{j-1} = W^j. \quad \square \end{aligned}$$

For each f -invariant subspace W^j , we consider the natural finite chain (W_i^j) .

DEFINITION 5.3. *In the conditions of Definition 5.1, we define for all $j \geq 0$*

$$\begin{aligned} W_i^j &= f^{-i}(W^j) \cap W^j = Y_i \cap W^j, \quad i \geq 0, \\ r_i^j &= \dim W_{i-1}^j - \dim W_i^j, \quad i \geq 1. \end{aligned}$$

Remark 5.4.

(1) It will be useful to bear in mind the following finite diagram:

$$\begin{array}{cccccccc} \cdots & \subset & f(W_1) & \subset & W & \subset & W^1 & \subset \cdots \subset W^{j-1} & \subset & W^j & \subset \cdots \subset W^{\ell(b)} & = & Y \\ & & \cup & & \cup & & \cup & & \cup & \cup & & \cup & \cup \\ \cdots & \subset & f(W_2) & \subset & W_1 & \subset & W_1^1 & \subset \cdots \subset W_1^{j-1} & \subset & W_1^j & \subset \cdots \subset W_1^{\ell(b)} & = & Y_1 \\ & & \cup & & \cup & & \cup & & \cup & \cup & & \cup & \cup \\ & & \vdots & & \vdots & & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ & & & & \cup & & \cup & & \cup & \cup & & \cup & \cup \\ \cdots & \subset & W_{i-1} & \subset & W_{i-1}^1 & \subset \cdots \subset W_{i-1}^{j-1} & \subset & W_{i-1}^j & \subset \cdots \subset W_{i-1}^{\ell(b)} & = & Y_{i-1} \\ & & \cup & & \cup & & \cup & & \cup & \cup & & \cup & \cup \\ \cdots & \subset & W_i & \subset & W_i^1 & \subset \cdots \subset W_i^{j-1} & \subset & W_i^j & \subset \cdots \subset W_i^{\ell(b)} & = & Y_i \\ & & \cup & & \cup & & \cup & & \cup & \cup & & \cup & \cup \\ & & \vdots & & \vdots & & \vdots & & \vdots & \vdots & & \vdots & \vdots \end{array}.$$

(2) Notice that

$$\begin{aligned} \frac{Y}{W} &\cong \bigoplus_{\substack{i \geq 0 \\ 1 \leq j \leq \ell(b)}} \frac{W_i^j}{W_i^{j-1} + W_{i+1}^j}, \\ \dim \frac{W_i^j}{W_i^{j-1} + W_{i+1}^j} &= \dim W_i^j - \dim W_i^{j-1} - \dim W_{i+1}^j + \dim W_{i+1}^j = r_{i+1}^j - r_{i+1}^{j-1}. \end{aligned}$$

These facts will guide the construction in the next section.

From the definitions and (3) of Proposition 5.2, it follows that $f(W_i^j) \subset W_{i-1}^{j-1}$ for all $i, j \geq 1$. Some basic properties of these maps are summarized in the following proposition.

PROPOSITION 5.5. *In the conditions of the above definition,*

- (1) $f^{-1}(W_{i-1}^{j-1}) = W_i^j$;
- (2) *the induced maps*

- (i) $\frac{W_i^j}{W_{i+1}^{j-1}} \longrightarrow \frac{W_{i-1}^{j-1}}{W_i^{j-1}}$,
- (ii) $\frac{W_i^{j+1}}{W_i^j} \longrightarrow \frac{W_{i-1}^j}{W_{i-1}^{j-1}}$

are injective for all $i, j \geq 1$.

Proof.

$$(1) f^{-1}(W_{i-1}^{j-1}) = f^{-1}(Y_{i-1} \cap W^{j-1}) = Y_i \cap f^{-1}(W^{j-1}) = Y_i \cap [f^{-1}(W^{j-1}) + W] = Y_i \cap W^j = W_i^j.$$

(2) It is a direct consequence of (1). \square

Finally, let us use the above construction to prove that condition (I'') of Corollary 4.5 implies (II) of Theorem 3.1. Let there be $(r^j)^*$, $0 \leq j \leq \ell(b)$, the observability indices of the restriction of f to each subspace W^j , that is to say

$$r^j = (r_1^j, r_2^j, \dots)$$

(see Definition 5.3). Notice that this restriction is observable (see Proposition 4.3 and (3) of Proposition 5.2); hence $|r^j| = \dim W^j$.

Let us see that, in fact, these partitions verify the properties in (II) of Theorem 3.1:

(a) Obviously $r^0 = r$, $r^{\ell(b)} = R$. Moreover,

$$\begin{aligned} |r^j| - |r^{j-1}| &= \dim W^j - \dim W^{j-1} \\ &= \dim \text{Ker } \tilde{f}^j - \dim \text{Ker } \tilde{f}^{j-1} = b_j \end{aligned}$$

for all $1 \leq j \leq \ell(b)$.

(b) From Proposition 4.3, $r_i^{j-1} \leq r_i^j$ for all $i, j \geq 1$. The equality holds for $i = 1$, according to Remark 4.6.

The inequality $r_i^j \leq r_{i-1}^{j-1}$ follows from the injectivity of (2)(i) in Proposition 5.5.

(c) Finally, condition (c) is a consequence of the injectivity of (2)(ii) in Proposition 5.5.

6. Proof of the sufficiency. Let partitions R, r, b and r^1, r^2, \dots, r^s be given, which verify (a)–(c) in (II) of Theorem 3.1. $f : Y \longrightarrow X, W \subset Y$ will be constructed in such a way that conditions (I'')(1), (I'')(2), and (I'')(3) in Corollary 4.5 are verified.

Let $W \subset Y \subset X$ be vector spaces having dimension $|r|, |R|$, and $|R| + R_1$, respectively. Let $\hat{f} : W \longrightarrow X$ be an observable linear map having observability indices r^* , so that condition (I'')(2) in Corollary 4.5 is verified. Also, a BK-basis of W is formed by $r_1 - r_2$ BK-chains having length 1, $r_2 - r_3$ BK-chains having length 2, etc. Let

$$\begin{aligned} B^0 &= \bigcup_{1 \leq i \leq \ell(r)} B_i^0; \\ B_i^0 &= \{e_{i,k}^0; 1 \leq k \leq r_i - r_{i+1}\} \end{aligned}$$

be a set of generators of these BK-chains. Hence

$$W_i = [B_{i+1}^0; B_{i+2}^0, f(B_{i+2}^0); B_{i+3}^0, f(B_{i+3}^0), f^2(B_{i+3}^0); \dots]$$

for all $1 \leq i \leq \ell(r)$.

Now, \hat{f} must be extended to $f : Y \longrightarrow X$ verifying (I'')(1) and (I'')(3) in Corollary 4.5. In order to achieve that, we consider any supplementary subspace \overline{W} of W in Y ,

and any basis B of it. Taking into account Remark 5.4, \overline{W} should be split into direct summands \overline{V}_i^j having a dimension

$$d_{i+1}^j = r_{i+1}^j - r_{i+1}^{j-1} \quad (i \geq 0, 1 \leq j \leq \ell(b)),$$

respectively, and then f will be defined on each of these subspaces \overline{V}_i^j . First, B is distributed (in any way) into subsets having cardinal d_{i+1}^j :

$$B = \bigcup_{\substack{i \geq 0 \\ 1 \leq j \leq \ell(b)}} B_{i+1}^j,$$

$$B_{i+1}^j = \{e_{i+1,k}^j; 1 \leq k \leq d_{i+1}^j\}.$$

(Notice that $B_{i+1}^j = \emptyset$ if $d_{i+1}^j = 0$; in particular, $B_1^j = \emptyset$, and $B_{i+1}^j = \emptyset$ if $i \geq \ell(r^j)$.)
 Second, $\overline{V}_i^j = [B_{i+1}^j]$ ($i \geq 0, 1 \leq j \leq \ell(b)$), so that

$$Y = W \oplus \overline{W} = W \oplus \left(\bigoplus_{\substack{i \geq 0 \\ 1 \leq j \leq \ell(b)}} \overline{V}_i^j \right),$$

$$\dim \overline{V}_i^j = d_{i+1}^j = r_{i+1}^j - r_{i+1}^{j-1}.$$

(Notice that $\overline{V}_0^j = \{0\}$, and $\overline{V}_i^j = \{0\}$ if $i \geq \ell(r^j)$.)

Considering the diagram

$$\begin{array}{ccccccc} & & \dots & & \dots & & \\ & & \oplus & & \oplus & & \\ \dots \oplus & \overline{V}_i^{j-1} & \oplus & \overline{V}_i^j & \oplus & \dots & \\ & \oplus & & \oplus & & & \\ \dots \oplus & \overline{V}_{i+1}^{j-1} & \oplus & \overline{V}_{i+1}^j & \oplus & \dots & \\ & \oplus & & \oplus & & & \\ & \dots & & \dots & & & \end{array}$$

and defining, for $i \geq 0, 0 \leq j \leq \ell(b)$,

$$V_i^j = W_i \oplus \left(\bigoplus_{\substack{\ell \geq i \\ 1 \leq h \leq j}} \overline{V}_\ell^h \right),$$

the following diagram is obtained:

$$\begin{array}{ccccccc} \dots \subset & W & \subset & V^1 & \subset \dots \subset & V^{\ell(b)} & = & Y \\ & \cup & & \cup & & \cup & & \cup \\ \dots \subset & W_1 & \subset & V_1^1 & \subset \dots \subset & V_1^{\ell(b)} & \equiv & V_1 \\ & \cup & & \cup & & \cup & & \cup \\ \dots & \dots & & \dots & & \dots & & \dots \end{array}$$

(where $V^j \equiv V_0^j$ and $V_i \equiv V_i^{\ell(b)}$), analogous to the one in Remark 5.4. Now, f will be defined on each \overline{V}_i^j in such a way that the corresponding subspaces W_i^j (according to Definition 5.3) are just V_i^j . Then, as desired, the observability indices of f will be R^* and the Segre characteristic of \tilde{f} will be $b^*, b_j = |r^j| - |r^{j-1}|$, so that the proof of the sufficiency will be finished.

To define f , in fact, two extensions, $f_*, f^* : Y \rightarrow X$ of \hat{f} , will be defined and then $f = \frac{1}{2}(f_* + f^*)$.

- (1) For each $i \geq 1$, f_* on \overline{V}_i^j will be defined by increasing recurrence over $1 \leq j \leq \ell(b)$.
 For $j = 1$,

$$f_*(e_{i+1,k}^1) = e_{i,k}^0 \in B_i^0 \subset W_{i-1}.$$

It is possible because the hypothesis (II)(b) implies

$$\dim \overline{V}_i^1 = r_{i+1}^1 - r_{i+1} \leq r_i - r_{i+1} = \#B_i^0.$$

For $j \geq 2$,

$$f_*(e_{i+1,k}^j) = e_{i,k}^{j-1} \in B_i^{j-1} \subset \overline{V}_{i-1}^{j-1}$$

if $1 \leq k \leq \min\{b_{i+1}^j, b_i^{j-1}\}$, and taking images

$$\begin{aligned} f_*(e_{i+1,k}^j) &\in B_i^{j-2} \cup B_i^{j-3} \cup \dots \cup B_i^0 \\ &\subset \overline{V}_{i-1}^{j-2} \oplus \dots \oplus \overline{V}_{i-1}^1 \oplus W_{i-1} \end{aligned}$$

in such a way that f_* is injective if $d_i^{j-1} < k \leq d_{i+1}^j$.
 It is possible because, as above,

$$\begin{aligned} \dim(\overline{V}_i^1 \oplus \dots \oplus \overline{V}_i^j) &= (r_{i+1}^1 - r_{i+1}) + (r_{i+1}^2 - r_{i+1}^1) + \dots + (r_{i+1}^j - r_{i+1}^{j-1}) \\ &= -r_{i+1} + r_{i+1}^j \leq -r_{i+1} + r_i^{j-1} \\ &= (r_i - r_{i+1}) + (r_i^1 - r_i) + \dots + (r_i^{j-1} - r_i^{j-2}) \\ &= \#B_i^0 + \#B_i^1 + \dots + \#B_i^{j-2} + \#B_i^{j-1}. \end{aligned}$$

- (2) Now

$$f^*(e_{i+1,k}^1) = f_*(e_{i+1,k}^1) = e_{i,k}^0.$$

For each $j \geq 2$, f^* is defined on \overline{V}_i^j by decreasing recurrence over $1 \leq i < \ell(r^j)$.

For $1 \leq k \leq \min\{d_{i+1}^j, d_i^{j-1}\}$,

$$f^*(e_{i+1,k}^j) = f_*(e_{i+1,k}^j) = e_{i,k}^{j-1}$$

and for $d_i^{j-1} < k \leq d_{i+1}^j$, taking images

$$\begin{aligned} f^*(e_{i+1,k}^j) &\in B_{i+1}^{j-1} \cup B_{i+2}^{j-1} \cup \dots \cup B_{\ell(r^{j-1})}^{j-1} \\ &\subset \overline{V}_i^{j-1} \oplus \overline{V}_{i+1}^{j-1} \oplus \dots \oplus \overline{V}_{\ell(r^{j-1})-1}^{j-1} \end{aligned}$$

in such a way that f^* is injective.

This is possible because of hypothesis (II)(c):

$$\dim \left(\bigoplus_{\ell \geq i} \overline{V}_\ell^j \right) = \sum_{\ell > i} (r_\ell^j - r_\ell^{j-1}) \leq \sum_{\ell > i-1} (r_\ell^{j-1} - r_\ell^{j-2}) = \dim \left(\bigoplus_{\ell \geq i-1} \overline{V}_\ell^{j-1} \right).$$

- (3) Finally, $f = \frac{1}{2}(f_* + f^*)$. Obviously, it is an extension of \widehat{f} .

The proof of Theorem 3.1 will be finished if $V_i^j = W_i^j$, or, equivalently, $Y_i = V_i$, $W^j = V^j$.

Obviously, $V_0 = Y$. Hence, it is sufficient to prove the following lemma.

LEMMA 6.1. *With the above notation,*

- (1) $f^{-1}(V^{j-1}) + W = V^j$ for all $j \geq 1$;
- (2) $f^{-1}(V_{i-1}) = V_i$ for all $i \geq 1$.

Proof. Previously notice that if a vector $e_{i,k}^{j-1} \in B$ belongs to $f_*(B)$ and also to $f^*(B)$, then either

$$e_{i,k}^{j-1} = f_*(e_{i+1,k}^j) = f^*(e_{i+1,k}^j) = f(e_{i+1,k}^j)$$

or there are some unique $h > 0$ and $\ell \geq 0$ such that

$$e_{i,k}^{j-1} \in f_*(B_{i+1}^{j+h}) \cap f^*(B_{i-\ell}^j).$$

- (1) By construction

$$f(V^j) \subset V^{j-1} + f(W).$$

Hence

$$V^j \subset f^{-1}(V^{j-1}) + W.$$

For the opposite inclusion, assume $x \notin V^j$ and let J be the maximum index $J > j$ such that x has some nonzero component in $\bar{V}^J \equiv \oplus_i \bar{V}_i^J$. Then $f^*(x)$ should have some nonzero component in $\bar{V}^{J-1} \equiv \oplus_i \bar{V}_i^{J-1}$. According to the previous note, and bearing in mind the definition of J , this component cannot be canceled by any component of $f_*(x)$, so that $f(x)$ has in fact some nonzero component in \bar{V}^{J-1} . Therefore, $f(x) \notin V^{j-1} + f(W)$.

(2) By construction, $f(V_i) \subset V_{i-1}$. Hence, $V_i \subset f^{-1}(V_{i-1})$. For the opposite inclusion, we proceed by increasing recurrence over i , in an analogous way to (1). \square

7. Construction of solutions. When condition (III) of Theorem 3.1 holds, explicit solutions Z verifying (I)–(I') can be obtained by means of the construction in section 6, starting on any sequence of partitions verifying (II). Two of such solutions Z, Z' will be called *equivalent* if the associated matrices can be transformed one into the other by means of a change of basis preserving their block structure, that is to say, if

$$\left(\begin{array}{cc|c} Q_1 & Q_{12} & S_1 \\ 0 & Q_2 & 0 \\ \hline 0 & 0 & T_1 \end{array} \right) \left(\begin{array}{cc} A_1 & Z \\ 0 & A_2 \\ \hline C_1 & C_3 \end{array} \right) \left(\begin{array}{cc} Q_1 & Q_{12} \\ 0 & Q_2 \end{array} \right)^{-1} = \left(\begin{array}{cc} A_1 & Z' \\ 0 & A_2 \\ \hline C_1 & C_3 \end{array} \right),$$

where Q_1, Q_2 , and T_1 are nonsingular. Clearly, different sequences of partitions as in (II) lead to nonequivalent solutions. Example 7.3 shows that nonequivalent solutions are possible even for the same sequence of partitions.

Example 7.1. Clearly, the partitions

$$R = (2, 2, 1), \quad r = (2), \quad b = (1, 1, 1)$$

verify condition (III) of Theorem 3.1. Two sequences of partitions verifying (II) are possible:

$$(2, 1), \quad (2, 2), \quad (2, 2, 1);$$

$$(2, 1), \quad (2, 1, 1), \quad (2, 2, 1).$$

According to the construction in section 6, they lead, respectively, to the following nonequivalent solutions:

0	0	0	0	1
0	0	0	1	0
0	0	0	0	0
0	0	1	0	0
0	0	0	1	0
1	0	0	0	0
0	1	0	0	0

0	0	0	0	1
0	0	1	0	0
0	0	0	0	0
0	0	1	0	0
0	0	0	1	0
1	0	0	0	0
0	1	0	0	0

Example 7.2. In general, condition (III) of Theorem 3.1 holds if $b = (1, 1, \dots, 1)$. It is not difficult to see that a sequence of partitions $r^j, 0 \leq j \leq \ell(b)$, verifying (II) can be constructed by recurrence as follows:

$$r_{i(j)}^j = r_{i(j)}^{j-1} + 1,$$

$$r_i^j = r_i^{j-1} \quad \text{if } i \neq i(j),$$

where

$$i(j) = \max\{i : r_i^{j-1} < R_i, \quad r_i^{j-1} < r_{i-1}^{j-1}\}.$$

Then, as in the previous example, explicit solutions can be obtained by means of the construction in section 6.

Example 7.3. Let us consider the partitions

$$R = (2, 2, 1, 1), \quad r = (2), \quad b = (1, 1, 1, 1).$$

It is a straightforward computation to see that the solutions

0	0	0	λ	0	1
0	0	0	0	1	0
0	0	0	0	0	0
0	0	1	0	0	0
0	0	0	1	0	0
0	0	0	0	1	0
1	0	0	0	0	0
0	1	0	0	0	0

are nonequivalent for different values of $\lambda \in \mathbb{C}$, although all of them correspond to the sequence of partitions

$$(2, 1), \quad (2, 2), \quad (2, 2, 1), \quad (2, 2, 1, 1).$$

In a similar way to the “condensation lemma” for the classical Carlson problem, let us see that many zero entries can be prescribed in the Z solutions.

LEMMA 7.4. *Let R, r, b be three partitions verifying the conditions in Theorem 3.1, and consider the particular case in (I'). Then*

- (1) for any Z solution, there is an equivalent Z' solution having nonzero entries only in the r_1 rows corresponding to the null ones in A_1 ;
- (2) moreover, Z' can be chosen in such a way that its entries in the b_1 columns corresponding to the null ones in A_2 are also 0, except one of them in each column, which can be valued 1 and are placed in different rows.

Proof. (1) It is immediate that each vector in the basis of Y , not in W , can be changed by adding a vector in W in such a way that its image by f would be a linear combination of the generators of the BK-chains of W . Explicitly, assume $P_1 = \begin{pmatrix} N \\ E \end{pmatrix}$ is a BK-matrix and $A_2 = J$ is a nilpotent Jordan matrix. Notice that

$$\left(\begin{array}{cc|c} I & Q_{12} & 0 \\ 0 & I & 0 \\ \hline 0 & 0 & I \end{array} \right) \begin{pmatrix} N & Z \\ 0 & J \\ \hline E & 0 \end{pmatrix} \begin{pmatrix} I & Q_{12} \\ 0 & I \end{pmatrix}^{-1} = \begin{pmatrix} 0 & Z - NQ_{12} + Q_{12}J \\ 0 & J \\ \hline E & EQ_{12} \end{pmatrix}.$$

Let us choose Q_{12} in such a way that $EQ_{12} = 0$ and $Z' = Z - NQ_{12} + Q_{12}J$ verifies the desired property. For the first, it is sufficient to make null the rows in Q_{12} corresponding to the lowest one in each block of N . The remaining rows of Q_{12} can be computed easily by recurrence in order to cancel all the rows in Z except those corresponding to null ones in N . For example, let

$$N = \begin{pmatrix} 0 & 0 & | & & & \\ 1 & 0 & | & & & \\ \hline & & | & 0 & 0 & 0 \\ & & | & 1 & 0 & 0 \\ & & | & 0 & 1 & 0 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & | & & & \\ \hline & | & 0 & 0 & 0 \\ & | & 1 & 0 & 0 \\ & | & 0 & 1 & 0 \end{pmatrix}, \quad Q_{12} = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ 0 & 0 & 0 & 0 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then $EQ_{12} = 0$,

$$-NQ_{12} + Q_{12}J = \begin{pmatrix} 0 & a_3 & a_4 & 0 \\ -a_1 & -a_2 & -a_3 & -a_4 \\ 0 & c_3 & c_4 & 0 \\ -c_1 & -c_1 + d_3 & -c_3 + d_4 & -c_4 \\ -d_1 & -d_2 & -d_3 & -d_4 \end{pmatrix}.$$

It is clear that Q_{12} can be chosen in such a way that $Z - NQ_{12} + Q_{12}J$ has zero entries in the second, fourth, and fifth rows.

(2) From Proposition 5.5, it follows immediately that, for all $i \geq 1$, the maps induced by f

$$\frac{W_i^1}{W_{i+1}^1 + W_i} \longrightarrow \frac{W_{i-1}}{W_i + f(W_i)}$$

are injective. Notice that the vectors in W^1/W are the eigenvectors of \tilde{f} . Thus, because of the above injectivities, the images of a basis of \tilde{f} -eigenvectors (in fact, of

a set of representative vectors in W^1) can be extended to a family of BK -generators of \widehat{f} . \square

Solutions having a minimal number of nonzero entries arise when the subspace W is “marked” [8], [4], that is to say, when there is some BK -basis of \widehat{f} extendible to a BK -basis of f .

COROLLARY 7.5. *Let R, r, b be three partitions verifying the conditions in Theorem 3.1 and Corollary 4.5. Then the following assertions are equivalent:*

- (1) *In terms of condition (I'), there is some solution Z whose only nonzero entries are those referred to in part (2) of Lemma 7.4, that is to say, b_1 1-valued entries placed in (different) columns corresponding to the null ones in J , and in (different) rows corresponding to the null ones in N .*
- (2) *There is an f -marked subspace W verifying (I'') of Corollary 4.5.*
- (3) *With the notation in (III): $R^* - r^* = b^*$.*

REFERENCES

- [1] I. BARAGAÑA AND I. ZABALLA, *Block similarity invariants of restrictions to (A, B) -invariant subspaces*, Linear Algebra Appl., 220 (1995), pp. 31–62.
- [2] I. BARAGAÑA AND I. ZABALLA, *Feedback invariants of supplementary pairs of matrices*, Automatica J. IFAC, 33 (1997), pp. 2119–2130.
- [3] I. BARAGAÑA AND I. ZABALLA, *Feedback Invariants of Restrictions and Quotients: Series and Parallel Connected Systems*, Preprint, 1998.
- [4] A. COMPTA AND J. FERRER, *On $(A, B)^t$ -invariant subspaces having extendible Brunovsky bases*, Linear Algebra Appl., 255 (1997), pp. 185–201.
- [5] J. FERRER AND F. PUERTA, *Similarity of non-everywhere defined linear maps*, Linear Algebra Appl., 168 (1992), pp. 27–55.
- [6] I. GOHBERG, AND M.A. KAASHOEK, *Unsolved problems in matrix and operator theory II. Partial multiplicities for products*, Integral Equations Operator Theory, 2 (1979), pp. 116–120.
- [7] I. GOHBERG, M.A. KAASHOEK, AND F. VAN SCHAGEN, *Partially Specified Matrices and Operators: Classification, Completion, Applications*, Oper. Theory Adv. Appl. 79, Birkhäuser-Verlag, Basel, 1995.
- [8] I. GOHBERG, P. LANCASTER, AND L. RODMAN, *Invariant Subspaces of Matrices with Applications*, John Wiley, New York, 1986.
- [9] T. KLEIN, *The multiplication of Schur functions and extensions of p -modules*, J. London Math. Soc., 43 (1968), pp. 280–284.
- [10] R.C. THOMPSON, *Smith invariants of a product of integral matrices*, Contemp. Math., 47 (1985), pp. 401–435.
- [11] W.M. WONHAM, *Linear Multivariable Control: A Geometric Approach*, Springer, New York, 1979.