Reviewing the closure hierarchy of orbits and bundles of system pencils and their canonical forms*

Stefan Johansson

UMINF 09.02

Department of Computing Science,
Umeå University,
SE-901 87 Umeå, Sweden.
stefanj@cs.umu.se

Abstract

Using a unifying terminology and notation an introduction to the theory of stratification for orbits and bundles of matrices, matrix pencils and system pencils with applications in systems and control is presented. Canonical forms of such orbits and bundles reveal the important system characteristics of the models under investigation. A stratification provides the qualitative information of which canonical structures are near each other in the sense of small perturbations. We discuss how fundamental concepts like controllability and observability of a system can be studied with the use of the stratification theory. Important results are presented in the form of the closure and cover relations for controllability and observability pairs. Furthermore, different canonical forms are considered from which we can derive the characteristics of a system. Specifically, we discuss how the Kronecker canonical form is related to the Brunovsky canonical form and its generalizations. Concepts and results are illustrated with several examples throughout the presentation.

Key words: Stratification, Jordan canonical form, Kronecker canonical form, Brunovsky canonical form, orbit, bundle, closure relations, cover relations, state-space system, system pencil, matrix pencil, matrix pair.

*Financial support has been provided by the *Swedish Foundation for Strategic Research* under the frame program grant A3 02:128.

Contents

1 Introduction

To study how linear state-space systems and models behave under small perturbations is a critical task, since computing the canonical structure of such a system is an ill-posed problem and is therefore sensitive to small perturbations. The canonical structure is, for example, of interest when computing the controllability and observability characteristics of a linear system.

We exemplify the problems that can arise by a steering system of an airplane. For such critical systems it is crucial that the system is controllable in all possible states. What we mean by that is, loosely speaking, that the steering should always (in any situation) react as predicted and should not collapse in an uncontrollable state so that the airplane no longer can be controlled. In a given time, the computed canonical structure of the steering system may indicate that we can control all rudders of the airplane. However, it may be so that a particular unexpected reaction from the pilot results in that one of the components in the steering systems no longer is controllable. Especially, it is important to know the canonical structure of these uncontrollable systems and how near they are our controllable system. Most likely, such uncontrollable systems are almost impossible to reach in practice.

In this paper, we consider *linear time-invariant, finite dimensional systems* (LTI systems) which in continuous time are represented as a *state-space model* by a system of differential equations

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t), \end{aligned} \tag{1.1}$$

where $A \in \mathbb{C}^{n\times n}$ is the system (state) matrix, $B \in \mathbb{C}^{n\times m}$ is the input (control) matrix, $C \in \mathbb{C}^{p\times n}$ is the output matrix, and $D \in \mathbb{C}^{p\times m}$ is the feedforward matrix. Moreover, $x(t)$ is the state vector, $u(t)$ is the input vector, and $y(t)$ is the output vector. We also consider the *generalized state-space system* (or *descriptor system*)

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t), \end{aligned} \tag{1.2}$$

where E can be singular. In the most general case, A and E can be rectangular matrices. However, in most cases we assume that they are square and E is nonsingular, and the generalized state-space system can (in theory) be transformed into the state-space form (1.1). In short form, the state-space system (1.1) is represented by the *quadruple of matrices* (A, B, C, D), and the generalized state-space system (1.2) by the 5-tuple (E, A, B, C, D).

An LTI system can also be represented and analyzed in terms of a general *matrix pencil* $G - \lambda H$, where G and H are $m_{\mathrm{p}} \times n_{\mathrm{p}}$ complex matrices and $\lambda \in \mathbb{C}$. A matrix pencil associated with a state-space system (1.1) is called a *system pencil* $S - \lambda T$ and has the form

$$\mathbf{S}(\lambda) = S - \lambda T = \begin{bmatrix} A & B \\ C & D \end{bmatrix} - \lambda \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \tag{1.3}$$

where S and T are of size $(n+p) \times (n+m)$.

We also consider the controllability pair (A, B) and the observability pair (A, C), associated with the particular systems

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad \text{and} \qquad \begin{aligned} \dot{x}(t) &= Ax(t), \\ y(t) &= Cx(t), \end{aligned}$$

respectively, of (1.1). The corresponding system pencils for the controllability and observability pairs are

$$\mathbf{S}_{\mathrm{C}}(\lambda) = \begin{bmatrix} A & B \end{bmatrix} - \lambda \begin{bmatrix} I_n & 0 \end{bmatrix}, \tag{1.4}$$

and

$$\mathbf{S}_{\mathrm{O}}(\lambda) = \begin{bmatrix} A \\ C \end{bmatrix} - \lambda \begin{bmatrix} I_n \\ 0 \end{bmatrix}. \tag{1.5}$$

These systems can also appear in generalized versions with the identity matrix I_n replaced by E as in (1.2). As the names indicate, the controllability and observability characteristics of an LTI system are revealed by the matrix pairs (A, B) and (A, C), respectively.

An LTI system (1.1) is said to be *controllable* if there exists an input signal $u(t)$, $t_0 \le t \le t_{\mathrm{f}}$, that takes every state variable from an initial state $x(t_0)$ to a desired final state $x(t_{\mathrm{f}})$ in finite time. Otherwise it is said to be *uncontrollable*. The dual concept of controllability is observability. System (1.1) is said to be *observable* if it is possible to find the initial state $x(t_0)$ from the input signal $u(t)$ and the output signal $y(t)$ measured over a finite interval $t_0 \le t \le t_{\mathrm{f}}$. Otherwise it is said to be *unobservable*.

Controllability and observability are two fundamental concepts in systems and control theory. Other fundamental concepts are for example poles, zeros, reachability, stability, and detectability. These system characteristics are reveled from the canonical structure information of the appropriate system pencil. An overview of canonical forms and structure information for matrices, matrix pencils, and system pencils constitute one major part of the paper. For an introduction to systems and control theory, we refer to [1, 17, 18, 56, 95, 97, 100] where also numerical aspects are discussed.

The next major part study how small perturbations can change the canonical structure of a matrix A, a matrix pencil $G - \lambda H$, and for independent system pencils $S - \lambda T$ associated with a controllability pair (A, B) and a observability pair (A, C), respectively. A *stratification* provides the qualitative information of which canonical structures are near each other in the sense of small perturbations [30, 31, 36, 72, 75]. To give a comprehensive review of existing results for stratification of orbits and bundles is one major contribution of this paper. The focus is on the stratification of matrices, matrix pencils, and matrix pairs. We present the stratification theory and its theoretical background, illustrated with several examples [30, 31]. Several other people have worked on the theory of stratifications and similar topics, e.g., see [7, 44, 59, 65, 96] and references there in. Furthermore, the related topic distance to uncontrollability [94] has been studied in, e.g., [9, 30, 32, 37, 58], and more recently in [14, 15, 29, 35, 61, 62, 88].

The stratification is the closure hierarchy of matrix (and matrix pencil) orbits and bundles of canonical structures. The hierarchy is obtained from the closure and cover relations of orbits and bundles, where the cover relations guarantee that two structures are nearest neighbours in the closure hierarchy. For example, a matrix orbit is the manifold of all similar matrices, and a bundle is the union of all orbits with the same canonical form but with unspecified eigenvalues [2]. The closure hierarchies can be analyzed and illustrated with the software tool StratiGraph [34, 70, 73, 74].

The rest of the paper is organized as follows. In Section 2, we review different canonical forms for matrices, matrix pencils and system pencils. These are the Jordan canonical form (JCF), Kronecker canonical form (KCF) and Brunovsky canonical form (BCF) with generalizations. Especially, in Section 2.7 we derive the permutation matrices that take a matrix pencil in Kronecker canonical form to (generalized) Brunovsky canonical form. In Section 3.1, we discuss existing numerical stable methods to compute the canonical structure information for matrices, matrix pencils and system pencils, using staircase-type algorithms. The following and related Section 3.2, considers the computation of the controllable and unobservable subspaces of a system. In Section 4, the geometry of the tangent and normal spaces of the orbits of matrices, matrix pencils, and system pencils, are considered. In the main section, Section 5, we present the theory of stratification of matrices, matrix pencils, and matrix pairs. In Section 5.1, we give a brief introduction to integer partitions and coin moves, which are used to define the stratification rules. In Section 5.3, the stratification rules for matrices and matrix pencils are presented and the stratification rules for matrix pairs are derived. Section 5 is ended with an extensive example illustrating the stratification of a state-space system. Finally, we give some concluding remarks in Section 6. As appendices, we present some important parts of the paper in a comprehensive and compact form. A summarizes the explicit expressions to compute the codimensions of orbits and bundles, and in B the stratification rules for matrices, matrix pencils and matrix pairs are summed up. Finally, we have summarized the most important notation used in this paper in C.

2 Canonical forms and invariants

In linear algebra, it is a well known fact that a matrix (or matrix pencil) can be transformed to different canonical forms in terms of similarity (or equivalence) transformations. In Section 2.1, we introduce the Schur form and the Jordan canonical form for matrices, and in Section 2.2 the Kronecker canonical form for matrix pencils is presented. In Section 2.5, we summarize the most common types of transformations used in systems and control theory. Then we introduce the Brunovsky canonical form with generalizations for system pencils in Section 2.6.

We discuss different representations and invariants for matrices, and matrix pencils in Sections 2.3 and 2.4. Moreover, in Section 2.7 we prove that it is possible with two permutation matrices to transform a matrix pencil in Kronecker canonical form to a corresponding system pencil in (generalized) Brunovsky canonical form, and vice versa.

2.1 Schur form and Jordan canonical form

For square matrices there exist two fundamental canonical forms, the *Schur form* and the *Jordan canonical form* (JCF) (also called Jordan normal form) [43, 55]. For many applications it is enough to compute the Schur form, which is both more numerically stable and less expensive to compute than JCF. To get the Schur form, in the complex case, we transform a matrix A to a similar upper triangular matrix such that $\widetilde{A} = QAQ^H$ with Q unitary, where the eigenvalues show up on the diagonal. In the real case, the matrix $\widetilde{A}$ is upper *quasi-triangular*, i.e., a block upper triangular matrix with 1-by-1 diagonal blocks corresponding to real eigenvalues and 2-by-2 blocks on the diagonal associated with complex conjugate pairs of eigenvalues.

For our purpose the Jordan canonical form is more adequate. If there exists a nonsingular matrix P such that $\widetilde{A} = PAP^{-1}$, then the matrices A and $\widetilde{A}$ are said to be *similar*[1]. For any matrix $A \in \mathbb{C}^{n\times n}$ there exists a similarity transformation such that

$$PAP^{-1} = \widetilde{A} = \operatorname{diag}(J(\mu_1), J(\mu_2), \ldots, J(\mu_q)),$$

and

$$J(\mu_i) = \operatorname{diag}(J_{h_1}(\mu_i), J_{h_2}(\mu_i), \ldots, J_{h_{g_i}}(\mu_i)), \quad h_1 \geq \cdots \geq h_{g_i} \geq 1,$$

where $J_{h_1}(\mu_i), \ldots, J_{h_{g_i}}(\mu_i)$ are *Jordan blocks for matrices* of size $h_k \times h_k$ with eigenvalue μ_i, and each Jordan block is defined as

$$J_{h_k}(\mu_i) = \begin{bmatrix} \mu_i & 1 & & \\ & \mu_i & \ddots & \\ & & \ddots & 1 \\ & & & \mu_i \end{bmatrix},$$

where left-out elements are zeros. The block diagonal matrix $\widetilde{A}$ is now said to be in Jordan canonical form with $q \leq n$ distinct (possibly multiple) eigenvalues.

The *algebraic multiplicity* a_i of the eigenvalue μ_i is the multiplicity of μ_i as a root of the characteristic equation $\det(A - \lambda I) = 0$. The *geometric multiplicity* g_i is the number of

[1] Notice, in order to conform with what is typically used in the matrix theory of state-space transformations, the transformation matrix applied to the right hand side of a matrix in similarity and equivalence transformations is expressed as a matrix inverse.

linearly independent eigenvectors associated with μ_i. We remark that $a_i = h_1 + \cdots + h_{g_i}$ and that $g_i \leq a_i$ corresponds to the number of Jordan blocks associated with the eigenvalue μ_i.

2.2 Kronecker canonical form

For general matrix pencils $G - \lambda H$ of size $m_\mathrm{p} \times n_\mathrm{p}$ we use the *Kronecker canonical form* (KCF), which is a generalization of JCF to general matrix pencils [43]. Two matrix pencils $G - \lambda H$ and $\widetilde{G} - \lambda\widetilde{H}$ are *strictly equivalent*[1] if there exist two nonsingular matrices U and V such that $\widetilde{G} - \lambda\widetilde{H} = U(G - \lambda H)V^{-1}$. Any matrix pencil can be transformed into KCF in terms of an equivalence transformation such that

$$\begin{aligned} &U(G - \lambda H)V^{-1} \\ &\quad = \operatorname{diag}(L_{\epsilon_1}, \ldots, L_{\epsilon_{r_0}}, J(\mu_1), \ldots, J(\mu_q), N_{s_1}, \ldots, N_{s_{g_\infty}}, L^T_{\eta_1}, \ldots, L^T_{\eta_{l_0}}), \end{aligned} \tag{2.6}$$

where $J(\mu_i) = \operatorname{diag}(J_{h_1}(\mu_i), \ldots, J_{h_{g_i}}(\mu_i))$, $i = 1, \ldots, q$. The blocks $J_{h_k}(\mu_i)$ are $h_k \times h_k$ *Jordan blocks for matrix pencils* associated with each distinct finite eigenvalue μ_i and the blocks N_{s_k} are $s_k \times s_k$ Jordan blocks for matrix pencils associated with the infinite eigenvalue. Moreover, g_i is the geometric multiplicity of the finite eigenvalues μ_i and g_∞ is the geometric multiplicity of the infinite eigenvalue. These two types of blocks constitute the *regular part* of a matrix pencil and are defined by

$$J_{h_k}(\mu_i) = \begin{bmatrix} \mu_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \mu_i \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & & \\ & \ddots & \ddots & \\ & & \ddots & 0 \\ & & & 1 \end{bmatrix}, \tag{2.7}$$

and

$$N_{s_k} = \begin{bmatrix} 1 & 0 & & \\ & \ddots & \ddots & \\ & & \ddots & 0 \\ & & & 1 \end{bmatrix} - \lambda \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}. \tag{2.8}$$

If $m_\mathrm{p} \neq n_\mathrm{p}$ or $\det(G - \lambda H) \equiv 0$ for all $\lambda \in \mathbb{C}$, then the matrix pencil also includes a *singular part* and we say that the matrix pencil is singular. The singular part of the KCF consists of the r_0 *right singular blocks* L_{ϵ_k} of size $\epsilon_k \times (\epsilon_k + 1)$ and the l_0 *left singular blocks* $L^T_{\eta_k}$ of size $(\eta_k + 1) \times \eta_k$, defined as

$$L_{\epsilon_k} = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix}, \tag{2.9}$$

and

$$L^T_{\eta_k} = \begin{bmatrix} 0 & & \\ 1 & \ddots & \\ & \ddots & 0 \\ & & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & & \\ 0 & \ddots & \\ & \ddots & 1 \\ & & 0 \end{bmatrix}. \tag{2.10}$$

An L_0 and an L_0^T block are of size 0×1 and 1×0, respectively, and each of them contributes to a column or row of zeros (see Example 1). The size of a matrix pencil is equal the sum of the sizes of all blocks in its KCF:

$$\begin{aligned} m_{\mathrm{p}} &= \sum_{j=1}^{r_0} \epsilon_j + \sum_{j=1}^{l_0} (\eta_j + 1) + \sum_{i=1}^{q} \sum_{j=1}^{g_i} h_j^{(i)} + \sum_{j=1}^{g_\infty} s_j, \text{ and} \\ n_{\mathrm{p}} &= m_p - l_0 + r_0 \end{aligned}$$

where $h_j^{(i)}$ are the sizes of the Jordan blocks associated with eigenvalue μ_i, $i = 1, \ldots, q$.

For consistency reasons, the L blocks always appear before the L^T blocks in the KCF. Apart from that the order of the blocks is arbitrary. Moreover, a general matrix pencil may only consist of a subset of the different types of canonical blocks mentioned above. For example, a regular pencil ($\det(G - \lambda H) \not\equiv 0$, except when λ is an eigenvalue) only has J and N blocks.

The transformation matrices used to compute the Kronecker canonical form can be very ill-conditioned, therefore it is more appropriate to compute a generalized Schur-staircase form of the matrix pencil, see Section 3.1. Notably, if the KCF is computed the elements represented by ones in the blocks $J_k(\mu_i)$, N_k, L_k and L_k^T are not forced to be ones, instead we just get them as nonzero entries. Moreover, the eigenvalues μ_i are computed as pairs of values (α_i, β_i), $\alpha_i \neq 0$ and/or $\beta_i \neq 0$, for $i = 1, \ldots, q$. If $\beta_i \neq 0$, for some i, then the eigenvalue $\mu_i = \alpha_i/\beta_i$, and if $\alpha_i \neq 0$ and $\beta_i = 0$ then μ_i is an infinite eigenvalue. In Section 3.1, we review how the eigenvalues are computed in practice (finite precision arithmetic).

2.3 Block structure notation

Both for matrices and matrix pencils we often use a compact notation, which we refer to as *block structure notation*, instead of expressing their canonical forms in matrix form. In general, a block diagonal matrix A with b blocks $A_1, A_2, \ldots, A_b$ can be represented as a direct sum

$$A \equiv A_1 \oplus A_2 \oplus \cdots \oplus A_b \equiv \bigoplus_{k=1}^{b} A_k.$$

Equation (2.6) can now compactly be rewritten as

$$U(G - \lambda H)V^{-1} \equiv \mathbb{L} \oplus \mathbb{L}^T \oplus \mathbb{J}(\mu_1) \oplus \cdots \oplus \mathbb{J}(\mu_q) \oplus \mathbb{N},$$

where

$$\begin{aligned} \mathbb{L} &= \bigoplus_{j=1}^{r_0} L_{\epsilon_j}, \qquad \mathbb{L}^T = \bigoplus_{j=1}^{l_0} L_{\eta_j}^T, \\ \mathbb{J}(\mu_i) &= \bigoplus_{j=1}^{g_i} J_{h_j}(\mu_i), \quad \text{and} \quad \mathbb{N} = \bigoplus_{j=1}^{g_\infty} N_{s_j}. \end{aligned}$$

Notably, in the block structure notation we reorder the blocks such that the L^T blocks appear directly after the L blocks.

Example 1

Consider a matrix pencil with two L_1 blocks, one L_0^T block and one $J_2(\alpha)$ block. The KCF of this matrix pencil is in block structure notation written as $2L_1 \oplus L_0^T \oplus J_2(\alpha)$. The corresponding representation in matrix form is

$$
\begin{aligned}
G - \lambda H &= \operatorname{diag}(L_1, L_1, J_2(\alpha), L_0^T) \\
&= \begin{bmatrix} 0 & 1 & & & & & \\ & & 0 & 1 & & & \\ & & & & \alpha & 1 & \\ & & & & 0 & \alpha & \\ & & & & & & \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & & & & & \\ & & 1 & 0 & & & \\ & & & & 1 & 0 & \\ & & & & 0 & 1 & \\ & & & & & & \end{bmatrix} \\
&= \begin{bmatrix} -\lambda & 1 & & & & & \\ & & -\lambda & 1 & & & \\ & & & & \alpha-\lambda & 1 & \\ & & & & 0 & \alpha-\lambda & \\ & & & & & & \end{bmatrix}.
\end{aligned}
$$

2.4 Invariants of matrices and matrix pencils

The matrix pencil characteristics can equivalently be expressed in terms of column/row minimal indices and finite/infinite elementary divisors. It follows that two matrix pencils are strictly equivalent if and only if they have the same minimal indices and elementary divisors or, equivalently, if they have the same KCF, i.e., the same L, L^T, J and N blocks [43]. Before defining these invariants, we introduce integer partitions which are used to represent the invariants.

An *integer partition* $\kappa = (\kappa_1, \kappa_2, \ldots)$ of an integer K is a monotonically decreasing sequence of integers ($\kappa_1 \geq \kappa_2 \geq \cdots \geq 0$) where $\kappa_1 + \kappa_2 + \cdots = K$. The *union* $\tau = (\tau_1, \tau_2, \ldots)$ of two integer partitions κ and ν is defined as $\tau = \kappa \cup \nu$ where $\tau_1 \geq \tau_2 \geq \cdots$, i.e., τ is composed from all elements of κ and ν in such order that τ becomes monotonically decreasing. For example, the union of $(5,4,4,1)$ and $(4,2)$ is $(5,4,4,4,2,1)$. The *difference* τ of two integer partitions κ and ν is defined as $\tau = \kappa \setminus \nu$, where τ includes the elements from κ except elements existing in both κ and ν, which are removed. Notably, elements in ν not appearing in κ do not contribute to the difference. For example, the difference $(5,4,4,1) \setminus (4,2)$ is $(5,4,1)$. Furthermore, the *conjugate partition* of κ is defined as $\nu = \operatorname{conj}(\kappa)$, where ν_i is equal to the number of integers in κ that is equal or greater than i, for $i = 1, 2, \ldots$. For example, the conjugate of $(4,4,2,1)$ is $(4,3,2,2)$.

The *normal rank* of $G - \lambda H$, $\operatorname{nrk}(G - \lambda H)$, is the order of the matrix pencil's greatest minor different from polynomial zero [42]. Given the KCF of an $m_\mathrm{p} \times n_\mathrm{p}$ matrix pencil, we have

$$\operatorname{nrk}(G - \lambda H) = n_\mathrm{p} - r_0 = m_\mathrm{p} - l_0,$$

where r_0 and l_0 are the number of right and left singular blocks, respectively. The *null space* of an $m \times n$ matrix A is denoted by null(A), and is defined by $\text{null}(A) = \{x \in \mathbb{C}^n \mid Ax = 0\}$ [55]. The complementary space to null(A^H) is the *range* of A, denoted by ran(A), and is defined by $\text{ran}(A) = \{y \in \mathbb{C}^m \mid y = Ax \text{ for some } x \in \mathbb{C}^n\}$ [55]. In some literature, the null space and the range of A are called the *kernel* and the *image* of A, respectively.

The four invariants, column/row minimal indices and finite/infinite elementary divisors, are defined as follows [43]:

(i) The *column* (*right*) *minimal indices* are $\epsilon = (\epsilon_1, \ldots, \epsilon_{r_0})$, where

$$\epsilon_1 \geq \epsilon_2 \geq \cdots \geq \epsilon_{r_1} > \epsilon_{r_1+1} = \cdots = \epsilon_{r_0} = 0,$$

define the sizes of the L_{ϵ_k} blocks, $\epsilon_k \times (\epsilon_k + 1)$, and $r_0 = n_\mathrm{p} - \text{nrk}\,(G - \lambda H)$. The conjugate partition $r = (r_1, \ldots, r_{\epsilon_1}, 0, \ldots)$ of ϵ defines the *r-numbers* of the matrix pencil. From these we define the integer partition $\mathcal{R}(G - \lambda H) = (r_0) \cup (r_1, \ldots, r_{\epsilon_1})$, which in Section 5 is used to characterize the sizes of the L blocks. If there are no $\epsilon_k = 0$ (i.e., no L_0 blocks) it follows that $r_0 = r_1$ and $\epsilon = (\epsilon_1, \ldots, \epsilon_{r_1})$, and if there are no column minimal indices then $\epsilon = \emptyset$ and $\mathcal{R}(G - \lambda H) = (0, 0, \ldots) = (\mathbf{0})$.

(ii) The *row* (*left*) *minimal indices* are $\eta = (\eta_1, \ldots, \eta_{l_0})$, where

$$\eta_1 \geq \eta_2 \geq \cdots \geq \eta_{l_1} > \eta_{l_1+1} = \cdots = \eta_{l_0} = 0,$$

define the sizes of the $L^T_{\eta_k}$ blocks, $(\eta_k + 1) \times \eta_k$, and $l_0 = m_\mathrm{p} - \text{nrk}\,(G - \lambda H)$. The conjugate partition $l = (l_1, \ldots, l_{\eta_1}, 0, \ldots)$ of η defines the *l-numbers* of the matrix pencil, and analogously to the column minimal indices, we define the integer partition $\mathcal{L}(G - \lambda H) = (l_0) \cup (l_1, \ldots, l_{\eta_1})$, where $l_0 = l_1$ if there are no L^T_0 blocks. If there are no left minimal indices it follows that $\eta = \emptyset$ and $\mathcal{L}(G - \lambda H) = (\mathbf{0})$.

(iii) The *finite elementary divisors* are of the form

$$(\lambda - \mu_1)^{h_1^{(1)}}, \ldots, (\lambda - \mu_1)^{h_{g_1}^{(1)}}, \ldots, (\lambda - \mu_q)^{h_1^{(q)}}, \ldots, (\lambda - \mu_q)^{h_{g_q}^{(q)}},$$

with $h_1^{(i)} \geq \cdots \geq h_{g_i}^{(i)} \geq 1$ for each q distinct finite eigenvalue μ_i, $i = 1, \ldots, q$. Here g_i is the geometric multiplicity of the eigenvalue μ_i and the sum of all $h_k^{(i)}$ for $k = 1, \ldots, g_i$ is the algebraic multiplicity of μ_i. The exponents of the finite elementary divisors for eigenvalue μ_i are represented by the integer partition $h_{\mu_i} = (h_1^{(i)}, \ldots, h_{g_i}^{(i)}, 0, \ldots)$ which is known as the *Segre characteristics*. The Segre characteristics correspond to the sizes $h_k^{(i)} \times h_k^{(i)}$ of the Jordan blocks for eigenvalue μ_i, and also give the orders $h_k^{(i)}$ of the finite zero at μ_i of the associated LTI system (1.1). The conjugate partition of h_{μ_i}, $\mathcal{J}_{\mu_i}(G - \lambda H) = (j_1, j_2, \ldots)$, is known as the *Weyr characteristics* of μ_i. Consequently, we get $j_1 = g_i$ for each μ_i, $i = 1, \ldots, q$. For matrices it follows that $j_1 = \dim(\text{null}(A - \mu_i I))$, $j_1 + j_2 = \dim(\text{null}(A - \mu_i I)^2)$, etc. In other words, j_1 is the number of eigenvectors of μ_i and j_k corresponds to the number of principal vectors of grade $k \geq 2$. Moreover, the trailing zeros in both h_{μ_i} and $\mathcal{J}_{\mu_i}(G - \lambda H)$ are left out, except for situations when they are explicitly used.

(iv) The *infinite elementary divisors* are of the form

$$\rho^{s_1}, \rho^{s_2}, \ldots, \rho^{s_{g_\infty}},$$

with $s_1 \geq \cdots \geq s_{g_\infty} \geq 1$, where g_∞ is the geometric multiplicity of the infinite eigenvalue and the sum of all s_k for $k = 1, \ldots, g_\infty$ is the algebraic multiplicity. The exponents represented by the integer partition $s = (s_1, \ldots, s_{g_\infty}, 0, \ldots)$ is the *Segre characteristics* for the infinite eigenvalue, and correspond to the sizes $s_k \times s_k$ of the N_{s_k} blocks. The orders of the zeros at infinity of the associated LTI system (1.1) is $s_k - 1$ [112], i.e., an infinite elementary divisor of order one (a simple eigenvalue) makes no contribution to the zeros at infinity. In the same way as for finite eigenvalues, the conjugate integer partition $\mathcal{N}(G - \lambda H) = (n_1, n_2, \ldots)$ is the *Weyr characteristics* for the infinite eigenvalue, and the trailing zeros in s and $\mathcal{N}(G - \lambda H)$ are normally left out, except when needed.

When it is clear from context, we use the abbreviated notation $\mathcal{R}$, $\mathcal{L}$, $\mathcal{J}$, and $\mathcal{N}$, for the above defined integer partitions. In the following, these integer partitions are referred to as *structure integer partitions*. Moreover, the integer partitions representing the minimal indices and elementary divisors give the largest block first, but in block structure notation (see Section 2.3) it is not unusual that the blocks are given in reverse order, i.e., the smallest block first. This actually is the same order in which the conjugate partitions $\mathcal{R}$, $\mathcal{L}$, $\mathcal{J}$, and $\mathcal{N}$ are interpreted. For example, the integer partition $\mathcal{R} = (4, 3, 3, 1)$ is read as: there are $4 - 3 = 1$ L_0 block, $3 - 3 = 0$ L_1 blocks, $3 - 1 = 2$ L_2 blocks, and $1 - 0 = 1$ L_3 block. The corresponding KCF in block structure notation would then be $L_0 \oplus 2L_2 \oplus L_3$. However, to be consistent with KCF we use the decreasing order of the block sizes in this paper.

The system pencils $\mathbf{S}(\lambda)$, $\mathbf{S}_{\mathrm{C}}(\lambda)$, and $\mathbf{S}_{\mathrm{O}}(\lambda)$, can also be expressed in terms of the above invariants and their associated structure integer partitions. However, in general their corresponding invariants are different. For example, the system pencil $\mathbf{S}_{\mathrm{C}}(\lambda)$ of a completely controllable system associated with the pair (A, B) can only have L blocks in its KCF while $\mathbf{S}(\lambda)$ (1.3) may have both types of singular invariants (blocks) as well as eigenvalues in its KCF.

Example 2

Let us again consider the matrix pencil in Example 1 with KCF $2L_1 \oplus L_0^T \oplus J_2(\alpha)$.

As defined above, the minimal indices and the elementary divisors give the sizes of the corresponding blocks. For this matrix pencil where we have two L blocks of size one the column (right) minimal indices are $\epsilon = (1, 1)$. Moreover, it has one L^T block of size zero and therefore the row (left) minimal indices are $\eta = (0)$, and the single Jordan block of size 2×2 corresponds to the Segre characteristics $h_\alpha = (2)$ for the finite eigenvalue α. The matrix pencil has no infinite eigenvalues and therefore no infinite elementary divisors.

We can also represent the KCF of the matrix pencil by its structure integer partitions $\mathcal{R}$, $\mathcal{L}$, $\mathcal{J}$, and $\mathcal{N}$. We start with the right singular blocks, $2L_1$. The first integer in $\mathcal{R}$ is the number of L blocks of size zero or greater, the second integer is the number of L blocks of size one or greater, and so on. This results in

$$2L_1 \;\Rightarrow\; \mathcal{R} = (2, 2, 0, \ldots),$$

where the trailing zeros normally are left out.

In the same way, we get the structure integer partitions $\mathcal{L}$, $\mathcal{J}$, and $\mathcal{N}$, with the exception that the first element in the integer partitions $\mathcal{J}$ and $\mathcal{N}$ represent blocks of size one or greater. Altogether, the integer partitions representing the canonical structure of the matrix pencil are:

$$\begin{aligned} \mathcal{R} &= (2, 2), \\ \mathcal{L} &= (1), \text{ and} \\ \mathcal{J}_\alpha &= (1, 1). \end{aligned}$$

In addition, we also consider the following invariants associated with the matrix polynomial $A - \lambda I$ corresponding to the $n \times n$ matrix A [43, Vol. 1]. Denote by D_k the *greatest common divisors* of all the minors of order k of the linear matrix polynomial $A - \lambda I$. Let $D_0 = 1$ and $D_k \equiv 0$ if all the minors of order k of $A - \lambda I$ are zeros. Then the *invariant factors* of the matrix A are defined by the polynomials given from the quotients

$$P_1 = \frac{D_n}{D_{n-1}}, \quad P_2 = \frac{D_{n-1}}{D_{n-2}}, \quad \ldots, \quad P_n = \frac{D_1}{D_0} = D_1. \tag{2.11}$$

Furthermore, from the decomposition of the invariant factors into irreducible factors the finite elementary divisors are defined:

$$P_j = \prod_{i=1}^{q} (\lambda - \mu_i)^{h_j^{(i)}}, \quad j = 1, \ldots, n, \tag{2.12}$$

where $\mu_1, \ldots, \mu_q$ are distinct eigenvalues and the exponents $h_j^{(i)}$ are the Segre characteristics $h_{\mu_i} = (h_1^{(i)}, \ldots, h_{g_i}^{(i)}, 0, \ldots)$. For square matrices it follows that $\sum_i \sum_j h_j^{(i)} = n$. From (2.11) and (2.12) we can derive the following relation:

$$\begin{aligned} D_j &= P_n P_{n-1} \cdots P_{n+2-j} P_{n+1-j} \\ &= \prod_{i=1}^{q} (\lambda - \mu_i)^{\sum_{k=1}^{j} h_{n+1-k}^{(i)}}, \qquad j = 1, \ldots, n. \end{aligned} \tag{2.13}$$

For each finite elementary divisor $\lambda - \mu_i$, $i = 1, \ldots, q$, define

$$d_j^{(i)} = \text{the multiplicity of } \lambda - \mu_i \text{ in } D_j,$$

where the integer sequence $d_{\mu_i} = (d_0^{(i)}, \ldots, d_n^{(i)})$ is increasing, i.e., $d_j^{(i)} \leq d_{j+1}^{(i)}$ for $j = 0, \ldots, n-1$ [64]. Note that the exponent $h_j^{(i)}$ is the multiplicity of the finite elementary divisor $\lambda - \mu_i$ in P_j and, unlike h_{μ_i} which has n elements, d_{μ_i} has $n+1$ elements. Furthermore, $d_0^{(i)} = 0$ and

$$\sum_{k=1}^{j} h_k^{(i)} = d_n^{(i)} - d_{n-j}^{(i)}, \quad j = 1, \ldots, n, \tag{2.14}$$

for each eigenvalue μ_i.

Example 3

Consider a matrix of size 9×9 with JCF $J_4(\alpha) \oplus 2J_2(\alpha) \oplus J_1(\beta)$. The corresponding elementary divisors are

$$(\lambda - \alpha)^4, \ (\lambda - \alpha)^2, \ (\lambda - \alpha)^2, \text{ and } (\lambda - \beta),$$

and the invariant factors are

$$\begin{aligned} &P_1 = (\lambda - \alpha)^4 (\lambda - \beta), \\ &P_2 = (\lambda - \alpha)^2, \\ &P_3 = (\lambda - \alpha)^2, \text{ and} \\ &P_4 = \cdots = P_9 = 1. \end{aligned}$$

Consequently, the Segre characteristics for the matrix are $h_\alpha = (4,2,2,0,0,0,0,0,0)$ and $h_\beta = (1,0,0,0,0,0,0,0,0)$. From (2.13) we can now derive the greatest common divisors:

$$
\begin{aligned}
&D_0 = \cdots = D_6 = 1,\\
&D_7 = P_9 \cdots P_3 = (\lambda-\alpha)^2,\\
&D_8 = P_9 \cdots P_2 = (\lambda-\alpha)^2(\lambda-\alpha)^2, \text{ and}\\
&D_9 = P_9 \cdots P_1 = (\lambda-\alpha)^2(\lambda-\alpha)^2(\lambda-\alpha)^4(\lambda-\beta),
\end{aligned}
$$

which give the integer sequences $d_\alpha = (0,0,0,0,0,0,0,2,4,8)$ and $d_\beta = (0,0,0,0,0,0,0,0,0,1)$.

2.5 State-space transformations

To manipulate an LTI system in the time domain several different types of transformations are used. Here we present some of the more common ones for the state-space system (1.1) with the system pencil (1.3):

$$\mathbf{S}(\lambda) = \begin{bmatrix} A & B \\ C & D \end{bmatrix} - \lambda \begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix}.$$

We only consider *structure preserving* transformations, that is, transformations that do not destroy or change the special block structure of a system pencil. Moreover, we only consider the complex case, i.e., matrices with complex entries, but several of the transformations and conditions in the following also hold for the real case. For simplicity, we use the notation $A \in Gl_n(\mathbb{C})$ to denote that the complex matrix A is $n \times n$ and nonsingular (where $Gl_n(\mathbb{C})$ is the linear group of order n over $\mathbb{C}$).

A system pencil $\mathbf{S}(\lambda)$ of a matrix quadruple is said to be *feedback equivalent* [19, 91, 104] to $\widetilde{\mathbf{S}}(\lambda)$ if there exist a $P \in Gl_n(\mathbb{C})$, $T \in Gl_p(\mathbb{C})$, $Q \in Gl_m(\mathbb{C})$, $S \in \mathbb{C}^{n\times p}$ and an $R \in \mathbb{C}^{m\times n}$, such that the nonsingular transformation matrices U and V are

$$U = \begin{bmatrix} P & S \\ 0 & T \end{bmatrix} \quad \text{and} \quad V^{-1} = \begin{bmatrix} P^{-1} & 0 \\ R & Q^{-1} \end{bmatrix},$$

and

$$\widetilde{\mathbf{S}}(\lambda) = U\mathbf{S}(\lambda)V^{-1}.$$

The feedback equivalence for matrix quadruples is a generalization of the feedback equivalence for matrix pairs and is the product of six elementary transformations defined for matrix quadruples. They are:

$$
\begin{aligned}
\textit{left multiplication}: &\quad (\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D}) = (PA, PB, C, D),\\
\textit{state-coordinate}: &\quad (\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D}) = (AP^{-1}, B, CP^{-1}, D),\\
\textit{input-coordinate}: &\quad (\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D}) = (A, BQ^{-1}, C, DQ^{-1}),\\
\textit{state-feedback}: &\quad (\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D}) = (A+BR, B, C+DR, D),\\
\textit{output-coordinate}: &\quad (\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D}) = (A, B, TC, TD),\\
\textit{output-injection}: &\quad (\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D}) = (A+SC, B+SD, C, D).
\end{aligned}
\tag{2.15}
$$

Taken together, the left multiplication and state-coordinate transformations form a similarity transformation of the system matrix A and are sometimes referred to as a general state-space transformation. This is also one of the most common transformations of a state-space system. The last two transformations, on the other hand, are not of interest for the actual control problem. The output-coordinate transformation is (mainly) used for rescaling the output and the output-injection has no meaningful usage for the control problem but are of more theoretical use.

We can now state the following important property for the state-space systems (1.1) (which also holds for feedback equivalence of all independent subsystems of (1.1) [53, 19, 117]).

Theorem 2.1 [19] *Two matrix quadruples* (A, B, C, D) *and* $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ *are* feedback equivalent if and only if *the corresponding system pencils* $\mathbf{S}(\lambda)$ *and* $\widetilde{\mathbf{S}}(\lambda)$ *are strictly equivalent.*

Two generalized state-space systems are said to be *restricted system equivalent* [17, 101] if there exist two matrices $P \in Gl_q(\mathbb{C})$ and $Z \in Gl_n(\mathbb{C})$ such that

$$\begin{bmatrix} P & 0 \\ 0 & I_p \end{bmatrix} \begin{bmatrix} A - \lambda E & B \\ C & D \end{bmatrix} \begin{bmatrix} Z^{-1} & 0 \\ 0 & I_m \end{bmatrix} = \begin{bmatrix} P(A - \lambda E)Z^{-1} & PB \\ CZ^{-1} & D \end{bmatrix}.$$

where $A, E \in \mathbb{C}^{q \times n}$.

For the controllability pair (A, B) the transformations 1 to 4 in (2.15) are applicable. Taken together, these transformations define the feedback equivalence for controllability pairs:

$$P \begin{bmatrix} A - \lambda I_n & B \end{bmatrix} \begin{bmatrix} P^{-1} & 0 \\ R & Q^{-1} \end{bmatrix} = \begin{bmatrix} P\left(A - \lambda I_n\right)P^{-1} + PBR & PBQ^{-1} \end{bmatrix}, \qquad (2.16)$$

where $P \in Gl_n(\mathbb{C}),\ Q \in Gl_m(\mathbb{C})$ and $R \in \mathbb{C}^{m \times n}$ (e.g., see [12, 59, 117]). Other names that appear in the literature for this equivalence relation are *block similar* [53] and Γ*-equivalence* [19].

These transformations can also be motivated from control theory. By adding a linear feedback $u(t) = Fx(t) + v(t)$ to $\dot{x}(t) = Ax(t) + Bu(t)$ we obtain the system

$$\dot{x}(t) = (A + BF)x(t) + Bv(t).$$

Then we introduce new coordinates $P^{-1}x(t)$ and $Q^{-1}v(t)$ for the state and input variables leading to a transformed system (with the same system behaviors):

$$\dot{x}(t) = P(A + BF)P^{-1}x(t) + PBQ^{-1}v(t),$$

which in matrix form is

$$P \begin{bmatrix} A - \lambda I_n & B \end{bmatrix} \begin{bmatrix} P & 0 \\ -QF & Q \end{bmatrix}^{-1} = P \begin{bmatrix} A - \lambda I_n & B \end{bmatrix} \begin{bmatrix} P^{-1} & 0 \\ FP^{-1} & Q^{-1} \end{bmatrix}.$$

By substituting FP^{-1} with R we obtain the equivalence transformation (2.16).

For the observability pair (A, C) the corresponding transformations are 1–2 and 5–6 in (2.15), which together define the feedback equivalence for observability pairs:

$$\begin{bmatrix} P & S \\ 0 & T \end{bmatrix} \begin{bmatrix} A - \lambda I_n \\ C \end{bmatrix} P^{-1} = \begin{bmatrix} P\left(A - \lambda I_n\right)P^{-1} + SCP^{-1} \\ TCP^{-1} \end{bmatrix},$$

where $P \in Gl_n(\mathbb{C}),\ T \in Gl_p(\mathbb{C})$ and $S \in \mathbb{C}^{n\times p}$.

For generalized matrix pairs (E, A, B), where $E,\ A \in \mathbb{C}^{q\times n}$ and $B \in \mathbb{C}^{q\times m}$, the restricted system equivalence, the input-coordinate and state-feedback transformations form the group of *proportional feedback transformations*, where two systems are said to be feedback equivalent (e.g., see [67, 86]). With the addition of the derivative-feedback transformation we have the group of *proportional* plus *derivative feedback transformations* where two systems are called pd-feedback equivalent (e.g., see [67, 86, 118]). To express the proportional plus derivative feedback transformations we need to separate the pencil $A - \lambda E$ into their A- and λ-parts, respectively, and apply a 3×3 block matrix from the right. For consistency, we also express the proportional feedback transformation in the same form.

Two generalized matrix pairs (E, A, B) and $(\widetilde{E}, \widetilde{A}, \widetilde{B})$ are *feedback equivalent* if there exists an equivalence transformation by two nonsingular matrices such that

$$\begin{aligned} &P\begin{bmatrix}-E & A & B\end{bmatrix}\begin{bmatrix}Z^{-1} & 0 & 0\\ 0 & Z^{-1} & 0\\ 0 & F_P Z^{-1} & Q^{-1}\end{bmatrix}\\ &\quad= \begin{bmatrix}-PEZ^{-1} & P(A + BF_P)Z^{-1} & PBQ^{-1}\end{bmatrix}\\ &\quad\equiv \begin{bmatrix}\widetilde{E} & \widetilde{A} & \widetilde{B}\end{bmatrix}.\end{aligned}$$

Two generalized matrix pairs (E, A, B) and $(\widetilde{E}, \widetilde{A}, \widetilde{B})$ are *pd-feedback equivalent* if there exists an equivalence transformation by two nonsingular matrices such that

$$\begin{aligned} &P\begin{bmatrix}-E & A & B\end{bmatrix}\begin{bmatrix}Z^{-1} & 0 & 0\\ 0 & Z^{-1} & 0\\ F_D Z^{-1} & F_P Z^{-1} & Q^{-1}\end{bmatrix}\\ &\quad= \begin{bmatrix}-P(E + BF_D)Z^{-1} & P(A + BF_P)Z^{-1} & PBQ^{-1}\end{bmatrix}\\ &\quad\equiv \begin{bmatrix}\widetilde{E} & \widetilde{A} & \widetilde{B}\end{bmatrix}.\end{aligned}$$

All the above transformations preserve the structure of (E, A, B), but when $q = n$ and $\det(A - \lambda E) \not\equiv 0$ the state-feedback and derivative-feedback transformations can destroy the regularity condition $\det(A - \lambda E) \not\equiv 0$ [66].

2.6 Brunovsky canonical form and generalizations

When considering canonical forms of the system pencil $\mathbf{S}(\lambda)$ associated with pairs, triples or quadruples of matrices, we are (mainly) interested in canonical forms obtained from structure-preserving transformations, see Section 2.5. One such example is the Brunovsky canonical form and its generalizations. These canonical forms explicitly reveal the system characteristics from the system pencils. This is in contrast to the KCF, which destroys the special block structure of $\mathbf{S}(\lambda)$ and only implicitly gives the system characteristics. Canonical forms for generalized state-space (or descriptor) systems studied in, e.g., [52, 63, 86, 105] are not considered in this paper.

Brunovsky formulated in 1970 a canonical form for completely controllable matrix pairs [12] (the results were published already in 1966 in a Russian article). He also derived the

r-numbers for a matrix pair (A, B) as[2] [12, 59]:

$$\begin{aligned} r_1 &= \operatorname{rank}(B), \\ r_j &= \operatorname{rank}\big(B, AB, \ldots, A^{j-1}B\big) - \operatorname{rank}\big(B, AB, \ldots, A^{j-2}B\big), \quad j = 2, \ldots, n. \end{aligned}$$

Kalman [81] pointed out that the Brunovsky invariants are equivalent to those of Kronecker [82] (see Section 2.4). The canonical form defined by Brunovsky has later been revised to include uncontrollable matrix pairs, see for example [53, Theorem 6.2.5] and [117, Theorem 2.11].

Given a matrix pair (A, B) associated with the state-space model

$$\dot{x}(t) = Ax(t) + Bu(t),$$

which does not need to be completely controllable, there exists a feedback equivalent matrix pair (A_B, B_B) in *Brunovsky canonical form* (BCF), such that

$$P\begin{bmatrix} A - \lambda I_n & B \end{bmatrix}\begin{bmatrix} P^{-1} & 0 \\ R & Q^{-1} \end{bmatrix} = \begin{bmatrix} A_B - \lambda I_n & B_B \end{bmatrix} = \left[\begin{array}{cc|cc} A_\epsilon & 0 & B_\epsilon & 0 \\ 0 & A_\mu & 0 & 0 \end{array}\right]. \tag{2.17}$$

The matrix pair (A_ϵ, B_ϵ) is controllable and the regular pencil A_μ consists of the *uncontrollable eigenvalues* (*modes*). Moreover, the column minimal indices of (A_ϵ, B_ϵ) are known as the *controllability indices* of (A, B). The next result follows immediately from Theorem 2.1.

Theorem 2.2 [117] *Two controllability pairs* (A, B) *and* $(\widetilde{A}, \widetilde{B})$ *are* feedback equivalent if and only if *they have the same controllability indices and finite elementary divisors, i.e., they are strictly equivalent.*

The dual form of BCF for the matrix pair (A, C) is

$$\begin{bmatrix} P & S \\ 0 & T \end{bmatrix}\begin{bmatrix} A - \lambda I_n \\ C \end{bmatrix}P^{-1} = \begin{bmatrix} A_B - \lambda I_n \\ C_B \end{bmatrix} = \left[\begin{array}{cc} A_\eta & 0 \\ 0 & A_\mu \\ \hline C_\eta & 0 \\ 0 & 0 \end{array}\right], \tag{2.18}$$

where (A_η, C_η) is observable and A_μ is regular and consists of the *unobservable eigenvalues* (*modes*). The row minimal indices of (A_η, C_η) are known as the *observability indices* of (A, C).

The BCF of a matrix pair is a special case of a more general canonical form proposed independently by Morse [91] for matrix triples and by Thorp [104] for matrix quadruples. This canonical from is defined as follows (see also [93] where a canonical form under similarity transformations is derived).

Let (A, B, C, D) be a matrix quadruple associated with the state-space model

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t). \end{aligned}$$

Moreover, let $\mathbf{S}(\lambda)$ be the associated system pencil with the following invariants:

- The column minimal indices $(\epsilon_1, \ldots, \epsilon_{r_1}, \epsilon_{r_1+1} \ldots, \epsilon_{r_0})$.
- The row minimal indices $(\eta_1, \ldots, \eta_{l_1}, \eta_{l_1+1} \ldots, \eta_{l_0})$.

[2]The l-numbers of the matrix pair (A, C) can similarly be determined from its observability matrix.

- The Segre characteristics $(h_1^{(i)}, \ldots, h_{g_i}^{(i)})$ for the finite eigenvalue μ_i, $i = 1, \ldots, q$ (the exponents of the finite elementary divisors).
- The Segre characteristics $(s_1, \ldots, s_{g_\infty})$ for the infinite eigenvalue (the exponents of the infinite elementary divisors). Let $\delta_i = s_i - 1$, such that $\delta_1 \geq \cdots \geq \delta_t > \delta_{t+1} = \cdots = \delta_{g_\infty} = 0$. The integer partition $(\delta_1, \ldots, \delta_t)$ corresponds to the t zeros at infinity of the associated LTI system, i.e., the t N_k blocks of size $k \geq 2$.

Alternatively, the system pencil $\mathbf{S}(\lambda)$ can be expressed in terms of the structure integer partitions $\mathcal{R}$, $\mathcal{L}$, $\mathcal{J}$ and $\mathcal{N}$ associated with the invariants above.

Now, there exists a feedback equivalence transformation of $\mathbf{S}(\lambda)$ such that

$$\begin{bmatrix} P & S \\ 0 & T \end{bmatrix} \begin{bmatrix} A - \lambda I_n & B \\ C & D \end{bmatrix} \begin{bmatrix} P^{-1} & 0 \\ R & Q^{-1} \end{bmatrix} = \begin{bmatrix} A_B - \lambda I_n & B_B \\ C_B & D_B \end{bmatrix}, \tag{2.19}$$

where (A_B, B_B, C_B, D_B) is in *generalized Brunovsky canonical form* (GBCF) [19, 90, 91, 104], defined by

$$
\left[\begin{array}{c|c} A_B - \lambda I_n & B_B \\ \hline C_B & D_B \end{array}\right] = \left[\begin{array}{cccc|cccc}
A_\epsilon & 0 & 0 & 0 & B_\epsilon & 0 & 0 & 0 \\
0 & A_\eta & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & A_\infty & 0 & 0 & B_\infty & 0 & 0 \\
0 & 0 & 0 & A_\mu & 0 & 0 & 0 & 0 \\ \hline
0 & C_\eta & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & C_\infty & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & D_\infty & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & D_0
\end{array}\right] =
$$

$$
= \left[\begin{array}{cccc|cccc}
\begin{smallmatrix} J_{\epsilon_1}(0) & & \\ & \ddots & \\ & & J_{\epsilon_{r_1}}(0) \end{smallmatrix} & 0 & 0 & 0 & \begin{smallmatrix} e_{\epsilon_1} & & \\ & \ddots & \\ & & e_{\epsilon_{r_1}} \end{smallmatrix} & 0 & 0 & 0 \\
0 & \begin{smallmatrix} J^T_{\eta_1}(0) & & \\ & \ddots & \\ & & J^T_{\eta_{l_1}}(0) \end{smallmatrix} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \begin{smallmatrix} J^T_{\delta_1}(0) & & \\ & \ddots & \\ & & J^T_{\delta_t}(0) \end{smallmatrix} & 0 & 0 & \begin{smallmatrix} f_{\delta_1} & & \\ & \ddots & \\ & & f_{\delta_t} \end{smallmatrix} & 0 & 0 \\
0 & 0 & 0 & \begin{smallmatrix} J(\mu_1) & & \\ & \ddots & \\ & & J(\mu_q) \end{smallmatrix} & 0 & 0 & 0 & 0 \\ \hline
0 & \begin{smallmatrix} e^T_{\eta_1} & & \\ & \ddots & \\ & & e^T_{\eta_{l_1}} \end{smallmatrix} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \begin{smallmatrix} e^T_{\delta_1} & & \\ & \ddots & \\ & & e^T_{\delta_t} \end{smallmatrix} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & I_{g_\infty - t} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right],
$$

where the J_k blocks are nilpotent matrices in its reduced Jordan form (2.7), D_0 is an $(l_0 - l_1) \times (r_0 - r_1)$ zero matrix,

$$
e_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \in \mathbb{C}^{i \times 1}, \qquad \text{and} \qquad f_i = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{C}^{i \times 1}.
$$

In the GBCF, the matrix pair (A_ϵ, B_ϵ) is controllable and corresponds to the L_k blocks, $k \geq 1$, in the KCF of $\mathbf{S}(\lambda)$. Similarly, the matrix pair (A_η, C_η) is observable and corresponds to the L_k^T blocks, $k \geq 1$. Moreover, the matrix A_μ corresponds to all Jordan blocks of the finite eigenvalues, where each block $J(\mu_i)$ in A_μ is block diagonal with the Jordan blocks for the specified finite eigenvalue μ_i. The matrix D_∞ corresponds to the N_1 blocks and the remaining parts, formed by

$$\begin{bmatrix} A_\infty & B_\infty \\ C_\infty & 0 \end{bmatrix},$$

correspond to the N_k blocks, $k \geq 2$. Furthermore, the number of columns in B_ϵ corresponds to the number of L_k blocks with $k \geq 1$, likewise, the number of rows in C_η corresponds to the number of L_k^T blocks with $k \geq 1$, and the number of columns in B_∞ or rows in C_∞ is the number of N blocks of size greater than one. The $(l_0 - l_1) \times (r_0 - r_1)$ matrix D_0 constitutes of the vectors $e_{\epsilon_{r_1+1}}, \ldots, e_{\epsilon_{r_0}}$ of size 0×1 and the vectors $e_{\eta_{l_1+1}}^T, \ldots, e_{\eta_{l_0}}^T$ of size 1×0. It follows that the number of columns of D_0 corresponds to the number of L_0 blocks, and the number of rows of D_0 corresponds to the number of L_0^T blocks.

Several system properties are reveled directly from the GBCF. The GBCF is composed of five decoupled subsystems as follows [7]:

(1) The subsystem $(A_\infty, B_\infty, C_\infty)$ corresponding to the infinite zero structure. It is controllable and observable and consists of decoupled chains of integrators with inputs and outputs.

(2) The subsystem (A_ϵ, B_ϵ) corresponding to the column minimal indices. It is controllable but not observable and consists of decoupled chains of integrators with inputs, but no outputs.

(3) The subsystem (A_η, C_η) corresponding to the row minimal indices. It is observable but not controllable and consists of decoupled chains of integrators with outputs, but no inputs.

(4) The subsystem A_μ corresponding to the finite structure. It is uncontrollable and unobservable and consists of the finite zeros of the original system.

(5) The "feedforward" subsystem $\operatorname{diag}(D_\infty, D_0)$. This subsystem passes $g_\infty - t$ inputs unchanged to $g_\infty - t$ outputs, annihilates $r_0 - r_1$ inputs, and generates $l_0 - l_1$ identically zero outputs.

A matrix triple is a special case of a matrix quadruple, where D in (2.19) is the zero matrix [19, 91]. For an LTI system it means that we have no feedforward matrix and consequently the subsystem (5) described above is absent. It follows that a matrix triple can have no infinite elementary divisors of order one, i.e., no N_1 blocks. Apart from this restriction, the invariants and the GBCF are the same as for matrix quadruples.

As said in the beginning of this section, it follows that the BCF for a matrix pair (A, B) (and (A, C)) is a subset of GBCF. The BCF (2.17) for (A, B) only includes the blocks A_ϵ, A_μ and B_ϵ. Similarly, the BCF (2.18) for (A, C) only includes the blocks A_η, A_μ and C_η. A consequence is that matrix pairs cannot have infinite eigenvalues (N blocks). Moreover, the controllability pair (A, B) has exactly m L blocks and the observability pair (A, C) has p L^T blocks. This can be verified from the fact that the controllability system pencil

$$\mathbf{S}_{\mathrm{C}}(\lambda) = \begin{bmatrix} A & B \end{bmatrix} - \lambda \begin{bmatrix} I_n & 0 \end{bmatrix}$$

has full row rank, i.e., the system pencil can have no left singular blocks (L^T blocks), and the number of columns in B_ϵ is equal to m. Similarly, the observability system pencil

$$\mathbf{S}_{\mathrm{O}}(\lambda) = \begin{bmatrix} A \\ C \end{bmatrix} - \lambda \begin{bmatrix} I_n \\ 0 \end{bmatrix}$$

has full column rank and therefore has no right singular blocks (L blocks), and the number of rows in C_η is equal to p.

Furthermore, the following system characteristics are revealed from the BCF of the controllability pair (A, B) and observability pair (A, C). If $\text{rank}(\mathbf{S}_\text{C}(\lambda)) < n$ for some $\lambda \in \mathbb{C}$ then (A, B) is uncontrollable and there exists a regular pencil A_μ whose eigenvalues correspond to the *uncontrollable eigenvalues* (*modes*). Likewise, if $\text{rank}(\mathbf{S}_\text{O}(\lambda)) < n$ for some $\lambda \in \mathbb{C}$ then (A, C) is unobservable and there exists a regular pencil A_μ whose eigenvalues correspond to the *unobservable eigenvalues* (*modes*). The number of L_0 blocks of (A, B) is $m-\text{rank}(B_B)$, where for each L_0 block one input signal $u_k(t)$ can be removed without loosing controllability of (A_ϵ, B_ϵ).[3] Likewise, the number L_0^T blocks of (A, C) is $p-\text{rank}(C_B)$, where for each L_0^T block one output signal $y_k(t)$ can be removed without loosing observability of (A_η, C_η).

Example 4

To exemplify the Brunovsky canonical form and its generalization we consider a state-space system with two states, three inputs and one output:

$$\begin{aligned} \dot{x}(t) &= \begin{bmatrix} 0 & 0 \\ -3 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 3 & 10 & 1 \\ 0.6 & 2 & 0.2 \end{bmatrix} u(t), \\ y(t) &= \begin{bmatrix} 0.6 & \gamma \end{bmatrix} x(t), \end{aligned} \tag{2.20}$$

where $\gamma > 0$. The system has the KCF $2L_0 \oplus J_1(\alpha) \oplus N_2$ with the corresponding GBCF

$$\mathbf{S}(\lambda) = \left[\begin{array}{cc|ccc} -\lambda & 0 & 0 & 0 & 1 \\ 0 & \alpha-\lambda & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 \end{array}\right],$$

where the finite eigenvalue α depends on the value of γ.

By inspecting the subsystems $\mathbf{S}_\text{C}(\lambda)$ and $\mathbf{S}_\text{O}(\lambda)$ of $\mathbf{S}(\lambda)$, we can derive the controllability and observability characteristics of the system. The controllability pair in BCF is

$$\mathbf{S}_\text{C}(\lambda) = \left[\begin{array}{cc|ccc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right] - \lambda \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right],$$

and has the KCF $L_2 \oplus 2L_0$, so the system is controllable. The observability pair in BCF is

$$\mathbf{S}_\text{O}(\lambda) = \left[\begin{array}{cc} 0 & 0 \\ 1 & 0 \\ \hline 0 & 1 \end{array}\right] - \lambda \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ \hline 0 & 0 \end{array}\right],$$

and has the KCF L_2^T, i.e., the system is also observable. Here we could be satisfied with knowing that the system is both controllable and observable. However, if we look at the system pencil of the observability pair

$$\mathbf{S}_\text{O}(\lambda) = \left[\begin{array}{cc} 0 & 0 \\ -3 & 0 \\ \hline 0.6 & \gamma \end{array}\right] - \lambda \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ \hline 0 & 0 \end{array}\right],$$

[3]However, for safety reasons it is customary to have redundancy in the actuation system and the corresponding control surface in critical systems.

we can see that the observability depends on the value of γ. As long as $\gamma > 0$ the system is observable, but when γ approaches 0 the observability pencil becomes closer and closer to being unobservable. Finally, when γ reaches zero the KCF of the observability pencil is $L_1^T \oplus J_1(0)$ with BCF

$$\mathbf{S}_{\mathrm{O}}(\lambda) = \left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \\ \hdashline 1 & 0 \end{array}\right] - \lambda \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ \hdashline 0 & 0 \end{array}\right],$$

which corresponds to an unobservable system with one unobservable mode at zero.

Even if the computed canonical structure is observable the original system may be unobservable or close to, since a zero element (e.g., γ in the above example) can become nonzero because of roundoff errors in the numerical methods or noise in the data. This is the reason why it is important to know the distance to the closest unobservable system (or uncontrollable system), or even better, to know all possible canonical structures which can be reached by a small perturbation and the distance to each of them.

2.7 Relation between KCF and GBCF

By comparing the KCF (2.6) and the GBCF (2.19) associated with the same system, we see that the two canonical forms are closely related to each other [81, 91, 104]. More precisely they are permutations of each other. From a general matrix pencil $G - \lambda H$ of size $m_{\mathrm{p}} \times n_{\mathrm{p}}$ in KCF the corresponding system pencil $S - \lambda T$ of size $(n+p) \times (n+m)$ in GBCF can be computed as $P_{row}(G - \lambda H)P_{col} = S - \lambda T$, where $m_{\mathrm{p}} = n + p$, $n_{\mathrm{p}} = n + m$, and P_{row} and P_{col} are permutation matrices of rows and columns, respectively (the identity matrix of conforming size with its rows or columns reordered). To get exactly the same form of $S - \lambda T$ as in (2.19), the blocks in the KCF of the matrix pencil $G - \lambda H$ must be ordered as (compare with (2.6)):

$$\begin{aligned} &G - \lambda H \\ &\quad = \mathrm{diag}(L_{\epsilon_1}, \ldots, L_{\epsilon_{r_0}}, L_{\eta_1}^T, \ldots, L_{\eta_{l_0}}^T, N_{s_1}, \ldots, N_{s_{g_\infty}}, J_{h_1}(\mu_1), \ldots, J_{h_{g_q}}(\mu_q)). \end{aligned} \tag{2.21}$$

The following two algorithms together determine permutation matrices P_{row} and P_{col} such that $P_{row}(G - \lambda H)P_{col}$ is in GBCF, where $G - \lambda H$ is assumed to be in the form (2.21).

Algorithm 1

The row-permutation matrix P_{row} is derived from $G - \lambda H$ ($m_{\mathrm{p}} \times n_{\mathrm{p}}$) in the form (2.21) with the following steps:

1. Create a matrix G_{row}, initially of size $m_{\mathrm{p}} \times n_{\mathrm{p}}$, where the elements corresponding to nonzero elements in G are all set to ones and the remaining elements are set to zeros.
2. Create a matrix H_{row}, initially of size $m_{\mathrm{p}} \times n_{\mathrm{p}}$, where the elements corresponding to nonzero elements in H are all set to ones and the remaining elements are set to zeros.

3. Set all rows of G_{row} to zero where the corresponding rows in H_{row} have a nonzero entry.
4. Let a and b be the number of columns in G_{row} and H_{row}, respectively, with only zero entries. Remove these zero columns such that G_{row} becomes $m_\mathrm{p} \times (n_\mathrm{p} - a)$ and H_{row} becomes $m_\mathrm{p} \times (n_\mathrm{p} - b)$.
5. Set $P_{row} = \begin{bmatrix} H_{row}^T \\ Q_{row} \\ G_{row}^T \end{bmatrix}$, where P_{row} is $m_\mathrm{p} \times m_\mathrm{p}$, and Q_{row} is $(m_\mathrm{p} - 2n_\mathrm{p} + a + b) \times m_\mathrm{p}$ and chosen such that P_{row} becomes a permutation matrix.

Algorithm 2

The column-permutation matrix P_{col} is derived from $G - \lambda H$ $(m_\mathrm{p} \times n_\mathrm{p})$ in the form (2.21) with the following steps:

1. Create a matrix G_{col}, initially of size $m_\mathrm{p} \times n_\mathrm{p}$, where the elements corresponding to nonzero elements in G are all set to ones and the remaining elements are set to zeros.
2. Create a matrix H_{col}, initially of size $m_\mathrm{p} \times n_\mathrm{p}$, where the elements corresponding to nonzero elements in H are all set to ones and the remaining elements are set to zeros.
3. Set all columns of G_{col} to zero where the corresponding columns in H_{col} have a nonzero entry.
4. Let c and d be the number of rows in G_{col} and H_{col}, respectively, with only zero entries. Remove these zero rows such that G_{col} becomes $(m_\mathrm{p} - c) \times n_\mathrm{p}$ and H_{col} becomes $(m_\mathrm{p} - d) \times n_\mathrm{p}$.
5. Set $P_{col} = [\, H_{col}^T \; Q_{col} \; G_{col}^T \,]$, where P_{col} is $n_\mathrm{p} \times n_\mathrm{p}$, and Q_{col} is $n_\mathrm{p} \times (n_\mathrm{p} - 2m_\mathrm{p} + c + d)$ and chosen such that P_{col} becomes a permutation matrix.

The matrices Q_{row} and Q_{col} can be chosen as zero matrices, since the 1's added in this step have no relevance for the permutation from KCF to GBCF (they correspond to permutations of rows or columns of zeros and are included only to make P_{row} and P_{col} permutation matrices). In step 4 of both algorithms, the dimensions $n_\mathrm{p} - b$ and $m_\mathrm{p} - d$, respectively, are equal to the number of states n of the system. Furthermore, for the controllability pair (A, B), $P_{row} = I_{m_\mathrm{p}}$, and for the observability pair (A, C), $P_{col} = I_{n_\mathrm{p}}$.

Proof of Algorithm 1. Let the $m_\mathrm{p} \times n_\mathrm{p}$ matrix pencil $G - \lambda H$ be in KCF with the block-order of (2.21) and with the corresponding $(n + p) \times (n + m)$ system pencil

$$S - \lambda T = \begin{bmatrix} A & B \\ C & D \end{bmatrix} - \lambda \begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix},$$

in GBCF, where $m_\mathrm{p} = n + p$ and $n_\mathrm{p} = n + m$. Note that $G - \lambda H$ and $S - \lambda T$ are associated with the same system and therefore have the same number of states n, i.e., the same number of nonzero columns in H and T, respectively. First consider the case when, in addition to the eigenvalues, $G - \lambda H$ and $S - \lambda T$ only have entries of ones and zeros.

We want to show that $P_{row}(G - \lambda H)P_{col} = S - \lambda T$, i.e.,

$$P_{row} G P_{col} = S, \tag{2.22}$$

and

$$P_{row} H P_{col} = T, \tag{2.23}$$

where P_{row} is the $(n+p) \times (n+p)$ row-permutation matrix constructed with Algorithm 1, and P_{col} is an $(n+m) \times (n+m)$ column-permutation matrix. We first consider (2.23) in part 1 and then (2.22) in part 2.

Part 1: Rewrite (2.23) as

$$H P_{col} = P_{row}^T T, \tag{2.24}$$

and let P_{row}^T be partitioned as

$$P_{row}^T = \begin{bmatrix} U_{row}^T & L_{row}^T \end{bmatrix}, \tag{2.25}$$

where U_{row}^T is $(n+p) \times n$ and L_{row}^T is $(n+p) \times p$. From the structure of T it follows that the row-permutation elements in P_{row}^T acting on T in (2.24) must all be in U_{row}^T, and we only need to consider the subproblem

$$H P_{col} = \begin{bmatrix} U_{row}^T & 0_{(n+p)\times p} \end{bmatrix} \begin{bmatrix} I_n & 0_{n\times m} \\ 0_{p\times n} & 0_{p\times m} \end{bmatrix} = \begin{bmatrix} U_{row}^T & 0_{(n+p)\times m} \end{bmatrix}.$$

Take P_{col} such that all m zero columns in H are moved to the trailing columns, i.e.,

$$H P_{col} = \begin{bmatrix} \widetilde{H} & 0_{(n+p)\times m} \end{bmatrix} = \begin{bmatrix} U_{row}^T & 0_{(n+p)\times m} \end{bmatrix}, \tag{2.26}$$

where $\widetilde{H}$ still has the same order of the nonzero columns as H (if $m = 0$ then $\widetilde{H} = H$). From (2.25) and (2.26) we now get that

$$P_{row} = \begin{bmatrix} \widetilde{H}^T \\ L_{row} \end{bmatrix}.$$

If P_{row} is taken as above with L_{row} as the zero matrix, then (2.23) is satisfied except for a column permutation (regardless of what P_{col} is). If $p = 0$, i.e., $G - \lambda H$ corresponds to a controllability pair (A, B), the proof of Algorithm 1 for P_{row} is complete (L_{row} is then $0 \times n$). Moreover, since the order of the L and J blocks are the same in the KCF and the GBCF it follows that $P_{row} = \widetilde{H}^T = I_n$. If $p > 0$, then continue with part 2. We remark that the above $\widetilde{H}$ is identical to H_{row} after step 4 in Algorithm 1.

Part 2: Now we consider the remaining part, L_{row}, of P_{row}. Equation (2.22) is equal to

$$P_{row} G P_{col} = \begin{bmatrix} U_{row} \\ L_{row} \end{bmatrix} G P_{col} = S. \tag{2.27}$$

Consequently, for each nonzero column in U_{row} the corresponding column in L_{row} must be zero, i.e., they cannot affect the same rows in G. Let $P_{col} = I_{n+m}$ and split G such that $G = G_1 + G_2$, where G_1 consists of the rows corresponding to nonzero columns in U_{row} (same as the nonzero rows in H) and G_2 consists of the rows corresponding to zero columns in U_{row}. Note that all eigenvalues of $G - \lambda H$ will be in G_1. The problem can now be rewritten as

$$\begin{bmatrix} U_{row} \\ 0 \end{bmatrix} G_1 + \begin{bmatrix} 0 \\ L_{row} \end{bmatrix} G_2 = \begin{bmatrix} S_1 \\ S_2 \end{bmatrix}.$$

Since U_{row} was determined in (2.26), we only need to consider the subproblem

$$L_{row} G_2 = S_2.$$

Let an $(n+p) \times q$ matrix $\widetilde{G}_2$ consist of only the q nonzero columns of G_2 (given in the same order). It follows that $q \leq p$. Take

$$L_{row} = \begin{bmatrix} 0 \\ \widetilde{G}_2^T \end{bmatrix}, \tag{2.28}$$

then $L_{row} G_2$ moves all nonzero rows in G_2 to the last q rows. The $p-q$ rows of zeros above $\widetilde{G}_2^T$ in L_{row} correspond to permutations of L_0^T blocks, i.e., rows of zeros in G for which there are no need to determine explicit permutations. Since the order of the blocks in (2.21) are the same as in GBCF, it follows that if L_{row} is chosen as in (2.28) and U_{row} as in step 1 then (2.27) is satisfied except for a column permutation (which is determined by Algorithm 2). Notably, the matrix $\widetilde{G}_2$ in (2.28) is identical to G_{row} after step 4 in Algorithm 1, and the rows of zeros above correspond to Q_{row} in step 5.

For the *general case* where G and H can have elements with values other than one or zero (additionally to the eigenvalues), we take $H_{row} \equiv \widetilde{H}$ and $G_{row} \equiv \widetilde{G}_2$ where the corresponding elements in H_{row} and G_{row} are set to one when $\widetilde{H}$ and $\widetilde{G}_2$ have a nonzero element, respectively. Let

$$P_{row} = \begin{bmatrix} H_{row}^T \\ Q_{row} \\ G_{row}^T \end{bmatrix},$$

where Q_{row} is chosen such that P_{row} becomes a permutation matrix. Then $P_{row}(G - \lambda H) I_{n+m}$ is equal to $S - \lambda T$ up to a column permutation, which is determined by P_{col} in Algorithm 2. $\square$

Proof of Algorithm 2. The proof of Algorithm 2 is similar to that of Algorithm 1. $\square$

Example 5

A Matlab function kcf2gbcf has been developed that given the KCF of a matrix pencil returns the GBCF and the permutation matrices which transform the matrix pencil in KCF (2.21) to GBCF.

Consider the 5×6 general matrix pencil given in Example 1 with KCF $2L_1 \oplus L_0^T \oplus J_2(\alpha)$, where the blocks have been reordered as in (2.21):

$$G - \lambda H = \begin{bmatrix} 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \boldsymbol{\alpha} & \mathbf{1} \\ 0 & 0 & 0 & 0 & 0 & \boldsymbol{\alpha} \end{bmatrix} - \lambda \begin{bmatrix} \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} \end{bmatrix}.$$

Let $\alpha = -5$. Then the output from the Matlab function is:

```
>> [S,T,Prow,Pcol] = kcf2gbcf(G,H)
S =
     0     0     0     0     1     0
     0     0     0     0     0     1
     0     0    -5     1     0     0
     0     0     0    -5     0     0
     0     0     0     0     0     0
```

```
T =
     1     0     0     0     0     0
     0     1     0     0     0     0
     0     0     1     0     0     0
     0     0     0     1     0     0
     0     0     0     0     0     0

Prow =
     1     0     0     0     0
     0     1     0     0     0
     0     0     0     1     0
     0     0     0     0     1
     0     0     1     0     0

Pcol =
     1     0     0     0     0     0
     0     0     0     0     1     0
     0     1     0     0     0     0
     0     0     0     0     0     1
     0     0     1     0     0     0
     0     0     0     1     0     0
```

We now show step by step how the permutation matrices P_{row} and P_{col} are constructed using Algorithms 1 and 2, respectively.

Algorithm 1: First we construct the matrices G_{row} and H_{row} from G and H, respectively (steps 1 and 2):

$$G_{row} = \begin{bmatrix} 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} & \mathbf{1} \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} \end{bmatrix}, \quad \text{and} \tag{2.29}$$

$$H_{row} = \begin{bmatrix} \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} \end{bmatrix}. \tag{2.30}$$

The third step is to set all rows of G_{row} to zero where the corresponding rows in H_{row} have a nonzero entry. For our example, all nonzero rows of G_{row} are set to zero and G_{row} becomes a zero matrix. Then, we remove all columns in G_{row} and H_{row} that only have entries of zeros:

$$G_{row} = 5 \times 0 \text{ empty matrix}, \quad \text{and}$$

$$H_{row} = \begin{bmatrix} \mathbf{1} & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & \mathbf{1} \end{bmatrix}.$$

The final P_{row} is constructed as:

$$P_{row} = \begin{bmatrix} H_{row}^T \\ Q_{row} \\ G_{row}^T \end{bmatrix} = \begin{bmatrix} \mathbf{1} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

where the last row is Q_{row}. If we set $Q_{row} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \end{bmatrix}$ then P_{row} becomes a permutation matrix.

Algorithm 2: The next step is to construct the column-permutation matrix P_{col}. This is done by first constructing the matrices G_{col} and H_{col}, which are the same as G_{row} and H_{row} given in (2.29) and (2.30), respectively. Then set all columns of G_{col} to zero where the corresponding columns in H_{col} have a nonzero entry:

$$G_{col} = \begin{bmatrix} 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Continue by removing all rows of zeros in G_{col} and H_{col}, which give

$$G_{col} = \begin{bmatrix} 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 0 \end{bmatrix}, \quad \text{and}$$

$$H_{col} = \begin{bmatrix} \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} \end{bmatrix}.$$

The resulting P_{col} is

$$P_{col} = \begin{bmatrix} H_{col}^T & Q_{col} & G_{col}^T \end{bmatrix} = \begin{bmatrix} \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 0 \\ 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 0 \end{bmatrix},$$

where Q_{col} is an empty matrix (6×0).

Finally, multiplying $G - \lambda H$ with P_{row} and P_{col} transform the matrix pencil in

KCF to the corresponding system pencil in GBCF:

$$P_{row}\left(\begin{bmatrix}0&1&0&0&0&0\\0&0&0&1&0&0\\0&0&0&0&0&0\\0&0&0&0&\alpha&1\\0&0&0&0&0&\alpha\end{bmatrix}-\lambda\begin{bmatrix}1&0&0&0&0&0\\0&0&1&0&0&0\\0&0&0&0&0&0\\0&0&0&0&1&0\\0&0&0&0&0&1\end{bmatrix}\right)P_{col}$$

$$=\left[\begin{array}{cccc:cc}0&0&0&0&1&0\\0&0&0&0&0&1\\0&0&\alpha&1&0&0\\0&0&0&\alpha&0&0\\\hdashline 0&0&0&0&0&0\end{array}\right]-\lambda\left[\begin{array}{cccc:cc}1&0&0&0&0&0\\0&1&0&0&0&0\\0&0&1&0&0&0\\0&0&0&1&0&0\\\hdashline 0&0&0&0&0&0\end{array}\right]$$

$$\equiv\begin{bmatrix}A_B&B_B\\C_B&D_B\end{bmatrix}-\lambda\begin{bmatrix}I_4&0\\0&0\end{bmatrix}.$$

3 Computing canonical structure information

In this section, we briefly discuss numerically stable methods to compute the canonical structure information of a matrix, matrix pencil or a system pencil. These methods transform the matrix or pencil to a so called staircase form (Section 3.1) from which we can extract the canonical structure information. However, in theoretical analyses we use the canonical forms (e.g., those considered in Section 2 like JCF, KCF and GBCF) because they describe the fine structure (canonical) elements of the given matrix, matrix pencil, or system pencil.

In Section 3.2, we consider the controllable and unobservable subspaces of matrix pencils and system pencils and how to compute them robustly. We also review how these subspaces directly are obtained from the BCF of a matrix pair.

3.1 Staircase-type forms

The computation of a canonical form like JCF, KCF, or GBCF is, in general, not a numerically stable process, because the transformation matrices that reduce, for example, a matrix to Jordan canonical form can be arbitrary ill-conditioned. Therefore it is not appropriate to use such canonical forms in practice. Instead we use so called *staircase-type forms* from which we can retrieve the same canonical structure information as from the canonical forms, by only using real orthogonal (or unitary in the complex case) transformation matrices and backward stable algorithms. An algorithm is *backward stable* [114] when it computes the exact canonical structure of a nearby (slightly perturbed) matrix or matrix pencil. Without going into any algorithmic details, we here present some of the staircase-type forms. More details of the different methods are given in [5, 13, 78, 109].

The staircase method was first introduced for matrices by Kublanovskaya in 1966 [83]. The resulting staircase form is called a *Jordan-Schur form*. The basic idea is to compute the null spaces of $(A - \mu I)^j$ for $j = 1, 2, \ldots$, for each eigenvalue μ of A, using unitary similarity transformations without explicitly computing the matrix powers $(A - \mu I)^j$. In [83], a normalized RQ factorization is used for rank decisions, and methods using the singular value decomposition (SVD) have later been developed [54, 79, 80, 102]. In addition to the Schur form (described in Section 2.1), the Jordan-Schur form gives detailed information of the Jordan structure of the matrix. For example, given a matrix A with one eigenvalue μ of multiplicity n, then $B = A - \mu I_n$ is nilpotent and has the only eigenvalue 0. Suppose the computed Jordan-Schur form for the matrix B is

$$\begin{array}{c}
\begin{array}{ccc} \overbrace{\hphantom{0\;\;0\;\;0}}^{m_1} & \overbrace{\hphantom{x\;\;x}}^{m_2} & \overbrace{\hphantom{x\;\;x}}^{m_3} \end{array}\\
\left[\begin{array}{ccc|cc|cc}
0 & 0 & 0 & x & x & x & x \\
 & 0 & 0 & x & x & x & x \\
 & & 0 & x & x & x & x \\ \hline
 & & & 0 & 0 & x & x \\
 & & & & 0 & x & x \\ \hline
 & & & & & 0 & 0 \\
 & & & & & & 0
\end{array}\right]
\end{array}.$$

Then the dimensions m_1, m_2 and m_3 are the Weyr characteristics of B for the eigenvalue 0, $\mathcal{J}_0 = (3, 2, 2)$. This corresponds to the JCF $J_3(0) \oplus J_3(0) \oplus J_1(0)$, and it follows that the matrix A has the JCF $J_3(\mu) \oplus J_3(\mu) \oplus J_1(\mu)$.

The staircase form for (singular) matrix pencils is called the *generalized Schur-staircase form*, which is the orthogonal counterpart of KCF. Other names used are *Kronecker-Schur form* and *GUPTRI form* (Generalized UPer TRIangular form) [26, 27]. The generalization of the staircase form for matrices to singular matrix pencils was done by Van Dooren [106, 108] using unitary equivalence transformations. This method has later been refined by a number of authors [4, 26, 27, 24, 77, 84, 85].

For example, an $m_{\rm p} \times n_{\rm p}$ singular matrix pencil $G - \lambda H$ can be transformed into the GUPTRI form [26, 27, 78]:

$$U(G - \lambda H)V^H = \begin{bmatrix} G_r - \lambda H_r & * & * \\ 0 & G_{reg} - \lambda H_{reg} & * \\ 0 & 0 & G_l - \lambda H_l \end{bmatrix}, \tag{3.31}$$

where U $(m \times m)$ and V $(n \times n)$ are unitary matrices and $*$ denotes arbitrary conforming submatrices. The rectangular block upper triangular $G_r - \lambda H_r$ and $G_l - \lambda H_l$ give the right and left singular structures of the matrix pencil, respectively. The remaining square upper triangular $G_{reg} - \lambda H_{reg}$ is regular and contains all the finite and infinite eigenvalues of $G - \lambda H$. Furthermore, the regular part $G_{reg} - \lambda H_{reg}$ is in the staircase form:

$$G_{reg} = \begin{bmatrix} G_z & * & * \\ 0 & G_f & * \\ 0 & 0 & G_i \end{bmatrix}, \qquad H_{reg} = \begin{bmatrix} H_z & * & * \\ 0 & H_f & * \\ 0 & 0 & H_i \end{bmatrix},$$

where $G_z - \lambda H_z$ and $G_i - \lambda H_i$ reveal the Jordan structures of the zero and infinite eigenvalues, and $G_f - \lambda H_f$, in generalized Schur form, includes the finite but nonzero eigenvalues.

As we have touched upon in the end of Section 2.2, the eigenvalues μ_i are computed as pairs of values, denoted by (α_i, β_i). If $\alpha_i \neq 0$ and $\beta_i \neq 0$ then μ_i is the finite nonzero eigenvalue $\mu_i = \alpha_i/\beta_i$, if $\alpha_i = 0$ and $\beta_i \neq 0$ then μ_i is a zero eigenvalue, and if $\alpha_i \neq 0$ and $\beta_i = 0$ then μ_i is an infinite eigenvalue. Notably, $\alpha_i = \beta_i = 0$ does not correspond to an eigenvalue, instead it belongs to the singular part of the matrix pencil. In the complex case of the GUPTRI form, the pairs of values (α_i, β_i) are given from the two corresponding diagonal elements of G_{reg} and H_{reg}:

$$G_{reg} = \begin{bmatrix} \ddots & & * \\ & \alpha_i & \\ 0 & & \ddots \end{bmatrix}, \quad \text{and} \quad H_{reg} = \begin{bmatrix} \ddots & & * \\ & \beta_i & \\ 0 & & \ddots \end{bmatrix}.$$

Consequently, the diagonal elements of G_f, G_i, H_z and H_f are nonzero, and those of G_z and H_i are zero.

The use of staircase-type forms or other types of condensed forms has a number of applications in systems and control theory, such as the computation of controllability, observability, minimality of state-space models, Kronecker structures, poles and zeros. To compute the Kronecker structure of a system pencil one of the staircase algorithms for matrix pencils can be used, but there exist efficient algorithms that exploit the special structure of a system pencil (e.g., see [13, 25, 89, 92, 107, 111]). To get a system pencil in a staircase form, the permuted system pencil

$$\widetilde{\mathbf{S}}(\lambda) = \begin{bmatrix} B & A - \lambda E \\ D & C \end{bmatrix}, \tag{3.32}$$

is usually considered. In the following, we take a closer look at staircase-type forms for the controllability pair (A, B), the observability pair (A, C) and the generalized state-space system (E, A, B, C, D).

Instead of computing the BCF of the controllability pair (A, B), the $n \times (n + m)$ system pencil $\mathbf{S}_{\rm C}(\lambda)$ is transformed into the so called *controllability staircase form* (block version of the *controller-Hessenberg form*) [8, 25, 94, 107, 109]. There exits a unitary matrix U $(n \times n)$, such that

$$
\begin{aligned}
&U \left[\begin{array}{c:c} B & A - \lambda I_n \end{array}\right] \begin{bmatrix} I_m & 0 \\ 0 & U^H \end{bmatrix} \\
&= \left[\begin{array}{c:ccccc} X_1 & A_{1,1} & * & \cdots & * & * \\ 0 & X_2 & A_{2,2} & & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & X_{\epsilon_1} & A_{\epsilon_1,\epsilon_1} & * \\ \hdashline 0 & \cdots & \cdots & 0 & 0 & A_{reg} \end{array}\right] - \lambda \left[\begin{array}{c:c} 0 & I_n \end{array}\right],
\end{aligned}
\tag{3.33}
$$

where the matrices $A_{i,i}$, $i = 1, \ldots, \epsilon_1$, are of size $r_i \times r_i$. The matrices X_i, $i = 1, \ldots, \epsilon_1$, are of size $r_i \times r_{i-1}$ with full row rank r_i, where $r_0 = m$. The sizes r_i form the integer partition $\mathcal{R}(A, B) = (r_0, r_1, \ldots, r_{\epsilon_1})$, where the conjugate of $(r_1, \ldots, r_{\epsilon_1})$ defines the controllability indices of (A, B). The matrix A_{reg} is regular and consists of the finite elementary divisors, i.e., the uncontrollable eigenvalues (modes) of (A, B).

The dual form is the *observability staircase form* (block version of the *observer-Hessenberg form*) for the $(n + p) \times n$ observability system pencil $\mathbf{S}_{\mathrm{O}}(\lambda)$ [8, 25, 94, 107, 109]. There exits a unitary matrix U $(n \times n)$, such that

$$
\begin{aligned}
&\begin{bmatrix} U & 0 \\ 0 & I_p \end{bmatrix} \left[\begin{array}{c} A - \lambda I_n \\ \hdashline C \end{array}\right] U^H \\
&= \left[\begin{array}{c:cccc} A_{reg} & * & \cdots & \cdots & * \\ 0 & A_{\eta_1,\eta_1} & \cdots & \cdots & * \\ 0 & Y_{\eta_1} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & A_{2,2} & * \\ \vdots & & \ddots & Y_2 & A_{1,1} \\ \hdashline 0 & \cdots & \cdots & 0 & Y_1 \end{array}\right] - \lambda \left[\begin{array}{c} I_n \\ \hdashline 0 \end{array}\right],
\end{aligned}
\tag{3.34}
$$

where the matrices $A_{i,i}$, $i = 1, \ldots, \eta_1$, are of size $l_i \times l_i$. The matrices Y_i, $i = 1, \ldots, \eta_1$, are of size $l_i \times l_{i-1}$ with full column rank l_i, where $l_0 = p$. The sizes l_i form the integer partition $\mathcal{L}(A, C) = (l_0, l_1, \ldots, l_{\eta_1})$, where the conjugate of $(l_1, \ldots, l_{\eta_1})$ defines the observability indices of (A, C). The matrix A_{reg} is regular and consists of the unobservable eigenvalues (modes) of (A, C).

There also exist staircase counterparts for the GBCF of a system pencil. The system pencil $\widetilde{\mathbf{S}}(\lambda)$ in (3.32) associated with a generalized state-space system (E, A, B, C, D) can be transformed into the *staircase Kronecker-like form* [110, 111] (or a similar staircase form, e.g., see [13]) using orthogonal matrices Q and Z, such that

$$
Q\widetilde{\mathbf{S}}(\lambda)Z = \left[\begin{array}{c:ccccc} B_r & A_r - \lambda E_r & * & * & * & * \\ 0 & 0 & A_\infty - \lambda E_\infty & * & * & * \\ 0 & 0 & 0 & D_i & * & * \\ 0 & 0 & 0 & 0 & A_f - \lambda E_f & * \\ 0 & 0 & 0 & 0 & 0 & A_l - \lambda E_l \\ \hdashline 0 & 0 & 0 & 0 & 0 & C_l \end{array}\right]. \tag{3.35}
$$

The generalized matrix pair (E_r, A_r, B_r) is controllable, and the system pencil $\begin{bmatrix} B_r & A_r - \lambda E_r \end{bmatrix}$ is in controllability staircase form and gives the right (column) minimal indices. The matrix E_r is invertible and upper-triangular and $A_r - \lambda E_r$ has full row-rank. Similarly, the generalized matrix pair (E_l, A_l, C_l) is observable, and the system pencil $\left[\begin{smallmatrix} A_l-\lambda E_l \\ C_l \end{smallmatrix}\right]$ is in observable staircase form and gives the left (row) minimal indices. The matrix E_l is also invertible and upper-triangular and $A_l - \lambda E_l$ has full column-rank.

Together the regular matrix pencil $A_\infty - \lambda E_\infty$ and the matrix D_i give the infinite elementary divisors, where A_∞ and D_i are invertible and upper-triangular, and E_∞ is nilpotent and upper-triangular. The matrix pencil $A_f - \lambda E_f$ gives the finite elementary divisors, where E_f is invertible and upper-triangular.

Example 6

We consider the state-space system (2.20) in Example 4 for $\gamma = 0$:

$$\begin{bmatrix} B & A-\lambda I_n \\ D & C \end{bmatrix} = \left[\begin{array}{ccc:cc} 3 & 10 & 1 & -\lambda & 0 \\ 0.6 & 2 & 0.2 & -3 & -\lambda \\ \hdashline 0 & 0 & 0 & 0.6 & 0 \end{array}\right].$$

The controllability staircase form is computed as (rounded to four decimals)

$$\begin{bmatrix} -0.9806 & -0.1961 \\ -0.1961 & 0.9806 \end{bmatrix} \left[\begin{array}{ccc:cc} 3 & 10 & 1 & 0 & 0 \\ 0.6 & 2 & 0.2 & -3 & 0 \end{array}\right] \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -0.9806 & -0.1961 \\ 0 & 0 & 0 & -0.1961 & 0.9806 \end{bmatrix}$$

$$= \left[\begin{array}{ccc:cc} -3.0594 & -10.1980 & -1.0198 & -0.5769 & -0.1154 \\ 0 & 0 & 0 & 2.8846 & 0.5769 \end{array}\right]$$

$$\equiv \left[\begin{array}{c:cc} X_1 & * & * \\ 0 & X_2 & * \end{array}\right].$$

We can now derive the Kronecker structure in the following way: $r_1 = \text{rank}(X_1) = 1$, $r_2 = \text{rank}(X_2) = 1$, and the system has 3 inputs; therefore $r_0 = 3$. There exists no regular part (A_{reg} is absent) so the controllability pair has $\mathcal{R}(A, B) = (3, 1, 1)$, i.e., the KCF $L_2 \oplus 2L_0$.

Similarly, the observability staircase form is computed as

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{cc} 0 & 0 \\ -3 & 0 \\ \hdashline 0.6 & 0 \end{array}\right] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \left[\begin{array}{cc} 0 & -3 \\ 0 & 0 \\ \hdashline 0 & 0.6 \end{array}\right] \equiv \left[\begin{array}{cc} A_{reg} & * \\ 0 & * \\ \hdashline 0 & Y_1 \end{array}\right],$$

where $l_1 = \text{rank}(Y_1) = 1$ and $l_0 = 1$. As we can see there exists a regular part of size 1×1 corresponding to the unobservable mode (here the eigenvalue 0). Consequently, $\mathcal{L}(A, C) = (1, 1)$ and $\mathcal{J}_0(A, C) = (1)$, which correspond to the KCF $L_1^T \oplus J_1(0)$.

3.2 Computing controllable and unobservable subspaces

The *controllable subspace* $\mathbf{C}_\mathcal{S}(A, B)$ and *unobservable subspace* $\overline{\mathbf{O}}_\mathcal{S}(A, C)$ of a state-space system (A, B, C, D) is defined, respectively, as (e.g., see [25, 107, 116])

$$\mathbf{C}_\mathcal{S}(A, B) = \inf\{\mathcal{S} \mid A\mathcal{S} \subset \mathcal{S};\ \text{ran}(B) \subset \mathcal{S}\} = \text{ran}(\mathbf{C}(A, B)),$$

$$\overline{\mathbf{O}}_\mathcal{S}(A, C) = \sup\{\mathcal{S} \mid A\mathcal{S} \subset \mathcal{S};\ \mathcal{S} \subset \text{null}(C)\} = \text{null}(\mathbf{O}(A, C)).$$

If $\dim(\mathbf{C}_{\mathcal{S}}(A,B)) = c < n$, then the system is uncontrollable and there exists an uncontrollable subspace. Analogously, if $\dim(\overline{\mathbf{O}}_{\mathcal{S}}(A,C)) = \bar{o} > 0$, then the system is unobservable and there exists an unobservable subspace.

With an $n \times n$ unitary transformation matrix U the controllability system pencil $\mathbf{S}_{\mathrm{C}}(\lambda)$ can be reduced to the controllability staircase form (3.33):

$$U \begin{bmatrix} B & A - \lambda I_n \end{bmatrix} \begin{bmatrix} I_m & 0 \\ 0 & U^H \end{bmatrix} = \left[\begin{array}{c|cc} B_c & A_c - \lambda I_c & * \\ 0 & 0 & A_{\bar{c}} - \lambda I_{\bar{c}} \end{array} \right],$$

where (A_c, B_c) is controllable, $A_{\bar{c}} - \lambda I_{\bar{c}}$ is regular and contains the uncontrollable modes, and the first c rows of U span $\mathbf{C}_{\mathcal{S}}(A,B)$ [107]. Dually, the observable subspace can be derived from the observability staircase form (3.34):

$$\begin{bmatrix} U & 0 \\ 0 & I_p \end{bmatrix} \begin{bmatrix} A - \lambda I_n \\ C \end{bmatrix} U^H = \left[\begin{array}{cc} A_{\bar{o}} - \lambda I_{\bar{o}} & * \\ 0 & A_o - \lambda I_o \\ \hline 0 & C_o \end{array} \right],$$

where (A_o, C_o) is observable, $A_{\bar{o}} - \lambda I_{\bar{o}}$ is regular and contains the unobservable modes, and the first $\bar{o}$ rows of U span $\overline{\mathbf{O}}_{\mathcal{S}}(A,C)$ [107].

It follows, that if a controllability pair (A,B) already is in BCF then

$$\begin{bmatrix} I_c & 0 \\ 0 & I_{\bar{c}} \end{bmatrix} \begin{bmatrix} B_B & A_B - \lambda I_n \end{bmatrix} \begin{bmatrix} I_m & 0 & 0 \\ 0 & I_c & 0 \\ 0 & 0 & I_{\bar{c}} \end{bmatrix} = \begin{bmatrix} B_\epsilon & A_\epsilon & 0 \\ 0 & 0 & A_\mu \end{bmatrix},$$

where A_ϵ has c rows/columns and A_μ has $\bar{c}$ rows/columns. Consequently,

$$\mathbf{C}_{\mathcal{S}}(A,B) = \operatorname{span}\left\{ \begin{bmatrix} I_c \\ 0 \end{bmatrix} \right\}.$$

Similarly, if an observability pair (A,C) is in BCF then

$$\begin{bmatrix} I_{\bar{o}} & 0 & 0 \\ 0 & I_o & 0 \\ 0 & 0 & I_p \end{bmatrix} \begin{bmatrix} A_B - \lambda I_n \\ C_B \end{bmatrix} \begin{bmatrix} I_{\bar{o}} & 0 \\ 0 & I_o \end{bmatrix} = \begin{bmatrix} A_\mu & 0 \\ 0 & A_\eta \\ 0 & C_\eta \end{bmatrix},$$

where A_μ has $\bar{o}$ rows/columns and A_η has o rows/columns, then

$$\overline{\mathbf{O}}_{\mathcal{S}}(A,C) = \operatorname{span}\left\{ \begin{bmatrix} I_{\bar{o}} \\ 0 \end{bmatrix} \right\}.$$

The subspaces can also be derived from a generalized Schur-staircase form where the structure of the system pencil is not preserved, in contrary to the controllability and observability staircase forms. For example, from the GUPTRI form of the corresponding general matrix pencil $G - \lambda H$ of a system, the subspaces can be derived as follows.

From the $m_{\mathrm{p}} \times n_{\mathrm{p}}$ matrix pencil $G - \lambda H$ in the GUPTRI form

$$U(G - \lambda H)V^H = \begin{bmatrix} G_r - \lambda H_r & * & * \\ 0 & G_{reg} - \lambda H_{reg} & * \\ 0 & 0 & G_l - \lambda H_l \end{bmatrix},$$

different pairs of *reducing subspaces* can be computed [25, 108]. Let $G_r - \lambda H_r$ be $m_r \times n_r$, $G_{reg} - \lambda H_{reg}$ be $m_{reg} \times n_{reg}$, and $G_l - \lambda H_l$ be $m_l \times n_l$. Let also U and V be partitioned as

$U = \begin{bmatrix} U_r & U_{reg} & U_l \end{bmatrix}$ and $V = \begin{bmatrix} V_r & V_{reg} & V_l \end{bmatrix}$, respectively, where the dimensions of each submatrix are given from the sizes of the blocks in the GUPTRI form of $G - \lambda H$. Then the *left* and *right reducing subspaces*, $\mathbf{U}$ and $\mathbf{V}$, form a pair of reducing subspaces spanned by the leading columns of U and V, respectively. The subspace is called *minimal* if it is spanned by the *minimal reducing subspace pair* $(\mathrm{span}\{U_r\}, \mathrm{span}\{V_r\})$, and *maximal* if it is spanned by the *maximal reducing subspace pair* $(\mathrm{span}\{U_r, U_{reg}\}, \mathrm{span}\{V_r, V_{reg}\})$.

If the $n \times (n+m)$ controllability system pencil is in the GUPTRI form

$$U \begin{bmatrix} B & A - \lambda I_n \end{bmatrix} V^H = \begin{bmatrix} A_r - \lambda B_r & * \\ 0 & A_{reg} - \lambda B_{reg} \end{bmatrix},$$

it follows that the controllable subspace is equal to the minimal left reducing subspace $\mathbf{U} \equiv \mathrm{span}\{U_r\}$, or equivalently, the bottom n rows of the minimal right reducing subspace $\mathbf{V} \equiv \mathrm{span}\{V_r\}$. For the generalized matrix pair (E, A, B), where E is nonsingular, the controllable subspace is equal to $E^{-1}\mathbf{U}$ [25].

Analogously, the unobservable subspace is equal to the maximal right reducing subspace of $\mathbf{S}_{\mathrm{O}}(\lambda)$, or equivalently, the first n rows of maximal left reducing subspace [25]. This follows from the duality

$$\begin{aligned}
&\text{unobservable subspace of } (A, C) \\
&\quad = (\text{controllable subspace of } (A^T, C^T))^{\perp} \\
&\quad = (\text{minimal left reducing subspace of } \begin{bmatrix} C^T & A^T - \lambda I_n \end{bmatrix})^{\perp} \\
&\quad = \text{maximal right reducing subspace of } \left[\begin{smallmatrix} A - \lambda I_n \\ C \end{smallmatrix}\right].
\end{aligned}$$

4 Matrix and pencil spaces

A matrix can be seen as a point in a *matrix space*, and the union of all similar matrices as a manifold in this space. We say that the matrix "lives" in the space spanned by the manifold, and the dimension of the manifold is given from the number of parameters of the matrix, where each fixed parameter gives one less degree of freedom. The dimension of the complementary space to the manifold is called the codimension, and as we will see has a vital role in the theory of stratification.

In this section, we consider the matrix space and the corresponding spaces for matrix pencils and system pencils. Moreover, it is shown how the dimensions and codimensions of these spaces are computed, and we also present a convenient way to get the codimension from the canonical structure information of a matrix, matrix pencil or system pencil.

4.1 The matrix space

A matrix A of size $n \times n$ has n^2 elements and therefore belongs to an n^2-dimensional (matrix) space, one dimension for each parameter. As mentioned above, a matrix A can be seen as a point in the n^2-dimensional space and consequently the union of all $n \times n$ matrices constitute the entire matrix space [30].

The *orbit* of a matrix, $\mathcal{O}(A)$, is the manifold of all similar matrices:

$$\mathcal{O}(A) = \{PAP^{-1} : \det(P) \neq 0\}. \tag{4.36}$$

This means that all matrices in the same orbit have the same canonical form, both the eigenvalues and the sizes of the Jordan blocks are fixed, and that $\mathcal{O}$(A) is a manifold in the n^2-dimensional space. A *bundle* defines the union of all orbits with the same canonical form but with the eigenvalues unspecified, $\bigcup_{\mu_i} \mathcal{O}(A)$ [2]. We denote the bundle of A by $\mathcal{B}(A)$.

The dimension of the space $\mathcal{O}(A)$ is equal to the dimension of the *tangent space* to $\mathcal{O}(A)$ at A, denoted by $\tan(A)$, and is defined in terms of A by the matrices of the form $T_A = XA - AX$, where X is an $n \times n$ matrix. Using the technique in [30] the tangent vectors T_A can be expressed in terms of the vec-operator and Kronecker products as:

$$\begin{aligned} \operatorname{vec}(T_A) &= \left(A^T \otimes I_n\right) \operatorname{vec}(X) - (I_n \otimes A) \operatorname{vec}(X) \\ &= \left(A^T \otimes I_n - I_n \otimes A\right) \operatorname{vec}(X). \end{aligned}$$

The orthogonal complement of the tangent space is the *normal space*, $\operatorname{nor}(A)$, which is the union of all $n \times n$ matrices Z that satisfy

$$A^H Z = Z A^H.$$

Figure 1 illustrates an orbit of a tentative matrix A together with the tangent and normal spaces to the orbit at A.

The dimension of the space complementary to the orbit is called the *codimension* of $\mathcal{O}(A)$, denoted by $\operatorname{cod}(\mathcal{O}(A))$ [23, 30, 113]. Consequently the codimension is equal to the dimension of the normal space and

$$\operatorname{cod}(\mathcal{O}(A)) = n^2 - \dim(\tan(A)).$$

The codimension of the $\mathcal{B}(A)$ is denoted $\operatorname{cod}(\mathcal{B}(A))$. When it is clear from context whether it is the codimension of an orbit or a bundle, we only write $\operatorname{cod}(A)$.

An explicit expression for the codimension for matrices was derived by Arnold [2] using *miniversal deformations*, including a parameterization of the normal space, with one parameter for each dimension of the normal space. It follows that the codimension can be obtained

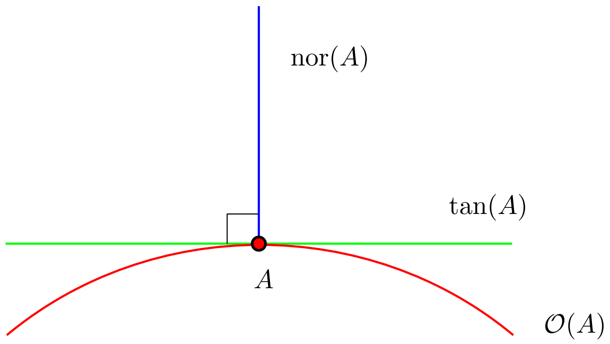


FIGURE 1: Illustration of the orbit, tangent space, and normal space for a tentative matrix A (marked with the dot).

as the number of linearly independent matrices X that solve $f(X) = XA - AX = 0$. For an introduction to and how the (mini)versal deformation is derived we refer to [3] and [30].

A more convenient way to determine the codimension of the orbit of A is based on the Jordan structure of the matrix (e.g., see [23]):

$$\operatorname{cod}(A) = c_{\text{Jor}}, \tag{4.37}$$

where

$$c_{\text{Jor}} = \sum_{i=1}^{q} \sum_{j=1}^{g_i} (2j-1)h_j^{(i)} = \sum_{i=1}^{q} (h_1^{(i)} + 3h_2^{(i)} + 5h_3^{(i)} + \cdots), \tag{4.38}$$

and $(h_1^{(i)}, \ldots, h_{g_i}^{(i)})$ are the Segre characteristics for the finite eigenvalue μ_i, as defined in Section 2.4, and q is the number of distinct eigenvalues.

Simple eigenvalues make no contribution to the codimension in the bundle case. Therefore, knowing the codimension of an orbit the codimension of the corresponding bundle is one less for each distinct eigenvalue:

$$\operatorname{cod}(\mathcal{B}(A)) = \operatorname{cod}(\mathcal{O}(A)) - (\text{number of distinct eigenvalues}).$$

For example, if we are interested in an $n \times n$ matrix A with k unspecified eigenvalues and the rest with known specified values, the codimension of $\mathcal{B}(A)$ is $\operatorname{cod}(\mathcal{O}(A)) - k$.

EXAMPLE 7

Given a matrix A with JCF $2J_2(\mu_1) \oplus J_1(\mu_1) \oplus 3J_5(\mu_2) \oplus J_2(\mu_3)$ with the corresponding Segre characteristics:

$$\begin{aligned} h_{\mu_1} &= (2, 2, 1), \\ h_{\mu_2} &= (5, 5, 5), \text{ and} \\ h_{\mu_3} &= (2). \end{aligned}$$

The codimension of the orbit of A is

$$\operatorname{cod}(A) = (2 + 3 * 2 + 5 * 1) + (5 + 3 * 5 + 5 * 5) + 2 = 60,$$

and the codimension of the bundle of A is $60 - 3 = 57$ (since A has three distinct eigenvalues).

4.2 The matrix pencil space

In the case of $m_\mathrm{p} \times n_\mathrm{p}$ matrix pencils $G - \lambda H$, we now have a $2m_\mathrm{p}n_\mathrm{p}$-dimensional space (two matrices, each with $m_\mathrm{p}n_\mathrm{p}$ elements), where the orbit is the manifold of strictly equivalent matrix pencils:

$$\mathcal{O}(G - \lambda H) = \{U(G - \lambda H)V^{-1} : \det(U) \cdot \det(V) \neq 0\}. \tag{4.39}$$

As for matrices, the bundle of $G - \lambda H$, $\mathcal{B}(G - \lambda H)$, is the set of matrix pencils with the same Kronecker canonical structure, i.e., equal left and right singular blocks and Jordan blocks of equal size but with unspecified eigenvalues.

The dimension of $\mathcal{O}(G - \lambda H)$ is equal to the dimension of the tangent space to $\mathcal{O}(G - \lambda H)$, which can be expressed by the pencils on the form

$$T_G - \lambda T_H = X(G - \lambda H) - (G - \lambda H)Y,$$

where X is an $m_\mathrm{p} \times m_\mathrm{p}$ matrix and Y is an $n_\mathrm{p} \times n_\mathrm{p}$ matrix. Edelman, Elmroth and Kågström [30] showed that by using Kronecker products the $2m_\mathrm{p}n_\mathrm{p}$ tangent vectors $T_G - \lambda T_H$ can be represented as

$$\begin{bmatrix} \mathrm{vec}(T_G) \\ \mathrm{vec}(T_H) \end{bmatrix} = \begin{bmatrix} G^T \otimes I_{m_\mathrm{p}} \\ H^T \otimes I_{m_\mathrm{p}} \end{bmatrix} \mathrm{vec}(X) - \begin{bmatrix} I_{n_\mathrm{p}} \otimes G \\ I_{n_\mathrm{p}} \otimes H \end{bmatrix} \mathrm{vec}(Y),$$

and the tangent space is the range of the $2m_\mathrm{p}n_\mathrm{p} \times ({m_\mathrm{p}}^2 + {n_\mathrm{p}}^2)$ matrix

$$T \equiv \begin{bmatrix} G^T \otimes I_{m_\mathrm{p}} & -I_{n_\mathrm{p}} \otimes G \\ H^T \otimes I_{m_\mathrm{p}} & -I_{n_\mathrm{p}} \otimes H \end{bmatrix}. \tag{4.40}$$

Then the normal space is

$$\mathrm{nor}(G - \lambda H) = \mathrm{null}(T^H) = \{Z_G - \lambda Z_H\},$$

where $Z_G G^H + Z_H H^H = 0$ and $G^H Z_G + H^H Z_H = 0$ [30]. The dimensions of the two complementary spaces can now be expressed in terms of the matrix T as

$$\dim(\tan(G - \lambda H)) = {m_\mathrm{p}}^2 + {n_\mathrm{p}}^2 - \dim(\mathrm{null}(T)),$$

and

$$\dim(\mathrm{nor}(G - \lambda H)) = \dim(\mathrm{null}(T^H)) = \dim(\mathrm{null}(T)) - (m_\mathrm{p} - n_\mathrm{p})^2.$$

As before, the codimension of the orbit is equal to the dimension of the normal space, which together with the tangent space makes up the complete $2m_\mathrm{p}n_\mathrm{p}$-dimensional space for the matrix pencil.

We recall from Section 2.4 the invariants associated with the KCF of a matrix pencil. These are the column minimal indices $(\epsilon_1, \ldots, \epsilon_{r_0})$, the row minimal indices $(\eta_1, \ldots, \eta_{l_0})$, the Segre characteristics $(h_1^{(i)}, \ldots, h_{g_i}^{(i)})$ for the finite eigenvalue μ_i for $i = 1, \ldots, q$, and the Segre characteristics $(s_1, \ldots, s_{g_\infty})$ for the infinite eigenvalue.

Knowing the KCF, Demmel and Edelman [23] derived explicit expressions for the codimension of a matrix pencil. They showed that it is a sum of separate codimensions:

$$\mathrm{cod}(G - \lambda H) = c_{\mathrm{Right}} + c_{\mathrm{Left}} + c_{\mathrm{Sing}} + c_{\mathrm{Jor}} + c_{\mathrm{Jor,Sing}}, \tag{4.41}$$

where

$$c_{\text{Right}} = \sum_{\epsilon_i > \epsilon_j} (\epsilon_i - \epsilon_j - 1), \qquad c_{\text{Left}} = \sum_{\eta_i > \eta_j} (\eta_i - \eta_j - 1),$$

$$c_{\text{Sing}} = \sum_{\epsilon_i, \eta_j} (\epsilon_i + \eta_j + 2), \qquad c_{\text{Jor}} = \sum_{i=1}^{q} \sum_{j=1}^{g_i} (2j-1) h_j^{(i)} + \sum_{j=1}^{g_\infty} (2j-1) s_j,$$

and

$$c_{\text{Jor,Sing}} = (r_0 + l_0) \left(\sum_{i=1}^{q} \sum_{j=1}^{g_i} h_j^{(i)} + \sum_{j=1}^{g_\infty} s_j \right).$$

The first two terms, c_{Right} and c_{Left}, come from the interaction between L blocks and L^T blocks, respectively. The term c_{Sing} comes from the interaction between the right and left singular blocks and is the summation over all pairs of L_{ϵ_i} and $L_{\eta_j}^T$ blocks. The term c_{Jor} comes from the Jordan blocks and corresponds to (4.38) for matrices, but also includes the infinite eigenvalues appearing in general matrix pencils. The last term $c_{\text{Jor,Sing}}$ is the product of the number of singular blocks and the total size of the regular part. As for matrices, the codimension of the corresponding bundle is given as:

$$\operatorname{cod}(\mathcal{B}(G - \lambda H)) = \operatorname{cod}(\mathcal{O}(G - \lambda H)) - (\text{number of distinct eigenvalues}).$$

Example 8

Given a matrix pencil $G - \lambda H$ with KCF $L_3 \oplus L_1 \oplus L_0 \oplus L_3^T \oplus L_0^T \oplus J_2(\alpha) \oplus J_1(\alpha) \oplus N_3$ with the corresponding integer partitions:

$$\begin{aligned} \epsilon &= (3, 1, 0), & \eta &= (3, 0), \\ h_\alpha &= (2, 1), \text{ and } & s &= (3). \end{aligned}$$

The codimension of the orbit of $G - \lambda H$ is the sum of the terms

$$\begin{aligned} c_{\text{Right}} &= (3 - 1 - 1) + (3 - 0 - 1) + (1 - 0 - 1) = 3, \\ c_{\text{Left}} &= 3 - 0 - 1 = 2, \\ c_{\text{Sing}} &= (3 + 3 + 2) + (3 + 0 + 2) + (1 + 3 + 2) + (1 + 0 + 2) \\ &\quad + (0 + 3 + 2) + (0 + 0 + 2) = 29, \\ c_{\text{Jor}} &= (2 + 3 * 1) + 3 = 8, \text{ and} \\ c_{\text{Jor,Sing}} &= (3 + 2)(2 + 1 + 3) = 30, \end{aligned}$$

which give $\operatorname{cod}(G - \lambda H) = 3+2+29+8+30 = 72$. It follows that the codimension of the bundle of $G - \lambda H$ is $72 - 2 = 70$, since we have two eigenvalues (one finite and one infinite).

Another approach to compute the codimension is from the singular value decomposition (SVD) of the matrix T in (4.40) [30]. It follows that

$$\operatorname{cod}(G - \lambda H) = \text{number of zero singular values of } T,$$

and that the left singular vectors corresponding to the zero singular value form an orthonormal basis for $\mathrm{nor}(G - \lambda H)$. The corresponding result for square matrices is

$$\mathrm{cod}(A) = \text{number of zero singular values of } A^T \otimes I_n - I_n \otimes A.$$

This is a robust but rather costly method for computing the codimension, e.g., to compute the SVD of T is an $O(m^3 n^3)$ operation. However, the main advantage of the SVD-based method is that the codimension can be computed without any knowledge of the canonical structure of the orbit.

Miniversal deformations for matrix pencils were derived by Edelman, Elmroth and Kågström [30] and partially by Berg and Kwatny [6]. Further studies on versal deformations of matrix pencils have, for example, been done in [46, 48], and [47] where the simplest miniversal deformation of matrices and matrix pencils is derived. Versal deformations of different kinds of system pencils (considered in the next section) have, for example, been studied in [7, 39, 50, 51, 103] and of invariant subspaces in [41, 99].

4.3 The system pencil space

Next, we consider pairs, triples and quadruples of matrices. An $(n + p) \times (n + m)$ matrix quadruple (A, B, C, D) belongs to an $((n+p)(n+m))$-dimensional space and a matrix triple (A, B, C) to an $(n^2 + np + nm)$-dimensional space. Similarly, the controllability pair (A, B) belongs to an $(n^2 + nm)$-dimensional space and the observability pair (A, C) belongs to an $(n^2 + np)$-dimensional space. Throughout this paper we are only considering orbits and bundles under feedback equivalence of these systems. For matrix quadruples (and matrix triples when $D \equiv 0$) such an orbit is defined as

$$\begin{aligned} &\mathcal{O}(A, B, C, D) \\ &= \left\{ \begin{bmatrix} P & S \\ 0 & T \end{bmatrix} \begin{bmatrix} A - \lambda I & B \\ C & D \end{bmatrix} \begin{bmatrix} P^{-1} & 0 \\ R & Q^{-1} \end{bmatrix} : \det(P) \cdot \det(T) \cdot \det(Q) \neq 0 \right\}. \end{aligned}$$

Similarly, the orbit for the controllability pairs is defined as

$$\mathcal{O}(A, B) = \left\{ P \begin{bmatrix} A - \lambda I & B \end{bmatrix} \begin{bmatrix} P^{-1} & 0 \\ R & Q^{-1} \end{bmatrix} : \det(P) \cdot \det(Q) \neq 0 \right\},$$

and for the observability pairs we have

$$\mathcal{O}(A, C) = \left\{ \begin{bmatrix} P & S \\ 0 & T \end{bmatrix} \begin{bmatrix} A - \lambda I \\ C \end{bmatrix} P^{-1} : \det(P) \cdot \det(T) \neq 0 \right\}.$$

The tangent space to $\mathcal{O}(A, B, C, D)$ at (A, B, C, D) is given by the system matrix of the form

$$\begin{bmatrix} T_A & T_B \\ T_C & T_D \end{bmatrix} = \begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} + \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} -X & 0 \\ V & W \end{bmatrix},$$

where X, Y, Z, V and W are matrices of conforming sizes [40]. Similar to the general matrix pencil case, we can express the tangent space of a matrix quadruple (A, B, C, D) as the range of the $(n^2 + nm + np + mp) \times (n^2 + np + p^2 + nm + m^2)$ matrix

$$T_{(A,B,C,D)} = \begin{bmatrix} A^T \otimes I_n - I_n \otimes A & C^T \otimes I_n & 0 & I_n \otimes B & 0 \\ B^T \otimes I_n & D^T \otimes I_n & 0 & 0 & I_m \otimes B \\ -I_n \otimes C & 0 & C^T \otimes I_p & I_n \otimes D & 0 \\ 0 & 0 & D^T \otimes I_p & 0 & I_m \otimes D \end{bmatrix},$$

where

$$\begin{bmatrix} \mathrm{vec}(T_A) \\ \mathrm{vec}(T_B) \\ \mathrm{vec}(T_C) \\ \mathrm{vec}(T_D) \end{bmatrix} = T_{(A,B,C,D)} \begin{bmatrix} \mathrm{vec}(X) \\ \mathrm{vec}(Y) \\ \mathrm{vec}(Z) \\ \mathrm{vec}(V) \\ \mathrm{vec}(W) \end{bmatrix}.$$

The matrix $T_{(A,B,C,D)}$ is derived using the technique for general matrix pencils [30]. The corresponding matrix representations of the tangent space of a triple (A, B, C), a controllability pair (A, B) and an observability pair (A, C) are ([15, 45]):

$$T_{(A,B,C)} = \begin{bmatrix} A^T \otimes I_n - I_n \otimes A & C^T \otimes I_n & 0 & I_n \otimes B & 0 \\ B^T \otimes I_n & 0 & 0 & 0 & I_m \otimes B \\ -I_n \otimes C & 0 & C^T \otimes I_p & 0 & 0 \end{bmatrix},$$

$$T_{(A,B)} = \begin{bmatrix} A^T \otimes I_n - I_n \otimes A & I_n \otimes B & 0 \\ B^T \otimes I_n & 0 & I_m \otimes B \end{bmatrix}, \quad \text{and}$$

$$T_{(A,C)} = \begin{bmatrix} A^T \otimes I_n - I_n \otimes A & C^T \otimes I_n & 0 \\ -I_n \otimes C & 0 & C^T \otimes I_p \end{bmatrix}.$$

As before, the dimension of the orbit is equal to the dimension of the tangent space to the orbit, and the codimension is equal to the dimension of the associated normal space. Expressed in terms of the T-matrix notation, we have for the different systems (see [51] for (A, B, C) and [39] for (A, B)):

$$\begin{aligned}
\dim(\tan(A,B,C,D)) &= n^2 + np + p^2 + nm + m^2 - \dim(\mathrm{null}(T_{(A,B,C,D)})), \\
\dim(\mathrm{nor}(A,B,C,D)) &= \dim(\mathrm{null}(T_{(A,B,C,D)})) - m^2 - p^2 + pm, \\
\dim(\tan(A,B,C)) &= n^2 + np + p^2 + nm + m^2 - \dim(\mathrm{null}(T_{(A,B,C)})), \\
\dim(\mathrm{nor}(A,B,C)) &= \dim(\mathrm{null}(T_{(A,B,C)})) - m^2 - p^2, \\
\dim(\tan(A,B)) &= n^2 + nm + m^2 - \dim(\mathrm{null}(T_{(A,B)})), \\
\dim(\mathrm{nor}(A,B)) &= \dim(\mathrm{null}(T_{(A,B)})) - m^2 - p^2 - np, \\
\dim(\tan(A,C)) &= n^2 + np + p^2 - \dim(\mathrm{null}(T_{(A,C)})), \text{ and} \\
\dim(\mathrm{nor}(A,C)) &= \dim(\mathrm{null}(T_{(A,C)})) - m^2 - p^2 - nm.
\end{aligned}$$

For the generalized case of the matrix quadruple where we also have restricted system equivalence [17, 101], the tangent space to $\mathcal{O}(E, A, B, C, D)$ is

$$\begin{bmatrix} T_E & T_A & T_B \\ 0 & T_C & T_D \end{bmatrix} = \begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix} \begin{bmatrix} E & A & B \\ 0 & C & D \end{bmatrix} + \begin{bmatrix} E & A & B \\ 0 & C & D \end{bmatrix} \begin{bmatrix} -S & 0 & 0 \\ 0 & -S & 0 \\ 0 & V & W \end{bmatrix}.$$

As for state-space systems we can get the tangent space from the range of

$$T_{(E,A,B,C,D)} = \begin{bmatrix} E^T \otimes I_n & -I_n \otimes E & 0 & 0 & 0 & 0 \\ A^T \otimes I_n & -I_n \otimes A & C^T \otimes I_n & 0 & I_n \otimes B & 0 \\ B^T \otimes I_n & 0 & D^T \otimes I_n & 0 & 0 & I_m \otimes B \\ 0 & -I_n \otimes C & 0 & C^T \otimes I_p & I_n \otimes D & 0 \\ 0 & 0 & 0 & D^T \otimes I_p & 0 & I_m \otimes D \end{bmatrix},$$

where

$$\begin{bmatrix} \mathrm{vec}(T_E) \\ \mathrm{vec}(T_A) \\ \mathrm{vec}(T_B) \\ \mathrm{vec}(T_C) \\ \mathrm{vec}(T_D) \end{bmatrix} = T_{(E,A,B,C,D)} \begin{bmatrix} \mathrm{vec}(X) \\ \mathrm{vec}(S) \\ \mathrm{vec}(Y) \\ \mathrm{vec}(Z) \\ \mathrm{vec}(V) \\ \mathrm{vec}(W) \end{bmatrix}.$$

This is an extension of the generalized matrix triple (E, A, B, C) considered in [15]. For (E, A, B, C) the group of proportional plus derivative feedback transformations [86, 118] acting on the E matrix is also of interest. The tangent space is now given as (see [15])

$$\begin{bmatrix} T_E & T_A & T_B \\ 0 & T_C & 0 \end{bmatrix} = \begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix} \begin{bmatrix} E & A & B \\ 0 & C & 0 \end{bmatrix} + \begin{bmatrix} E & A & B \\ 0 & C & 0 \end{bmatrix} \begin{bmatrix} -S & 0 & 0 \\ 0 & -S & 0 \\ U & V & W \end{bmatrix},$$

with the corresponding T matrix in Kronecker product representation:

$$T_{(E,A,B,C)}$$
$$= \begin{bmatrix} E^T \otimes I_n & -I_n \otimes E & 0 & 0 & I_n \otimes B & 0 & 0 \\ A^T \otimes I_n & -I_n \otimes A & C^T \otimes I_n & 0 & 0 & I_n \otimes B & 0 \\ B^T \otimes I_n & 0 & 0 & 0 & 0 & 0 & I_m \otimes B \\ 0 & -I_n \otimes C & 0 & C^T \otimes I_p & 0 & 0 & 0 \end{bmatrix},$$

where

$$\begin{bmatrix} \mathrm{vec}(T_E) \\ \mathrm{vec}(T_A) \\ \mathrm{vec}(T_B) \\ \mathrm{vec}(T_C) \end{bmatrix} = T_{(E,A,B,C)} \begin{bmatrix} \mathrm{vec}(X) \\ \mathrm{vec}(S) \\ \mathrm{vec}(Y) \\ \mathrm{vec}(Z) \\ \mathrm{vec}(U) \\ \mathrm{vec}(V) \\ \mathrm{vec}(W) \end{bmatrix}.$$

Knowing the canonical structure, the explicit expression for the codimension of the orbit of a controllability pair (A, B) is derived in [39], see also [38]. By rewriting the result, it is obvious that the computation of the codimension of $\mathcal{O}(A, B)$ can be done using parts of the expression (4.41) for matrix pencils. Following the notation of (4.41) and the invariants associated with (A, B), the codimension of $\mathcal{O}(A, B)$ is

$$\mathrm{cod}(A, B) = c_{\mathrm{Right}} + c_{\mathrm{Jor}} + c_{\mathrm{Jor,Right}}, \tag{4.42}$$

where

$$c_{\mathrm{Right}} = \sum_{\epsilon_k > \epsilon_l} (\epsilon_k - \epsilon_l - 1),\ c_{\mathrm{Jor}} = \sum_{i=1}^{q} \sum_{k=1}^{g_i} (2k-1) h_k^{(i)},\ \text{and } c_{\mathrm{Jor,Right}} = r_0 \sum_{i=1}^{q} \sum_{k=1}^{g_i} h_k^{(i)}.$$

The codimension of the orbit of an observability pair (A, C) is easily derived by its duality to (A, B). The codimension of $\mathcal{O}(A, C)$ is

$$\mathrm{cod}(A, C) = c_{\mathrm{Left}} + c_{\mathrm{Jor}} + c_{\mathrm{Jor,Left}}, \tag{4.43}$$

where

$$c_{\mathrm{Left}} = \sum_{\eta_k > \eta_l} (\eta_k - \eta_l - 1),\ c_{\mathrm{Jor}} = \sum_{i=1}^{q} \sum_{k=1}^{g_i} (2k-1) h_k^{(i)},\ \text{and } c_{\mathrm{Jor,Left}} = l_0 \sum_{i=1}^{q} \sum_{k=1}^{g_i} h_k^{(i)}.$$

Expressed in terms of the structure invariants of the system, the codimension for matrix quadruples and matrix triples were derived by García-Planas and Magret. The explicit expression for the codimension of a matrix triple is derived in [51] and the explicit expression for a matrix quadruple is presented, but not derived, in [49]. However, parts of the results provided in [49] and [51] seem to be incorrect. The terms coming from the interaction between the N blocks should depend on the existence of L and L^T blocks[4]. This can be seen by studying the versal deformations of the corresponding system pencil. We have included our revised version of their results in A.

When computing the codimension of the corresponding bundle, the same relation holds for pairs, triples and quadruples associated with a system pencil $\mathbf{S}(\lambda)$, as for matrices and matrix pencils:

$$\mathrm{cod}(\mathcal{B}(\mathbf{S}(\lambda))) = \mathrm{cod}(\mathcal{O}(\mathbf{S}(\lambda))) - (\text{number of distinct eigenvalues}).$$

[4]See Equations (A.66), (A.67), (A.76), and (A.77) of A. See also the original rules in [51, Table 1; Eq. (2) and (10)] and [49, p. 881].

5 Stratification of orbits and bundles

Computing the canonical structure of a system is an ill-posed problem; the system may be sensitive to small perturbations, e.g., small changes in the input data may drastically change the computed canonical structure. Besides knowing the canonical structure, it is equally important to be able to identify nearby canonical structures in order to explain the behavior of the state-space system under small perturbations. For example, a state-space system which is found to be controllable may be very close to an uncontrollable system, and can therefore by only a small change in some data, e.g., due to round-off errors, become uncontrollable.

A *stratification* gives the closure hierarchy of orbits and bundles of canonical structures, i.e., it shows which structures are near to each other (in the sense of small perturbations) and their relation to other structures. For square matrices, Arnold [2] examined nearby structures by small perturbations using versal deformations. For matrix pencils, the theory was first introduced for the set of 2-by-3 matrix pencils by Elmroth and Kågström [37] and later extended in collaboration with Edelman to general matrices and matrix pencils [30, 31]. In line of this work, the theory has further been developed in [34] by Elmroth, P. Johansson and Kågström, and for matrix pairs together with S. Johansson in [33, 36, 76]. Other interesting papers have been published by Berg and Kwatny [6, 7], Boley [10], Garcia-Planas and Magret [44], and Pervouchine [96].

In the following, when it is clear from context we sometimes use the shorter term structure when we refer to a canonical structure. Moreover, in the graph representation used below a *downward path* is defined as a path for which all edges start in a node and end in another node below in the graph. Similarly, an *upward path* is a path in the opposite direction.

Based on the theory in [30, 31, 36], a software tool, StratiGraph [34, 70, 73, 74], has been developed for computing and visualizing the stratification. The stratification is represented as a connected graph where the nodes correspond to orbits (or bundles) of different canonical structures and the edges to their covering relations. Given a node for a canonical structure, its closure is represented by the node itself and all nodes which can be reached by a downward path. In Figure 2, we can see how such a stratification can be represented graphically. The graph illustrates the complete stratification of bundles of all 111 structurally different 7×7 matrices, where each bundle of the different Jordan structures is represented by one of the 111 nodes. Indeed, the size of a graph grows exponentially with the matrix size.

In the graph, it is always possible to go from any canonical structure to another higher up in the graph by a small perturbation if and only if they are connected by an upward path. The other way around is normally not possible, i.e., a structure does not have to be near a structure below in the graph. However, the cases when a structure below in the hierarchy actually is nearby is often of particular interest, as it shows that a more degenerate structure can be found by a small perturbation.

The stratification can be characterized as follows. First, the codimension determines the level in the graph on which the canonical structure resides. We remark that several structures can have the same codimension and therefore are on the same level in the graph. In Figure 2, the codimension is shown on the left side of the graph. Second, the cover relations give the connected structure(s) above or below in the closure hierarchy and guarantee that there is no structure in between. Two structures that have the same codimension cannot cover each other, instead they belong to different branches in the graph. Third, the *most generic* structure is the one with the lowest codimension and is therefore the topmost node in the graph. In Figure 2, the top node corresponds to the most generic orbit of 7×7 diagonalizable matrices with 7 distinct eigenvalues (JCF: $J_1(\mu_1) \oplus J_1(\mu_2) \oplus \cdots \oplus J_1(\mu_7)$; $\mu_i \neq \mu_j$ for $i \neq j$). The *most degenerate* (or the least generic) structure is the one with the highest codimension

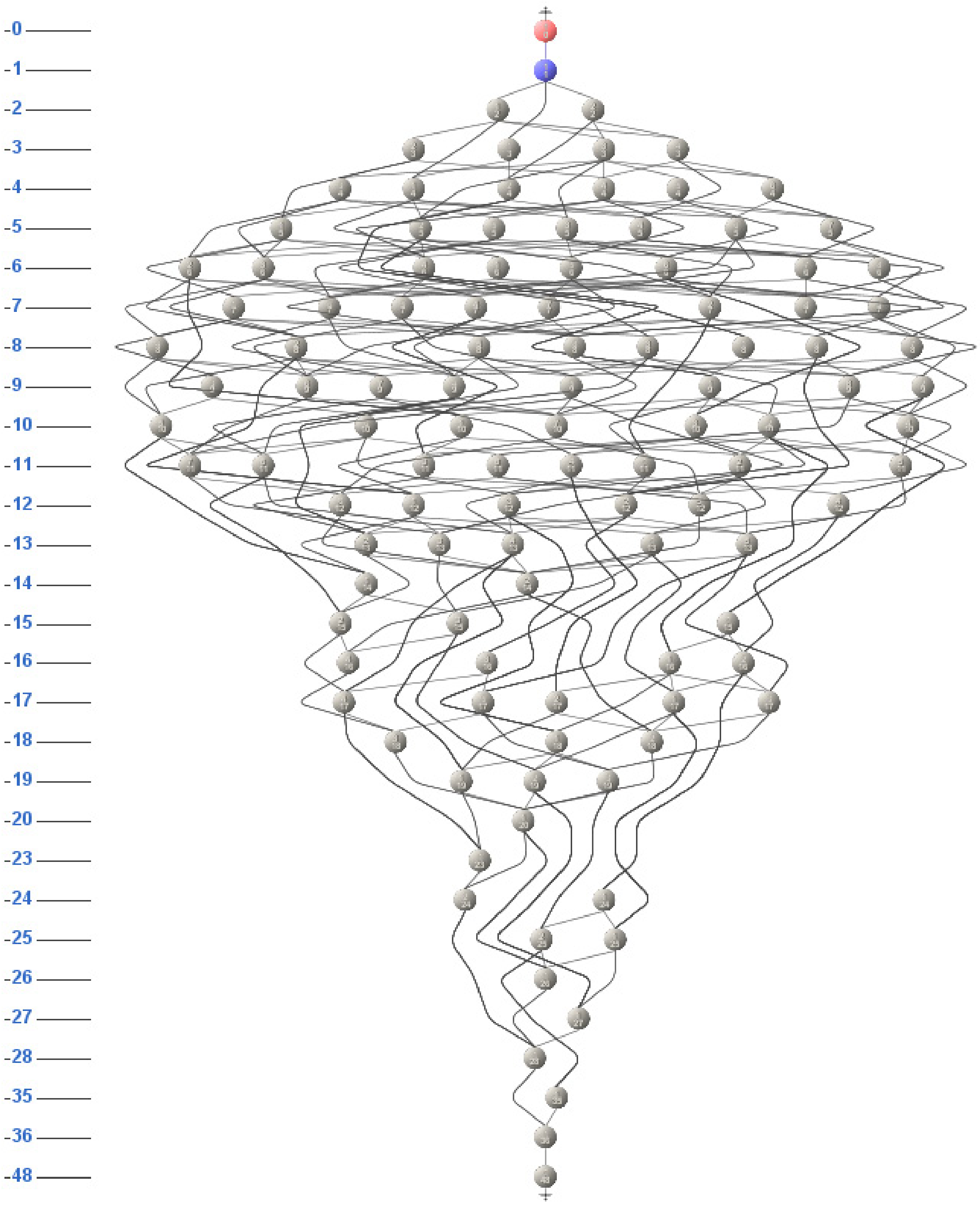


FIGURE 2: The complete stratification of the bundle to a 7×7 matrix give 111 nodes and 313 edges. The numbers on the left show the codimension of the nodes on each level. The graph is generated using StratiGraph v 3.0.

and is consequently the bottom node. In Figure 2, the bottom node corresponds to the most degenerate orbit and Jordan structure, i.e., a nilpotent matrix in diagonal form (the zero matrix with JCF: $J_1(\mu_1) \oplus \cdots \oplus J_1(\mu_7)$; $\mu_i = 0$, $\forall i$). All other 109 nodes represent matrix orbits corresponding to all possible combinations of Jordan blocks (and eigenvalues), whose sizes add up to seven. The edges represent closure relations in the orbit hierarchy.

How the codimensions are computed have already been discussed in Section 4, the most generic and degenerate cases are considered in Section 5.2, and the cover relations in Section 5.3. We end this section by discussing the stratification of a small state-space system in Section 5.4.

5.1 Integer partitions and coins

Before we go any further into the theory of stratification, we define some more properties for integer partitions, which were introduced in Section 2.4. Following [30, 31], the integer partitions are used to express the stratification rules with combinatorial rules acting on these partitions. These rules are called minimum coin moves.

For integer partitions we use standard vector operations and if $\kappa = (\kappa_1, \kappa_2, \ldots)$, $\kappa_1 \geq \kappa_2 \geq \cdots \geq 0$, is an integer partition of an integer K and m is a scalar, then we denote the sum $\kappa_1 + \kappa_2 + \cdots$ as $\sum \kappa$ and $(\kappa_1 + m, \kappa_2 + m, \ldots)$ as $\kappa + m$. We also recall from Section 2.4 the following operations on integer partitions. The union of two integer partitions κ and ν is denoted by $\kappa \cup \nu$, the difference by $\kappa \setminus \nu$, and the conjugate of κ is denoted by $\mathrm{conj}(\kappa)$.

If $\nu = (\nu_1, \nu_2, \ldots)$ is a second integer partition (not necessarily of the same integer K as κ) and $\kappa_1 + \cdots + \kappa_i \geq \nu_1 + \cdots + \nu_i$ for $i = 1, 2, \ldots$, then $\kappa \geq \nu$. Note that, if $\sum \kappa = \sum \nu$ then $\kappa \leq \nu$ if and only if $\mathrm{conj}(\kappa) \geq \mathrm{conj}(\nu)$. We say that κ *dominates* ν or $\kappa > \nu$, if $\kappa \geq \nu$ and $\kappa \neq \nu$. If κ, ν and τ are integer partitions of the same integer K and there does not exist any τ such that $\kappa > \tau > \nu$ where $\kappa > \nu$, then κ *covers* ν. It follows that κ covers ν if and only if $\kappa > \nu$ and $\mathrm{conj}(\kappa) < \mathrm{conj}(\nu)$.

A weaker definition of cover is *adjacent* [22, 65], where κ and ν can be partitions of different integers. We say that $\kappa > \nu$ are adjacent partitions if either κ covers ν or if $\kappa = \nu \cup (1)$.

An integer partition $\kappa = (\kappa_1, \ldots, \kappa_n)$ can also be represented by n piles of coins, where the first pile has κ_1 coins, the second κ_2 coins and so on. This representation is used by Edelman, Elmroth and Kågström [31] to construct the stratification rules. They also defined the following sets of rules on the coin representation.

- *Minimum rightward coin move* on κ: Move one coin *one* column rightward or *one* row downward, and keep κ monotonically decreasing.

- *Minimum leftward coin move* on κ: Move one coin *one* column leftward or *one* row upward, and keep κ monotonically decreasing.

In Figure 3, a *Hasse diagram* and the corresponding piles of coins are illustrated for the integer partition of $K = 6$, where two covering partitions are nearest neighbours. For example, the integer partition $\kappa = (5, 1)$ covers $\nu = (4, 2)$.

In [31], it was shown that the two coin moves defined above can be used to find covering partitions above and below a given partition (see Figure 4).

Theorem 5.1 [16, 31]

(a) An integer partition κ covers *ν if ν can be obtained from κ by a minimum rightward coin move on κ.*

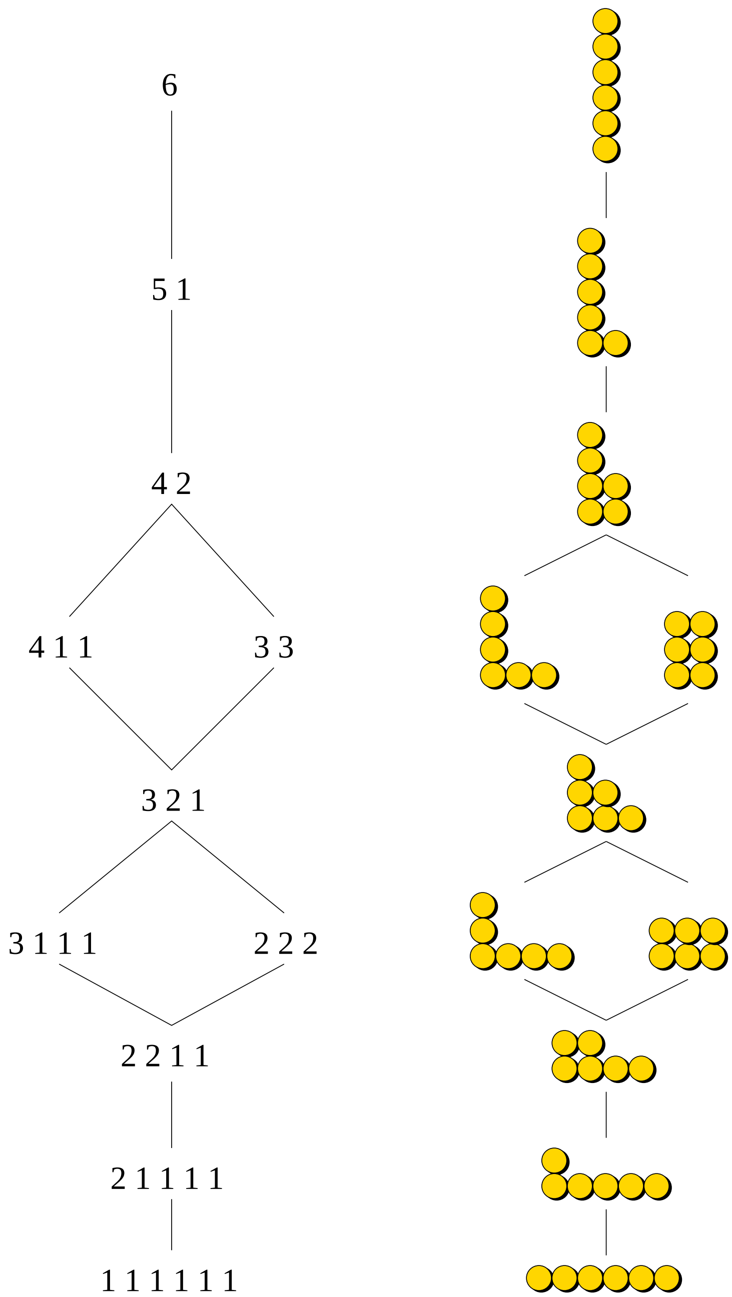


FIGURE 3: Example of a covering relationship with six coins.

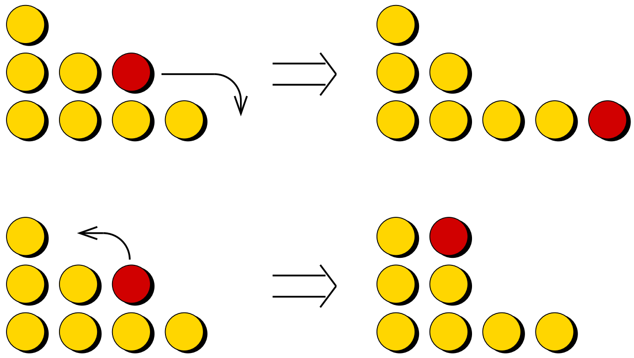


FIGURE 4: Minimum rightward and leftward coin moves illustrate that $\kappa = (3, 2, 2, 1)$ covers $\nu = (3, 2, 1, 1, 1)$ and $\kappa = (3, 2, 2, 1)$ is covered by $\tau = (3, 3, 1, 1)$.

(b) An integer partition κ is covered by *τ if τ can be obtained from κ by a minimum leftward coin move on κ.*

We can also illustrate the conjugate operation with coins, which is obtained by transposing the coins on the anti-diagonal as in Figure 5.

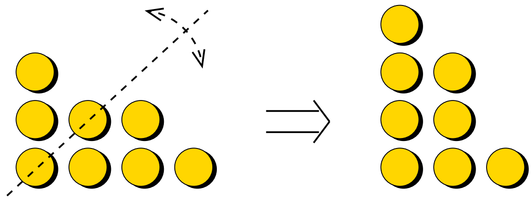


FIGURE 5: Conjugate of the partition $(3, 2, 2, 1)$ is $(4, 3, 1)$.

5.2 Most and least generic cases

Almost all systems of the same size and type (matrices, matrix pencils, etc.) have the same canonical structure. This canonical structure corresponds to the most generic case and has the lowest codimension in the closure hierarchy. This follows from the concept of generic (above called most generic) in [115].

Definition 5.1 [44, 115] *A submanifold Y of a manifold X is called* generic *(or most generic) if it is open, dense and its boundary is the union of submanifolds of (strictly) lower dimensions.*

The opposite case, which in the nilpotent matrix case corresponds to the zero matrix, is the least generic case, or equivalently, the most degenerate case and it has the highest codimension. In the closure hierarchy graph, the most generic case is represented by the topmost node and the most degenerate case by the bottom node. The canonical structures in between correspond to degenerate (or non-generic) cases. In the following, the most and least generic structures for matrices, matrix pencils and system pencils are expressed in their canonical structure and the structure integer partitions $\mathcal{R}$ and $\mathcal{L}$, see Section 2.1–2.4.

We remark that the formulae, discussed below, to compute the most generic structure only hold if there are no restrictions on the matrix, matrix pencil, or system pencil. Otherwise, for example when the matrix pencil has a special structure or fixed rank, the restrictions must be considered when determining the most and least generic cases. There can even exist several most generic structures, but only one with codimension 0 (if it exists). This has recently been studied for general matrix pencils in, e.g., [20, 21, 69].

For general $n \times n$ matrices, the most generic canonical structure has n J_1 blocks corresponding to n distinct eigenvalues. The associated orbit has codimension n and consequently the bundle has codimension $n - n = 0$. For orbits of nilpotent matrices the most generic case has one Jordan block of size $n \times n$ and the codimension of its orbit is n. The most degenerate structure is the one with n J_1 blocks corresponding to a single eigenvalue of multiplicity n, which for orbits has the codimension n^2 and for bundles $n^2 - 1$. Hence, the orbit corresponding to the most degenerate Jordan structure is only a point in the matrix space. In the bundle case, the matrix has one degree of freedom given by the unspecified value of the eigenvalue. See for example Figure 2, where the most degenerate structure of a bundle for a 7×7 matrix has codimension $7 \cdot 7 - 1 = 48$.

For a non-square matrix pencil of size $m_{\mathrm{p}} \times n_{\mathrm{p}}$ the most generic case with $d = n_{\mathrm{p}} - m_{\mathrm{p}} > 0$ has $\mathcal{R} = (r_0, \ldots, r_{\alpha+1})$ where $r_0 = \cdots = r_\alpha = d$ and $r_{\alpha+1} = c$ with $\alpha = \lfloor m_{\mathrm{p}}/d \rfloor$ and $c = m_{\mathrm{p}} \bmod d$ [23, 106]. The same statement holds for $d = m_{\mathrm{p}} - n_{\mathrm{p}} > 0$ by only replacing the partition $\mathcal{R}$ with $\mathcal{L}$. It follows that the most generic structures for non-square matrix pencils are equivalent to

$$G - \lambda H = \begin{bmatrix} 0 & I_{m_{\mathrm{p}}} \end{bmatrix} - \lambda \begin{bmatrix} I_{m_{\mathrm{p}}} & 0 \end{bmatrix}, \qquad \text{if } m_{\mathrm{p}} < n_{\mathrm{p}},$$

and

$$G - \lambda H = \begin{bmatrix} 0 \\ I_{n_{\mathrm{p}}} \end{bmatrix} - \lambda \begin{bmatrix} I_{n_{\mathrm{p}}} \\ 0 \end{bmatrix}, \qquad \text{if } m_{\mathrm{p}} > n_{\mathrm{p}}.$$

The most generic canonical structure for a square matrix pencil of size $n_{\mathrm{p}} \times n_{\mathrm{p}}$ consists only of a regular part with n_{p} distinct finite eigenvalues, i.e., it is diagonalizable and $\det(G - \lambda H) = 0$ if and only if λ is an eigenvalue. The most generic structures for square singular matrix pencils have $r_0 = \cdots = r_j = 1$ and $l_0 = \cdots = l_{n_{\mathrm{p}}-j-1} = 1$, $j = 0, \ldots, n_{\mathrm{p}} - 1$ [113], i.e., the number of most generic square singular matrix pencils is n_{p}. The most degenerate case, both for a square and non-square matrix pencil, corresponds to the zero pencil $G - \lambda H = \mathbf{0}_{m_{\mathrm{p}} \times n_{\mathrm{p}}} - \lambda \mathbf{0}_{m_{\mathrm{p}} \times n_{\mathrm{p}}}$ and has $\mathcal{R} = (n_{\mathrm{p}})$ and $\mathcal{L} = (m_{\mathrm{p}})$, i.e., n_{p} L_0 blocks and m_{p} L_0^T blocks.

The most generic cases for a matrix quadruple and a matrix triple depend on the dimensions of the corresponding $(n + p) \times (n + m)$ system pencil

$$\mathbf{S}(\lambda) = \begin{bmatrix} A & B \\ C & D \end{bmatrix} - \lambda \begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix}.$$

For matrix quadruples the most generic case is [19]:

(1) If $m > p$, let $d = m - p$, $\alpha = \lfloor n/d \rfloor$ and $c = n \bmod d$. Then the most generic structure has p N_1 blocks and $\mathcal{R} = (r_0, \ldots, r_{\alpha+1})$ where $r_0 = \cdots = r_\alpha = d$ and $r_{\alpha+1} = c$.

(2) If $p > m$, let $d = p - m$, $\alpha = \lfloor n/d \rfloor$ and $c = n \bmod d$. Then the most generic structure has m N_1 blocks and $\mathcal{L} = (l_0, \ldots, l_{\alpha+1})$ where $l_0 = \cdots = l_\alpha = d$ and $l_{\alpha+1} = c$.

(3) If $m = p$, the most generic structure has m N_1 blocks and n J_1 blocks with distinct eigenvalues.

For matrix triples the most generic case is [19]:

(1) If $m > p$ and $n \geq p$, let $d = m - p$, $\alpha = \lfloor (n-p)/d \rfloor$ and $c = (n-p) \bmod d$. Then the most generic structure has p N_2 blocks and $\mathcal{R} = (r_0, \ldots, r_{\alpha+1})$ where $r_0 = \cdots = r_\alpha = d$ and $r_{\alpha+1} = c$.

(2) If $p > m$ and $n \geq m$, let $d = p - m$, $\alpha = \lfloor (n-m)/d \rfloor$ and $c = (n-m) \bmod d$. Then the most generic structure has m N_2 blocks and $\mathcal{L} = (l_0, \ldots, l_{\alpha+1})$ where $l_0 = \cdots = l_\alpha = d$ and $l_{\alpha+1} = c$.

(3) If $m = p$ and $n \geq m$, the most generic structure has m N_2 blocks and $n - m$ J_1 blocks with distinct eigenvalues.

(4) If $n < \min\{m, p\}$, the most generic structure has n N_2 blocks, $m - n$ L_0 blocks and $p - n$ L_0^T blocks.

We remark that in the orbit case for both quadruples and triples, where the regular part is nilpotent, case (3) gives one Jordan block of size n and $n - m$, respectively. Moreover, the most generic matrix quadruple has N_1 blocks, while the most generic triple has N_2 blocks. The larger N blocks for matrix triples are a consequence of that the smallest N block a matrix triple can have is of size 2×2, see Section 2.6.

The most degenerate cases for triples and quadruples have n J_1 blocks with equal eigenvalues, m L_0, and p L_0^T blocks.

Example 9

Given a system pencil $\mathbf{S}(\lambda)$ with $n = 2$, $m = 3$, and $p = 1$, we illustrate how the most generic and the most degenerate canonical structures for matrix quadruples and matrix triples are derived.

We begin with the matrix quadruple associated with $\mathbf{S}(\lambda)$. Since $m > p$ we use case (1) above:

$$\begin{aligned} d &= m - p = 3 - 1 = 2, \\ \alpha &= \lfloor n/d \rfloor = \lfloor 2/2 \rfloor = 1, \text{ and} \\ c &= n \bmod d = 2 \bmod 2 = 0. \end{aligned}$$

We get $\mathcal{R} = (2, 2, 0)$ corresponding to two L_1 blocks, and one ($p = 1$) N_1 block, i.e., the most generic structure for a matrix quadruple associated with $\mathbf{S}(\lambda)$ has the KCF $2L_1 \oplus N_1$.

For the associated matrix triple we also use the corresponding case (1) ($m > p$ and $n \geq p$):

$$\begin{aligned} d &= m - p = 3 - 1 = 2, \\ \alpha &= \lfloor (n-p)/d \rfloor = \lfloor (2-1)/2 \rfloor = 0, \text{ and} \\ c &= (n-p) \bmod d = (2-1) \bmod 2 = 1. \end{aligned}$$

We get $\mathcal{R} = (2, 1)$ corresponding to one L_0 block and one L_1 block, and one ($p = 1$) N_2 block, i.e., the KCF $L_1 \oplus L_0 \oplus N_2$.

The most degenerate structure, both for the matrix quadruple and matrix triple, is $3L_0 \oplus L_0^T \oplus 2J_1(\mu_1)$.

The most generic structure of the controllability pair (A, B) has $\mathcal{R} = (r_0, \dots, r_\alpha, r_{\alpha+1})$ where $r_0 = \cdots = r_\alpha = m$, $r_{\alpha+1} = n \bmod m$, and $\alpha = \lfloor n/m \rfloor$ [57]. For the observability pair (A, C) the most generic structure has $\mathcal{L} = (l_0, \dots, l_\alpha, l_{\alpha+1})$ where $l_0 = \cdots = l_\alpha = p$, $l_{\alpha+1} = n \bmod p$, and $\alpha = \lfloor n/p \rfloor$. The most degenerate case of (A, B) has m L_0 blocks and n Jordan blocks of size 1×1 corresponding to an eigenvalue of multiplicity n. Similarly, (A, C) has p L_0^T blocks and n 1×1 Jordan blocks. In other words, the most generic cases correspond to completely controllable and observable systems, while the most degenerate cases correspond to systems with n uncontrollable and n unobservable multiple modes, respectively.

Example 10

We use the same system pencil as in Example 9 to illustrate how the most generic and the most degenerate canonical structures for the matrix pairs (A, B) and (A, C), respectively, are computed.

For the matrix pair (A, B) with $n = 2$ and $m = 3$ we have $\alpha = \lfloor 2/3 \rfloor = 0$ and $r_1 = 2 \bmod 3 = 2$. Hence, the most generic structure has $\mathcal{R} = (3, 2)$, giving the KCF $2L_1 \oplus L_0$. The most degenerate structure has the KCF $3L_0 \oplus 2J_1(\mu_1)$.

The matrix pair (A, C) with $n = 2$ and $p = 1$ has $\alpha = \lfloor 2/1 \rfloor = 2$ and $l_3 = 2 \bmod 1 = 0$. So, the most generic structure has $\mathcal{L} = (1, 1, 1, 0)$, i.e., the KCF L_2^T, and the most degenerate structure is $L_0^T \oplus 2J_1(\mu_1)$.

Example 11

In this example, we illustrate how a fixed structure of the system matrices A and B in a state-space system $\dot{x}(t) = Ax(t) + Bu(t)$ can restrict the form of the most and least generic cases.

Consider a controllability pair (A, B) of the same size as in Example 10 associated with the state-space system

$$\dot{x}(t) = \begin{bmatrix} 1 & 0 \\ 1.45 & 2.5 \end{bmatrix} x(t) + \begin{bmatrix} 3.0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} u(t), \tag{5.44}$$

with the corresponding 2×5 controllability system pencil

$$\mathbf{S}_{\mathrm{C}}(\lambda) = \left[\begin{array}{cc|ccc} 1 & 0 & 3.0 & 0 & 0 \\ 1.45 & 2.5 & 1 & 0 & 0 \end{array}\right] - \lambda \begin{bmatrix} I_2 & 0 \end{bmatrix}.$$

Moreover, let all zeros and ones in the system matrices A and B be fixed.

We first determine the most generic structure. For a controllability pair the number of L_0 blocks is $m - \operatorname{rank}(B)$ (see end of Section 2.6). It follows that for the system (5.44) the number of L_0 blocks must be at least $3 - \operatorname{rank}(B) = 2$ (we have two fixed columns of zeros). Since the most generic structure has two L_0 blocks: $r_0 = \cdots = r_\alpha = m$,

$$\begin{aligned} r_{\alpha+1} &= m - 2 = 3 - 2 = 1, \text{ and} \\ r_{\alpha+1} &= n \bmod m = n - \lfloor n/m \rfloor m = n - \alpha m = 2 - 3\alpha. \end{aligned}$$

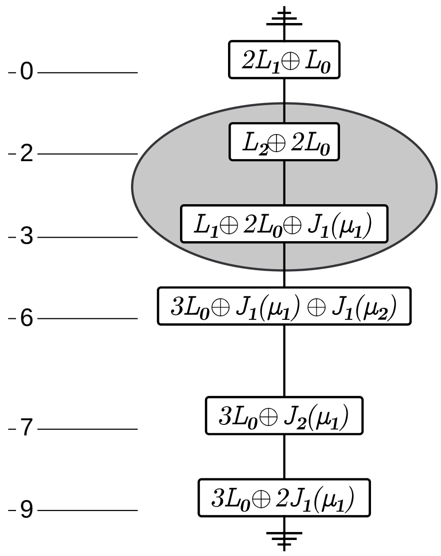


FIGURE 6: The complete bundle stratification of a 2×5 controllability pencil. The grey area marks the possible canonical structures for the state-space system (5.44). The numbers on the left show the codimension of the orbits on each level.

From the above two equations $\alpha = 1$ and, consequently, the most generic structure has $\mathcal{R} = (3, 3, 1)$ corresponding to the KCF $L_2 \oplus 2L_0$.

The most degenerate structure is determined by studying the system with all free variables in A and B set to zero:

$$\dot{x}(t) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} u(t),$$

which has the KCF $L_1 \oplus 2L_0 \oplus J_1(1)$. That the most degenerate structure has one L_k block, where $k > 0$, follows from that we have one nonzero column in B. In this case $k = 1$. Consequently, the total size of the Jordan structure is $n - k = 2 - 1 = 1$.

Figure 6 displays the complete stratification of the bundles of 2×5 controllability pencils. The grey area marks the possible canonical structures for the state-space system (5.44) with all zeros and ones fixed. As we can see, the only possible structures are the two we have derived above. How the complete bundle stratification in Figure 6 is determined is explained in Section 5.4.

5.3 Closure and cover relations

To determine the closure hierarchy for $n \times n$ matrices we stratify the n^2-dimensional matrix space into similarity orbits (or bundles). Similarly, the closure hierarchy for $m_{\mathrm{p}} \times n_{\mathrm{p}}$ matrix pencils is given by the stratification of strictly equivalence orbits (or bundles) in the $2m_{\mathrm{p}}n_{\mathrm{p}}$-dimensional matrix pencil space. The stratification of orbits or bundles is given from the closure relations and further the cover relations between these manifolds, see Arnold [2] and

[30, 31]. An orbit *covers* another orbit if its closure includes the closure of the other orbit and there is no orbit in between in the closure hierarchy, i.e., they are nearest neighbours in the hierarchy. The closure and cover relations for bundles are defined analogously.

We want a procedure to decide if an orbit closure is a (proper) subset of another orbit closure. We start by observing that if two systems with the corresponding orbits $\mathcal{O}_1$ and $\mathcal{O}_2$ are equivalent, then $\overline{\mathcal{O}}_1 = \overline{\mathcal{O}}_2$ where $\overline{\mathcal{O}}$ denotes the orbit closure. Furthermore, an orbit $\mathcal{O}_2$ which lies in the closure of $\mathcal{O}_1$ is less generic, i.e., $\dim(\mathcal{O}_2) < \dim(\mathcal{O}_1)$. In general, this follows from the *closed orbit lemma* [68], where the differentiable manifold X can be a matrix A, a matrix pencil $G - \lambda H$, or a system pencil $S - \lambda T$.

Theorem 5.2 [68] *Let an equivalence transformation act on a differentiable manifold X. Then each orbit is a smooth, locally closed subset of X, whose boundary is a union of orbits of strictly lower dimension. In particular, orbits of minimal dimension are closed (so closed orbits exist).*

In the following, we show the requirements on the canonical forms corresponding to $\mathcal{O}_1$ and $\mathcal{O}_2$ such that $\overline{\mathcal{O}}_1 \supseteq \overline{\mathcal{O}}_2$, i.e., that the closure of $\mathcal{O}_2$ lies in the closure of $\mathcal{O}_1$, for matrices, matrix pencils and matrix pairs. The closure and cover relations for orbits of matrix quadruples and matrix triples are not yet completely determined, and are therefore not considered in this paper.

Starting with matrices, the theory behind the closure decision problem for orbits of nilpotent matrices goes back to 1961. If the matrix has well clustered eigenvalues but is not nilpotent, the blocks associated with the same eigenvalue can be shifted to a nilpotent matrix and the same theory can be used. For example, given a matrix A with eigenvalues $\mu_1, \ldots, \mu_q$. Order the Jordan blocks such that $A = \operatorname{diag}(A_1, \ldots, A_q)$, where A_i contains all Jordan blocks associated with the eigenvalue μ_i, for $i = 1, \ldots, q$. In order to study closure and cover relations related to the eigenvalue μ_i, the matrix can be shifted ($\widetilde{A} = A - \mu_i I$) so that the block $\widetilde{A}_i = A_i - \mu_i I$ is nilpotent.

The closure conditions for orbits of matrices are given by the following theorem, where the integer partition h_{μ_i} represents the Segre characteristics and $\mathcal{J}_{\mu_i}$ the Weyr characteristics for the finite eigenvalue μ_i, and q is the number of distinct eigenvalues.

Theorem 5.3 [2, 31] $\overline{\mathcal{O}}(A_1) \supseteq \overline{\mathcal{O}}(A_2)$ if and only if $\mathcal{J}_{\mu_i}(A_1) \leq \mathcal{J}_{\mu_i}(A_2)$ *and* $h_{\mu_i}(A_1) \geq h_{\mu_i}(A_2)$*, for all* $\mu_i \in \mathbb{C}$*,* $i = 1, \ldots, q$*.*

From Theorem 5.3 it follows that the number of eigenvalues and the total size of all blocks associated with the same eigenvalue, are the same for all orbits in the closure hierarchy. This in contrast to the bundle case where eigenvalues can coalesce or split apart.

Next, we consider the cover relations for orbits of matrices. This can be obtained from Theorem 5.3 and the definition of covering partitions. Here we give the cover relations in form of coin moves on the structure integer partition $\mathcal{J}_{\mu_i}$ as presented in [31].

Theorem 5.4 [2, 31] $\mathcal{O}(A_1)$ *covers* $\mathcal{O}(A_2)$ if and only if *some* $\mathcal{J}_{\mu_i}(A_2)$ *can be obtained from* $\mathcal{J}_{\mu_i}(A_1)$ *by a minimum leftward coin move, and* $\mathcal{J}_{\mu_j}(A_2) = \mathcal{J}_{\mu_j}(A_1)$ *for all* $\mu_j \neq \mu_i$*.*

In the case of not well-clustered eigenvalues, we have to consider the bundle case as defined by Arnold [2]. The solution to the closure decision problem for matrix bundles is given in Theorem 5.5, where coalescing two eigenvalues α and β is equivalent to take the union of the two corresponding integer partitions $\mathcal{J}_\alpha$ and $\mathcal{J}_\beta$. We remark that, even if testing for closure relations between nilpotent matrices is trivial, deciding if one bundle is in the closure of another bundle is an NP-complete problem [60] (see also [31]). The conditions for covering relations expressed in terms of coin moves are given in Theorem 5.6. We have summarized the stratification rules to find a covering or a covered orbit/bundle to a matrix in Table 2 of B.

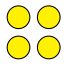
Theorem 5.5 **[28, 31, 87]** *If $\mathcal{B}(A_1)$ has at least as many distinct eigenvalues as $\mathcal{B}(A_2)$, then $\overline{\mathcal{B}}(A_1) \supseteq \overline{\mathcal{B}}(A_2)$* if and only if *it is possible to coalesce eigenvalues and apply the dominance ordering coin moves to the structure integer partitions of the bundle defined by A_1 to reach that of A_2.*

Theorem 5.6 **[31]** *$\mathcal{B}(A_1)$ covers $\mathcal{B}(A_2)$* if and only if *some $\mathcal{J}_{\mu_i}(A_2)$ either can be obtained from $\mathcal{J}_{\mu_i}(A_1)$ by a minimum leftward coin move or by coalescing the partitions from two distinct eigenvalues (and $\mathcal{J}_{\mu_j}(A_2) = \mathcal{J}_{\mu_j}(A_1)$ for all other eigenvalues μ_j).*

Example 12

Let A be a 7×7 matrix with JCF $2J_2(\mu_1) \oplus J_1(\mu_2) \oplus J_1(\mu_3) \oplus J_1(\mu_4)$. Using the stratification rules in Table 2.C (also given in Theorem 5.6 above) and Table 2.D of B, we show how to derive all nearest neighbours to the matrix A in a bundle stratification. The complete bundle stratification of all 7×7 matrixces is shown in Figure 2, where the matrix A is represented by one of the nodes with codimension 7 (the fourth node from the right).

The JCF of A is expressed in the Weyr characteristics and its corresponding sets of coins as

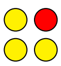
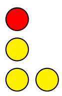
$\mathcal{J}_{\mu_1} = (2, 2);$  ,

$\mathcal{J}_{\mu_2} = (1);$  ,

$\mathcal{J}_{\mu_3} = (1);$  , and

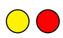
$\mathcal{J}_{\mu_4} = (1);$  .

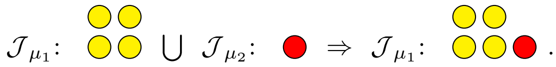
We start with the two rules in Table 2.C to find all covered matrix bundles. First, rule C.1 (minimal leftward coin move) is applied to all sets of coins for which it is feasible. It follows that it can only be applied to the set $\mathcal{J}_{\mu_1}$:

$\mathcal{J}_{\mu_1}$:  $\Rightarrow$ $\mathcal{J}_{\mu_1}$:  .

The remaining sets are unchanged, showing that $\mathcal{B}(A)$ covers the bundle of the matrices with JCF $J_2(\mu_1) \oplus 2J_1(\mu_1) \oplus J_1(\mu_2) \oplus J_1(\mu_3) \oplus J_1(\mu_4)$.

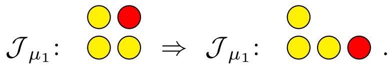
Then we apply rule C.2 (take the union of two sets) to all the sets of coins. Here we have two possibilities, either we take the union of two sets with one coin (whichever two of $\mathcal{J}_{\mu_2}$, $\mathcal{J}_{\mu_3}$ and $\mathcal{J}_{\mu_4}$), e.g.,

$\mathcal{J}_{\mu_2}$:  $\bigcup$ $\mathcal{J}_{\mu_3}$:  $\Rightarrow$ $\mathcal{J}_{\mu_2}$:  , or we take the union of $\mathcal{J}_{\mu_1}$ and one of the sets with one coin, e.g.,

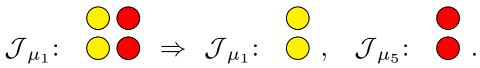


It follows that $\mathcal{B}(A)$ also covers the bundles of the structures with JCF $2J_2(\mu_1) \oplus J_2(\mu_2) \oplus J_1(\mu_4)$ and $J_3(\mu_1) \oplus J_2(\mu_1) \oplus J_1(\mu_3) \oplus J_1(\mu_4)$.

To find all covering matrix bundles the two rules in Table 2.D are used. Rule D.1 (minimal rightward coin move) can only be applied to the set $\mathcal{J}_{\mu_1}$:



With all other sets unchanged, we get that $\mathcal{B}(A)$ is covered by the bundle of the structures with JCF $J_3(\mu_1) \oplus J_1(\mu_1) \oplus J_1(\mu_2) \oplus J_1(\mu_3) \oplus J_1(\mu_4)$. The second rule (divide one set into two) can also only be applied to the set $\mathcal{J}_{\mu_1}$:



This shows that $\mathcal{B}(A)$ is covered by the bundle of the structures with JCF $2J_1(\mu_1) \oplus J_1(\mu_2) \oplus J_1(\mu_3) \oplus J_1(\mu_4) \oplus 2J_1(\mu_5)$.

The bundle closure hierarchy with all covered and covering bundles to $\mathcal{B}(A)$ is shown in Figure 7.

The closure decision problem for orbits of general matrix pencils was solved by Pokrzywa [98] and later reformulated by De Hoyos [19]. Independently, Bongartz [11] derived a similar solution to the problem. Here follows the theorem given in [19] formulated as in [31], where $\mathcal{R} = (r_0, \ldots, r_{\epsilon_1})$, $\mathcal{L} = (l_0, \ldots, l_{\eta_1})$, and $\mathcal{J}_{\mu_i} = (j_1^{(i)}, \ldots, j_{g_i}^{(i)})$ are the structure integer partitions defined in Section 2.4. Moreover, denote by$r_0(G - \lambda H)$ the number of column minimal indices (r_0) for $G - \lambda H$.

Theorem 5.7 [31, 19, 98] $\overline{\mathcal{O}}(G - \lambda H) \supseteq \overline{\mathcal{O}}(\widetilde{G} - \lambda \widetilde{H})$ if and only if *the following relations hold:*

(1) $\mathcal{R}(G - \lambda H) + \mathrm{nrk}\,(G - \lambda H) \geq \mathcal{R}(\widetilde{G} - \lambda \widetilde{H}) + \mathrm{nrk}(\widetilde{G} - \lambda \widetilde{H})$.

(2) $\mathcal{L}(G - \lambda H) + \mathrm{nrk}\,(G - \lambda H) \geq \mathcal{L}(\widetilde{G} - \lambda \widetilde{H}) + \mathrm{nrk}(\widetilde{G} - \lambda \widetilde{H})$.

(3) $\mathcal{J}_{\mu_i}(G - \lambda H) + r_0(G - \lambda H) \leq \mathcal{J}_{\mu_i}(\widetilde{G} - \lambda \widetilde{H}) + r_0(\widetilde{G} - \lambda \widetilde{H})$,
for all $\mu_i \in \overline{\mathbb{C}}$, $i = 1, 2 \ldots$, *where* $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$.

From matrix bundles it follows that deciding if a bundle of a matrix pencil is in the closure of another is also an NP-complete problem. The necessary conditions for an orbit or a bundle of two matrix pencils to be closest neighbours were derived in [11, 19, 98], which was later complemented with the sufficient conditions in [31].

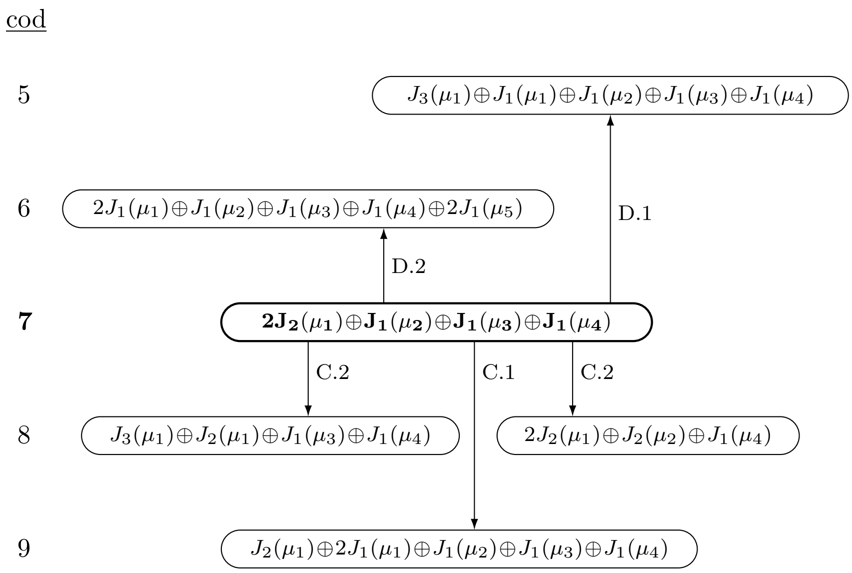


FIGURE 7: All nearest neighbours in the closure hierarchy to the bundle of the matrix with JCF $2J_2(\mu_1) \oplus J_1(\mu_2) \oplus J_1(\mu_3) \oplus J_1(\mu_4)$. Which rule in Table 2 of B used is marked at each edge, and the bundle codimensions are shown to the left.

Theorem 5.8 [31] *Given the structure integer partitions $\mathcal{L}$, $\mathcal{R}$ and $\mathcal{J}_{\mu_i}$ of $G - \lambda H$, where $\mu_i \in \overline{\mathbb{C}}$, one of the following* if-and-only-if *rules finds $\widetilde{G} - \lambda\widetilde{H}$ such that$\mathcal{O}(G - \lambda H)$ covers $\mathcal{O}(\widetilde{G} - \lambda\widetilde{H})$:*

(1) Minimum rightward *coin move in $\mathcal{R}$ (or $\mathcal{L}$).*

(2) If the rightmost column in $\mathcal{R}$ (or $\mathcal{L}$) is one single coin, move that coin to a new rightmost column of some $\mathcal{J}_{\mu_i}$ (which may be empty initially).

(3) Minimum leftward *coin move in any $\mathcal{J}_{\mu_i}$.*

(4) Let k denote the total number of coins in all of the longest (= lowest) rows from all of the $\mathcal{J}_{\mu_i}$. Remove these k coins, add one more coin to the set, and distribute $k+1$ coins to r_p, $p = 0, \dots, t$ and l_q, $q = 0, \dots, k-t-1$ such that at least all nonzero columns of $\mathcal{R}$ and $\mathcal{L}$ are given coins.

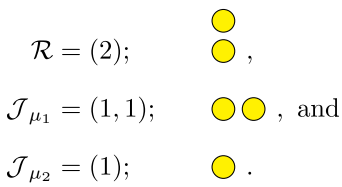
Rules 1 and 2 are not allowed to make coin moves that affect r_0 (or l_0).

Notice that in the above two theorems for matrix pencils, the eigenvalue μ_i corresponding to the structure integer partition $\mathcal{J}_{\mu_i}$ belongs to the extended complex plane, i.e., $\mu_i \in \mathbb{C} \cup \{\infty\}$. Moreover, in Theorem 5.8 the restriction for rules (1) and (2) implies that the number of left and right singular blocks remain fixed, while rule (4) adds one new block of each kind and rule (3) corresponds to the nilpotent case. We also remark that rule (4) cannot be applied if the total number of nonzero columns in $\mathcal{R}$ and $\mathcal{L}$ are more than $k+1$. If the rule can be applied, at least one coin must be assigned to $\mathcal{R}$ and $\mathcal{L}$, respectively. In Table 3 of B, the complete set of rules for cover relations of matrix pencils is given (both for covered/covering orbits and bundles).

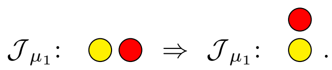
Example 13

Here we consider the orbit stratification of the 3×5 matrix pencil $G - \lambda H$ associated with the state-space system (2.20) in Example 4. The KCF of $G - \lambda H$ is $2L_0 \oplus J_2(\alpha) \oplus J_1(\beta)$, where in (2.20) the eigenvalue $\alpha = \infty$ and the eigenvalue β depends on the value of γ (if $\gamma = 0$ then $\beta = 0$). However, we consider the general case where α and β can be any complex number (including infinity) and therefore denote the eigenvalues by μ_1 and μ_2, respectively. In the following, we show how all matrix pencils covered by $\mathcal{O}(G - \lambda H)$ can be found using the rules in Theorem 5.8 (or Table 3.A of B).

First, we express the KCF $2L_0 \oplus J_2(\mu_1) \oplus J_1(\mu_2)$ in the structure integer partitions $\mathcal{R}$ and $\mathcal{J}$, and its corresponding sets of coins:

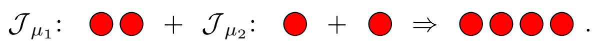


Rule (1) cannot be applied to $\mathcal{R}$ because the rule is not allowed to do any coin moves that affect r_0 (the first column in $\mathcal{R}$). The second rule cannot be used because $\mathcal{R}$ has not a single coin in its rightmost column. However, rule (3) can be applied to the set $\mathcal{J}_{\mu_1}$:

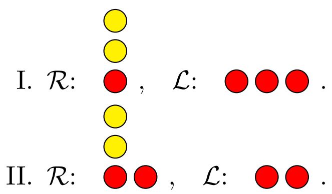


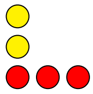
This shows that $\mathcal{O}(G - \lambda H)$ covers the orbit of the matrix pencils with KCF $2L_0 \oplus 2J_1(\mu_1) \oplus J_1(\mu_2)$.

We can also apply the fourth rule which says the following. Take all coins on the lowest rows of all $\mathcal{J}$ and add one coin to the set:



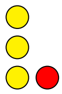
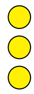
Then distribute the coins on $\mathcal{R}$ and $\mathcal{L}$, such that at least all nonzero columns in $\mathcal{R}$ and $\mathcal{L}$ get one coin each and $\mathcal{R}$ and $\mathcal{L}$ get no less than one coin. The four coins can be distributed in three different ways on the sets $\mathcal{R}$ and $\mathcal{L}$:



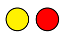
III. $\mathcal{R}$:  , $\mathcal{L}$:  .

The three cases correspond to the matrix pencils $3L_0 \oplus L_2^T$, $L_1 \oplus 2L_0 \oplus L_1^T$, and $L_2 \oplus 2L_0 \oplus L_0^T$, respectively.

To illustrate rule (2), we also derive the orbits of matrix pencils covered by the orbit of $L_1 \oplus 2L_0 \oplus L_1^T$ (case II above). Rule (1) cannot be applied because there are no minimal rightward coin moves that do not affect r_0. For rule (2), there are two choices; either the single rightmost coin in $\mathcal{R}$ or the single rightmost coin in $\mathcal{L}$ is moved to a new set $\mathcal{J}_{\mu_1}$:

$\mathcal{R}$:  $\Rightarrow$ $\mathcal{R}$:  , $\mathcal{J}_{\mu_1}$:  ,

or

$\mathcal{L}$:  $\Rightarrow$ $\mathcal{L}$:  , $\mathcal{J}_{\mu_1}$:  .

The two cases give the matrix pencils with KCF $3L_0 \oplus L_1^T \oplus J_1(\mu_1)$ and $L_1 \oplus 2L_0 \oplus L_0^T \oplus J_1(\mu_1)$. Furthermore, rule (3) and (4) cannot be applied because there is no regular part.

If we derive the orbits that are covered by the orbit of the remaining structures $2L_0 \oplus L_2 \oplus L_0^T$, $2L_0 \oplus 2J_1(\mu_1) \oplus J_1(\mu_2)$, and $3L_0 \oplus L_2^T$, we actually get no more structures than those already derived. The orbit stratification derived above is shown in Figure 8 (the complete stratification of the bundles of 3×5 matrix pencils is shown in Figure 9). We remark that rule (1) is not used in this example, but it is similar to rule (3) and should be straightforward to apply.

Closure conditions for controllability pairs, both necessary and sufficient, have been studied by Gracia, De Hoyos and Zaballa [59], and later by Hinrichsen and O'Halloran [65, 66]. As shown below, the closure conditions are a subset of those for general matrix pencils. Here we give our reformulation and slight modification of the theorem originally presented in [66, Theorem 4.6].

Theorem 5.9 [66, 75] $\overline{\mathcal{O}}(A,B) \supseteq \overline{\mathcal{O}}(\widetilde{A},\widetilde{B})$ if and only if *the following conditions hold:*

(1) $\mathcal{R}(A,B) \geq \mathcal{R}(\widetilde{A},\widetilde{B})$.

(2) $\mathcal{J}_{\mu_i}(A,B) \leq \mathcal{J}_{\mu_i}(\widetilde{A},\widetilde{B})$, *for all* $\mu_i \in \mathbb{C}$, $i = 1,\ldots,q$.

The closure conditions for the observability pair (A,C) are, from the duality with (A,B), equal to those for (A,B) except that $\mathcal{R}$ is replaced by $\mathcal{L}$.

In [66, Theorem 4.6], condition (1) is given as $\mathrm{conj}(\epsilon) \geq \mathrm{conj}(\widetilde{\epsilon})$ and condition (2) as $D_j(A,B)$ *divides* $D_j(\widetilde{A},\widetilde{B})$ for all $j = 1,\ldots,n$, where the integer partition ϵ is the column minimal indices and $D_j(A,B)$ are the greatest common divisors of all minors of (A,B), as defined in Section 2.4. Furthermore, they only prove the theorem for $\overline{\mathcal{O}}(A,B) \supseteq (\widetilde{A},\widetilde{B})$ instead of the more rigid condition $\overline{\mathcal{O}}(A,B) \supseteq \overline{\mathcal{O}}(\widetilde{A},\widetilde{B})$. In the proof, we show that these two closure relations are indeed equal for matrix pairs and that the two conditions in [66,

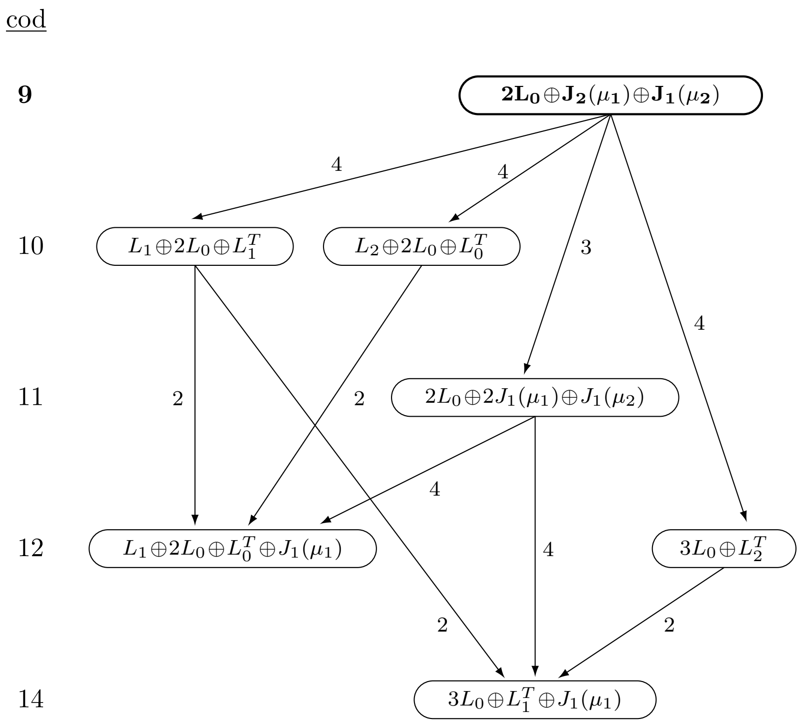


FIGURE 8: A subgraph of all orbits that are in closure of the orbit of the matrix pencil with KCF $2L_0 \oplus J_2(\mu_1) \oplus J_1(\mu_2)$. Which rule of Theorem 5.8 used is marked at each edge, and the orbit codimensions are shown to the left.

Theorem 4.6] can be reformulated as those in Theorem 5.9. The proof is originally presented in [75].

Proof of Theorem 5.9. Notably, in order to conform with our formulations of Theorems 5.3 and 5.7 we write $\overline{\mathcal{O}}(A,B) \supseteq \overline{\mathcal{O}}(\widetilde{A},\widetilde{B})$ instead of $\overline{\mathcal{O}}(A,B) \supseteq (\widetilde{A},\widetilde{B})$ as originally written in [66, Theorem 4.6]. This can be done since $\overline{\mathcal{O}}(A,B)$ consists of the set of all controllability pairs with the canonical form of (A,B) (i.e., $\mathcal{O}(A,B)$) and more degenerate orbits in the closure of $\mathcal{O}(A,B)$. The same holds for $\overline{\mathcal{O}}(\widetilde{A},\widetilde{B})$. Since $\mathcal{O}(\widetilde{A},\widetilde{B})$ is in the closure of $\mathcal{O}(A,B)$, $\overline{\mathcal{O}}(\widetilde{A},\widetilde{B})$ is in $\overline{\mathcal{O}}(A,B)$.

It remains to show that conditions (1) and (2) given in [66, Theorem 4.6] are equal to conditions (1) and (2), respectively, of Theorem 5.9.

Condition (1): Show that $\mathrm{conj}(\epsilon) \geq \mathrm{conj}(\widetilde{\epsilon})$ in [66, Theorem 4.6] is equivalent to $\mathcal{R}(A,B) \geq \mathcal{R}(\widetilde{A},\widetilde{B})$ in Theorem 5.9.

Knowing that $\mathcal{R} = (r_0) \cup \mathrm{conj}(\epsilon)$ where r_0 is the number of L blocks, and that (A,B) always has m L blocks, it follows directly that $\mathrm{conj}(\epsilon) \geq \mathrm{conj}(\widetilde{\epsilon})$ is equivalent to $\mathcal{R}(A,B) \geq \mathcal{R}(\widetilde{A},\widetilde{B})$.

Condition (2): Show that $D_j(A,B)$ divides $D_j(\widetilde{A},\widetilde{B})$ $(j = 1,\ldots,n)$ in [66, Theorem 4.6] is equivalent to $\mathcal{J}_{\mu_i}(A,B) \leq \mathcal{J}_{\mu_i}(\widetilde{A},\widetilde{B})$ (for all $\mu_i \in \mathbb{C}$, $i = 1,\ldots,q$) in Theorem 5.9.

The BCF of the matrix pairs (A, B) and $(\widetilde{A}, \widetilde{B})$ are

$$\begin{bmatrix} A_\epsilon & 0 & B_\epsilon \\ 0 & A_\mu & 0 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} \widetilde{A}_\epsilon & 0 & \widetilde{B}_\epsilon \\ 0 & \widetilde{A}_\mu & 0 \end{bmatrix},$$

respectively, where (A_ϵ, B_ϵ) and $(\widetilde{A}_\epsilon, \widetilde{B}_\epsilon)$ consist of the singular parts and the square pencils A_μ and $\widetilde{A}_\mu$ consist of the regular parts of (A, B) and $(\widetilde{A}, \widetilde{B})$, respectively.

First of all, the size of A_μ is always less or equal to the size of $\widetilde{A}_\mu$. It follows from $\operatorname{cod}(A, B) \leq \operatorname{cod}(\widetilde{A}, \widetilde{B})$ (equality if (A, B) and $(\widetilde{A}, \widetilde{B})$ have the same KCF).

If A_μ and $\widetilde{A}_\mu$ are of the same size we can use the closure conditions in Theorem 5.3 for matrices. We follow the steps by Hinrichsen and O'Halloran in [64, p. 614] to prove the equivalence with the elementary divisors.

It follows from Theorem 5.3 that $\overline{\mathcal{O}}(A_\mu) \supseteq \overline{\mathcal{O}}(\widetilde{A}_\mu)$ if and only if $\mathcal{J}_{\mu_i}(A_\mu) \leq \mathcal{J}_{\mu_i}(\widetilde{A}_\mu)$ and $h_{\mu_i}(A_\mu) \geq h_{\mu_i}(\widetilde{A}_\mu)$, for each eigenvalue μ_i. Recall from Section 2.4 that $h_k^{(i)}$ (in $h_{\mu_i}(A_\mu)$ and $h_{\mu_i}(\widetilde{A}_\mu)$) is the multiplicity of the elementary divisor $\lambda - \mu_i$ in P_k of A_μ and $\widetilde{A}_\mu$, respectively. We get from (2.14) and the closure condition in Theorem 5.3 that $\overline{\mathcal{O}}(A_\mu) \supseteq \overline{\mathcal{O}}(\widetilde{A}_\mu)$ if and only if $D_j(A_\mu)$ divides $D_j(\widetilde{A}_\mu)$, i.e.,

$$\mathcal{J}_{\mu_i}(A_\mu) \leq \mathcal{J}_{\mu_i}(\widetilde{A}_\mu) \quad \textit{if and only if} \quad D_j(A_\mu) \text{ divides } D_j(\widetilde{A}_\mu).$$

Finally, we consider the case when the size of $\widetilde{A}_\mu$ is larger than for A_μ. Let $\widetilde{A}_{reg}$ be a submatrix of $\widetilde{A}_\mu$ with the same eigenvalues as A_μ, where the total size of all blocks corresponding to each of those eigenvalues are at least as large as in A_μ. As shown above it follows that $D_j(A_\mu)$ divides $D_j(\widetilde{A}_{reg})$ and therefore must $D_j(A_\mu)$ divide $D_j(\widetilde{A}_\mu)$. Expressed in Weyr characteristics this corresponds to $\mathcal{J}_{\mu_i}(A_\mu) \leq \mathcal{J}_{\mu_i}(\widetilde{A}_{reg})$, and consequently $\mathcal{J}_{\mu_i}(A, B) = \mathcal{J}_{\mu_i}(A_\mu) \leq \mathcal{J}_{\mu_i}(\widetilde{A}_{reg}) \leq \mathcal{J}_{\mu_i}(\widetilde{A}, \widetilde{B})$, for all $\mu_i \in \mathbb{C}$, $i = 1, \ldots, q$.
□

Theorem 5.10 [36] *If $\mathcal{B}(A, B)$ has at least as many distinct eigenvalues as $\mathcal{B}(\widetilde{A}, \widetilde{B})$, then* $\overline{\mathcal{B}}(A, B) \supseteq \overline{\mathcal{B}}(\widetilde{A}, \widetilde{B})$ if and only if *the following conditions hold:*

(1) $\mathcal{R}(A, B) \geq \mathcal{R}(\widetilde{A}, \widetilde{B})$.

(2) It is possible to coalesce eigenvalues and apply the dominance ordering coin moves to $\mathcal{J}_{\mu_i}(A, B)$, for any μ_i, to reach $(\widetilde{A}, \widetilde{B})$.

Proof. The theorem follows directly from Theorem 5.9 and the closure condition for matrix bundles presented in [31]. □

In [65], also the necessary conditions for cover relations of matrix pencils with no row minimal indices have been derived. They are summarized in Proposition 5.11 with some minor reformulations. We remark that a matrix pencil $G - \lambda H$ with no row minimal indices can have infinite elementary divisors, which is not the case for a controllability pair (A, B). As defined in Section 2.4, the invariants of $G - \lambda H$ and $\widetilde{G} - \lambda \widetilde{H}$ (column minimal indices and Segre characteristics of finite and infinite eigenvalues) are

$$\begin{aligned} &\epsilon = (\epsilon_1, \ldots, \epsilon_{r_0}),\ h_{\mu_i} = (h_1^{(i)}, \ldots, h_{g_i}^{(i)}),\ s = (s_1, \ldots, s_{g_\infty}), \quad \text{and} \\ &\widetilde{\epsilon} = (\widetilde{\epsilon}_1, \ldots, \widetilde{\epsilon}_{\widetilde{r}_0}),\ \widetilde{h}_{\mu_j} = (\widetilde{h}_1^{(j)}, \ldots, \widetilde{h}_{\widetilde{g}_j}^{(j)}),\ \widetilde{s} = (\widetilde{s}_1, \ldots, \widetilde{s}_{\widetilde{g}_\infty}), \end{aligned}$$

respectively. Notably, the integer partitions associated with the same invariants of $G - \lambda H$ and $\widetilde{G} - \lambda \widetilde{H}$, e.g. ϵ and $\widetilde{\epsilon}$, can be of different length.

Proposition 5.11 [65] *Let $G - \lambda H$ and $\widetilde{G} - \lambda \widetilde{H}$ be two $n \times (n+m)$ matrix pencils with no row minimal indices. If $\mathcal{O}(G - \lambda H)$ covers $\mathcal{O}(\widetilde{G} - \lambda \widetilde{H})$ then one of the following conditions holds:*

(1) $\operatorname{conj}(\epsilon) > \operatorname{conj}(\widetilde{\epsilon})$ *are adjacent,* $h_{\mu_i} = \widetilde{h}_{\mu_i}$ *for all eigenvalues* μ_i*, and* $s = \widetilde{s}$*.*

(2) $\sum_{i=1}^m \epsilon_i > \sum_{i=1}^m \widetilde{\epsilon}_i$, $\operatorname{conj}(\epsilon) > \operatorname{conj}(\widetilde{\epsilon})$ *are adjacent,* $\widetilde{h}_1^{(i)} = h_1^{(i)} + 1$ *for some eigenvalue* μ_i *(where* μ_i *can be a new eigenvalue), and* $s = \widetilde{s}$*.*

(3) $\sum_{i=1}^m \epsilon_i > \sum_{i=1}^m \widetilde{\epsilon}_i$, $\operatorname{conj}(\epsilon) > \operatorname{conj}(\widetilde{\epsilon})$ *are adjacent,* $h_{\mu_i} = \widetilde{h}_{\mu_i}$ *for all eigenvalues* μ_i*, and* $\widetilde{s}_1 = s_1 + 1$ *(where* s *and* $\widetilde{s}$ *can be empty partitions).*

(4) $\epsilon = \widetilde{\epsilon}$, $h_{\mu_i} > \widetilde{h}_{\mu_i}$ *for all eigenvalues* μ_i*, and* $s = \widetilde{s}$*.*

(5) $\epsilon = \widetilde{\epsilon}$, $h_{\mu_i} = \widetilde{h}_{\mu_i}$ *for all eigenvalues* μ_i*, and* $s > \widetilde{s}$*.*

From Theorem 5.9, Proposition 5.11, and the cover conditions for matrix pencils (see Theorem 5.8 and B), it is possible to derive both necessary and sufficient conditions for a cover relation between two controllability pairs (A, B). The proof is organized as follows. We modify Proposition 5.11 so that it fulfills the restrictions given by the structure of the controllability pair and then, where required, strengthen each condition so that they become not only necessary but also sufficient.

Theorem 5.12 [36] *$\mathcal{O}(A, B)$ covers $\mathcal{O}(\widetilde{A}, \widetilde{B})$* if and only if *one of the following conditions holds:*

(1) $\mathcal{R}(A, B)$ *covers* $\mathcal{R}(\widetilde{A}, \widetilde{B})$ *where* $r_0(A, B) = r_0(\widetilde{A}, \widetilde{B})$*, and* $\mathcal{J}_{\mu_i}(A, B) = \mathcal{J}_{\mu_i}(\widetilde{A}, \widetilde{B})$ *for all eigenvalues* μ_i*.*

(2) If $r_{\epsilon_1} = 1$ *and* $\epsilon_1 \geq 1$ *for* $\mathcal{R}(A, B)$*, then* $\mathcal{R}(\widetilde{A}, \widetilde{B}) = \mathcal{R}(A, B) \setminus (r_{\epsilon_1})$, $\mathcal{J}_{\mu_i}(\widetilde{A}, \widetilde{B}) = \mathcal{J}_{\mu_i}(A, B) \cup (1)$ *for some eigenvalue* μ_i *(where* $\mathcal{J}_{\mu_i}(A, B)$ *can be an empty partition), and* $\mathcal{J}_{\mu_j}(A, B) = \mathcal{J}_{\mu_j}(\widetilde{A}, \widetilde{B})$ *for all* $\mu_j \neq \mu_i$*.*

(3) $\mathcal{R}(A, B) = \mathcal{R}(\widetilde{A}, \widetilde{B})$, $\mathcal{J}_{\mu_i}(A, B)$ *covers* $\mathcal{J}_{\mu_i}(\widetilde{A}, \widetilde{B})$ *for one eigenvalue* μ_i*, and* $\mathcal{J}_{\mu_j}(A, B) = \mathcal{J}_{\mu_j}(\widetilde{A}, \widetilde{B})$ *for all* $\mu_j \neq \mu_i$*.*

Theorem 5.13 [36] *$\mathcal{B}(A, B)$ covers $\mathcal{B}(\widetilde{A}, \widetilde{B})$* if and only if *one of the following conditions holds:*

(1) $\mathcal{R}(A, B)$ *covers* $\mathcal{R}(\widetilde{A}, \widetilde{B})$ *where* $r_0(A, B) = r_0(\widetilde{A}, \widetilde{B})$*, and* $\mathcal{J}_{\mu_i}(A, B) = \mathcal{J}_{\mu_i}(\widetilde{A}, \widetilde{B})$ *for all eigenvalues* μ_i*.*

(2) If $r_{\epsilon_1} = 1$ *and* $\epsilon_1 \geq 1$ *for* $\mathcal{R}(A, B)$*, then* $\mathcal{R}(\widetilde{A}, \widetilde{B}) = \mathcal{R}(A, B) \setminus (r_{\epsilon_1})$, $\mathcal{J}_{\mu_i}(\widetilde{A}, \widetilde{B}) = (1)$ *for a new eigenvalue* μ_i*, and* $\mathcal{J}_{\mu_j}(A, B) = \mathcal{J}_{\mu_j}(\widetilde{A}, \widetilde{B})$ *for all* $\mu_j \neq \mu_i$*.*

(3) $\mathcal{R}(A, B) = \mathcal{R}(\widetilde{A}, \widetilde{B})$, $\mathcal{J}_{\mu_i}(A, B)$ *covers* $\mathcal{J}_{\mu_i}(\widetilde{A}, \widetilde{B})$ *for one eigenvalue* μ_i*, and* $\mathcal{J}_{\mu_j}(A, B) = \mathcal{J}_{\mu_j}(\widetilde{A}, \widetilde{B})$ *for all* $\mu_j \neq \mu_i$*.*

(4) $\mathcal{R}(A, B) = \mathcal{R}(\widetilde{A}, \widetilde{B})$, $\mathcal{J}_{\mu_i}(\widetilde{A}, \widetilde{B}) = \mathcal{J}_{\mu_i}(A, B) \cup \mathcal{J}_{\mu_j}(A, B)$ *for one pair of eigenvalues* μ_i *and* μ_j, $\mu_i \neq \mu_j$*, and* $\mathcal{J}_{\mu_k}(A, B) = \mathcal{J}_{\mu_k}(\widetilde{A}, \widetilde{B})$ *for all* $\mu_k \neq \mu_i, \mu_j$*.*

Notably, Theorem 5.13 has four rules, i.e., one extra compared to Theorem 5.12. The additional rule (4) follows from that eigenvalues can coalesce in the bundle case.

From the dual relation between the controllability pair (A, B) and the observability pair (A, C), it follows that replacing partition $\mathcal{R}$ with $\mathcal{L}$ in Theorems 5.12 and 5.13 gives the cover conditions for the observability pair (A, C). We remark that the theorems are only valid for independent (decoupled) matrix pairs (A, B) and (A, C), respectively. This means that Theorems 5.12 and 5.13 cannot be applied straightforwardly to the related matrix triple (A, B, C) or matrix quadruple (A, B, C, D). The covering relations for orbits and bundles of the controllability pair in terms of coin rules are given in Corollaries 5.14 and 5.15. The reformulations are done using the definition of integer partitions and Theorem 5.1. In Tables 4 and 5 of B, these and the remaining covering relations for matrix pairs are summarized. A larger example illustrating the usage of Corollary 5.15 is presented in Section 5.4.

Corollary 5.14 *Given the structure integer partitions $\mathcal{R}$ and $\mathcal{J}_{\mu_i}$ of (A, B), one of the following* if-and-only-if rules *finds $(\widetilde{A}, \widetilde{B})$ such that $\mathcal{O}(A, B)$ covers $\mathcal{O}(\widetilde{A}, \widetilde{B})$:*

(1) Minimum rightward *coin move in $\mathcal{R}$.*

(2) If the rightmost column in $\mathcal{R}$ is one single coin, move that coin to a new rightmost column of some $\mathcal{J}_{\mu_i}$ (which may be empty initially).

(3) Minimum leftward *coin move in any $\mathcal{J}_{\mu_i}$.*

Rules 1 and 2 are not allowed to do coin moves that affect r_0.

Corollary 5.15 *Given the structure integer partitions $\mathcal{R}$ and $\mathcal{J}_{\mu_i}$ of (A, B), one of the following* if-and-only-if rules *finds $(\widetilde{A}, \widetilde{B})$ such that $\mathcal{B}(A, B)$ covers $\mathcal{B}(\widetilde{A}, \widetilde{B})$:*

(1) Minimum rightward *coin move in $\mathcal{R}$.*

(2) If the rightmost column in $\mathcal{R}$ is one single coin, move that coin to the first column of $\mathcal{J}_{\mu_i}$ for a new eigenvalue μ_i.

(3) Minimum leftward *coin move in any $\mathcal{J}_{\mu_i}$.*

(4) Let any pair of eigenvalues coalesce, i.e., take the union of their sets of coins.

The major difference between the rules for matrix pencils and matrix pairs, is that rule (4) (both for orbits and bundles) in Theorem 5.8 does not apply to matrix pairs, since there is only one type of singular blocks (L or L^T) in each matrix pair type. Moreover, in rules (1) and (2) of Corollaries 5.14 and 5.15, the pair (A, B) applies to the $\mathcal{R}$ partition only.

5.4 Illustrating the stratification of a state-space system

We illustrate the concept of stratification by considering a general state-space system of the same size as the one used in Example 4, with two states, three inputs and one output ($n = 2$, $m = 3$ and $p = 1$):

$$\mathbf{S}(\lambda) = \begin{bmatrix} A & B \\ C & D \end{bmatrix} - \lambda \begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix}, \tag{5.45}$$

where $A \in \mathbb{C}^{2\times 2}$, $B \in \mathbb{C}^{2\times 3}$, $C \in \mathbb{C}^{1\times 2}$ and $D \in \mathbb{C}^{1\times 3}$. We assume that the system does not have well-defined poles and zeros (eigenvalues), so we have to work with the bundle

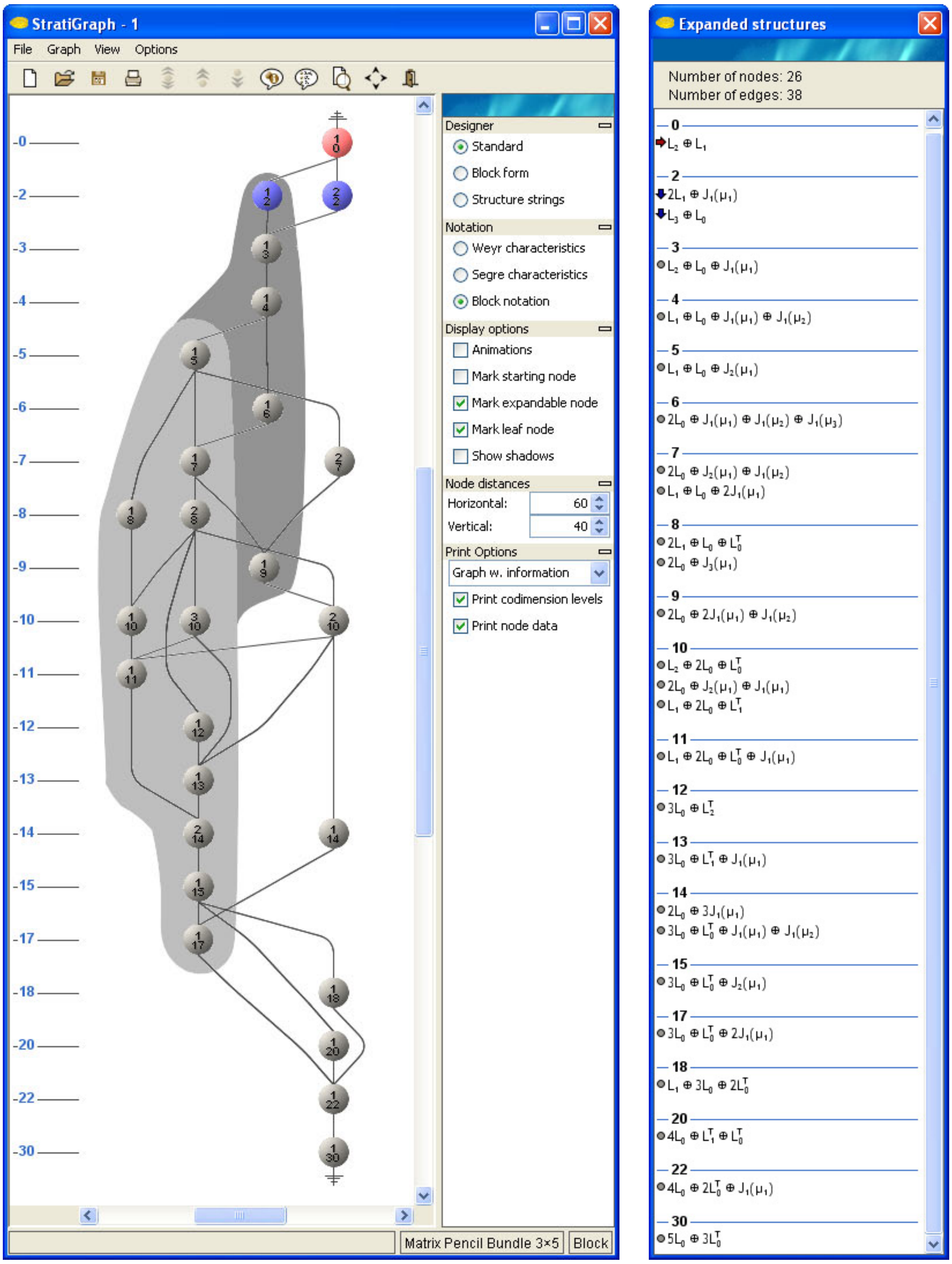


FIGURE 9: The window on the left shows the complete stratification of the bundles of 3×5 general matrix pencils. The light grey area marks the possible canonical structures for a matrix triple and together with the dark grey area for a matrix quadruple. The window on the right shows the corresponding canonical structures associated with the nodes in the graph.

stratification. To compute and visualize the closure hierarchy graphs for the different system pencils in the examples we use the software tool StratiGraph v 3.0 [70, 74].

The bundle stratification of general matrix pencils $G - \lambda H$ of size 3×5 can be represented by a graph with 26 different (canonical) structures, as in Figure 9. The graph spans from the most generic case, $L_2 \oplus L_1$ with codimension 0, to the most degenerate case, $5L_0 \oplus 3L_0^T$ with codimension 30. This stratification does not, however, consider the special structure of the system pencil $\mathbf{S}(\lambda)$ and therefore generates canonical structures in the closure hierarchy that are in fact not possible for the system (5.45). Moreover, for matrix pencils the stratification procedure makes no distinction between finite and infinite elementary divisors (eigenvalues).

Instead we associate the state-space system with a matrix quadruple (A, B, C, D). Even though we do not have the covering relations for matrix quadruples we can generate all 18 possible structures for $\mathbf{S}(\lambda)$. They are listed in Table 1 with their corresponding bundle codimensions in the leftmost column, where the most and least generic structures are derived in Example 9. In Figure 9, the corresponding structures are highlighted by the union of the dark and light grey areas. Let $c\!:\!k$ denote node $\genfrac{}{}{0pt}{}{k}{c}$ in Figure 9, where c is the codimension of the corresponding bundle and k is an order number that identifies individual nodes with the same codimension. As examples, we show why the orbits with KCF $L_3 \oplus L_0$ (2:2) and $L_1 \oplus L_0 \oplus 2J_1(\mu_1)$ (7:2) are not possible for system (5.45). The reasoning follows directly from the characteristics of the GBCF (2.19). The node with KCF $L_3 \oplus L_0$ has only two L blocks and therefore $m = 2$ (the system has two inputs), which contradicts with the number of inputs of $\mathbf{S}(\lambda)$. The node with KCF $L_1 \oplus L_0 \oplus 2J_1(\mu_1)$ can either have a finite or infinite eigenvalue μ_1. If μ_1 is finite then $n = 3$ and $m = 2$, and if μ_1 is infinite then $n = 1$ and $m = 4$. In both cases, the dimensions of the state-space matrices contradicts with $\mathbf{S}(\lambda)$.

Moreover, the stratification procedure for matrix quadruples identifies infinite and finite elementary divisors and treat them separately. For example, the structure $2L_0 \oplus J_2(\mu_1) \oplus J_1(\mu_2)$ (node 7:2) for the general matrix pencil splits into two different structures in the matrix quadruple case:

- $2L_0 \oplus J_2(\mu_1) \oplus N_1$ corresponding to a system with one finite elementary divisor of order two (a finite zero at μ_1 of order two), one infinite elementary divisor of order one, and the column minimal indices $\epsilon_1 = 0$ and $\epsilon_2 = 0$.

- $2L_0 \oplus J_1(\mu_1) \oplus N_2$ corresponding to a system with one finite elementary divisor of order one (a finite zero at μ_1 of order one), one infinite elementary divisor of order two (an infinite zero of order one), and the column minimal indices $\epsilon_1 = 0$ and $\epsilon_2 = 0$.

By only considering the subsystem of $\mathbf{S}(\lambda)$ corresponding to the matrix triple (A, B, C), we get the subset of structures in Table 1 with no infinite elementary divisors of order one, i.e., no N_1 blocks (as we concluded in the end of Section 2.6). In Table 1, the codimensions are presented for the bundles of the matrix triples when they can appear in the closure hierarchy, and in Figure 9 the possible canonical structures for triples are highlighted by the light grey area. The most and least generic triples are derived in Example 9. The bundle of the matrix triple associated with the second order state-space system (2.20) considered in Example 4, with KCF $2L_0 \oplus J_1(\mu_1) \oplus N_2$, is represented by the node 7:2 in Figure 9 and as we can see in Table 1 it has codimension 2.

Now, consider the independent subsystems of (5.45) associated with the controllability pair (A, B) and the observability pair (A, C), where the controllability system pencil

$$\mathbf{S}_{\mathrm{C}}(\lambda) = \begin{bmatrix} A & B \end{bmatrix} - \lambda \begin{bmatrix} I_2 & 0 \end{bmatrix},$$

TABLE 1: All possible canonical structures for a state-space system with two states, three inputs and one output. The bundle codimensions for the associated matrix quadruple and triple are listed in the first two columns, and the label for the corresponding node in Figure 9 in the last column.

cod($\mathcal{B}(*)$)			Node label
(A,B,C,D)	(A,B,C)	Canonical structure (KCF)	in Fig. 9
0	–	$2L_1 \oplus N_1$	2:1
1	–	$L_2 \oplus L_0 \oplus N_1$	3:1
2	–	$L_1 \oplus L_0 \oplus J_1(\mu_1) \oplus N_1$	4:1
3	0	$L_1 \oplus L_0 \oplus N_2$	5:1
4	–	$2L_0 \oplus J_1(\mu_1) \oplus J_1(\mu_2) \oplus N_1$	6:1
5	–	$2L_0 \oplus J_2(\mu_1) \oplus N_1$	7:2
5	2	$2L_0 \oplus J_1(\mu_1) \oplus N_2$	7:2
5	2	$2L_1 \oplus L_0 \oplus L_0^T$	8:1
6	3	$2L_0 \oplus N_3$	8:2
7	–	$2L_0 \oplus 2J_1(\mu_1) \oplus N_1$	9:1
7	4	$L_2 \oplus 2L_0 \oplus L_0^T$	10:1
7	4	$L_1 \oplus 2L_0 \oplus L_1^T$	10:3
8	5	$L_1 \oplus 2L_0 \oplus L_0^T \oplus J_1(\mu_1)$	11:1
9	6	$3L_0 \oplus L_2^T$	12:1
10	7	$3L_0 \oplus L_1^T \oplus J_1(\mu_1)$	13:1
11	8	$3L_0 \oplus L_0^T \oplus J_1(\mu_1) \oplus J_1(\mu_2)$	14:1
12	9	$3L_0 \oplus L_0^T \oplus J_2(\mu_1)$	15:1
14	11	$3L_0 \oplus L_0^T \oplus 2J_1(\mu_1)$	17:1

is of size 2×5 and the observability system pencil

$$\mathbf{S}_{\mathrm{O}}(\lambda) = \begin{bmatrix} A \\ C \end{bmatrix} - \lambda \begin{bmatrix} I_2 \\ 0 \end{bmatrix},$$

is 3×2. The stratification of bundles of the matrix pairs (A, B) and (A, C) are illustrated by graphs (a) and (c) in Figure 10, and in graphs (b) and (d) we show the stratification for orbits. These closure hierarchy graphs are computed by using the stratification rules in Tables 4 and 5 of B.

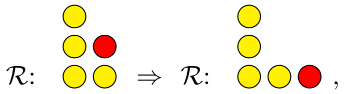
We now show step by step the procedure to get the complete bundle stratification of the controllability pair (A, B), as shown in graph (a) of Figure 10. We can, e.g., start by determining the most generic case which corresponds to a controllable system (or the most degenerate case if we work the opposite way). In Example 11, we have shown how the most and least generic cases can be determined if the system matrices A and B have a fixed structure. However, we are now interested in the stratification of a general controllability pair. As we have shown in Example 10 the most generic structure has the KCF $2L_1 \oplus L_0$ with the corresponding BCF:

$$\begin{bmatrix} A_B & B_B \end{bmatrix} - \lambda \begin{bmatrix} I_2 & 0 \end{bmatrix} = \left[\begin{array}{cc:ccc} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right] - \lambda \left[\begin{array}{cc:ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right].$$

By considering the controllability pair in BCF we can directly see that (A, B) is controllable; $\operatorname{rank}\left(\begin{bmatrix} A_B & B_B \end{bmatrix} - \lambda \begin{bmatrix} I_2 & 0 \end{bmatrix}\right) = 2$ for all $\lambda \in \mathbb{C}$. We can also see that it has three $(m = 3)$ L blocks of which one $(3 - \operatorname{rank}(B_B) = 1)$ is an L_0 block.

This is the topmost node in the graph and by using the appropriate formula in A we can also determine that it has codimension 0. The next step is to decide which bundle(s) that is (are) covered by $\mathcal{B}(2L_1 \oplus L_0)$. This is done by using the rules (1)–(4) in the bottom left part of Table 4 of B.

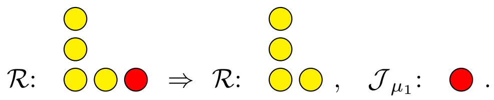
The first rule tells us to do a minimum rightward coin move on $\mathcal{R}$. The only possible choice is to move the topmost coin from r_1 to r_2:



which gives the structure $L_2 \oplus 2L_0$. The second rule is not applicable because the rightmost coin in $\mathcal{R}$ is not a single coin, as well as the third and fourth rules because we have no Jordan blocks. So the only bundle covered by $\mathcal{B}(2L_1 \oplus L_0)$ is the bundle with KCF $L_2 \oplus 2L_0$ and codimension 2. The associated matrix pair is controllable, which also can be seen from its BCF:

$$\begin{bmatrix} A_B & B_B \end{bmatrix} - \lambda \begin{bmatrix} I_2 & 0 \end{bmatrix} = \left[\begin{array}{cc|ccc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right] - \lambda \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right].$$

We continue by repeating the procedure for $L_2 \oplus 2L_0$. The first rule is not applicable because the only possible minimum rightward coin move affects r_0 which is not allowed. As before, since there are no Jordan blocks in $L_2 \oplus 2L_0$ the last two rules can also not be applied. However, with rule (2) we can remove the last coin from $\mathcal{R}$ and create a new set $\mathcal{J}_{\mu_1} = (1)$:



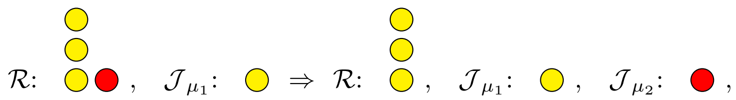
The new structure has the KCF $L_1 \oplus 2L_0 \oplus J_1(\mu_1)$ which corresponds to a system with one uncontrollable mode at μ_1. Looking at the corresponding BCF:

$$\begin{bmatrix} A_B & B_B \end{bmatrix} - \lambda \begin{bmatrix} I_2 & 0 \end{bmatrix} = \left[\begin{array}{cc|ccc} 0 & 0 & 1 & 0 & 0 \\ 0 & \mu_1 & 0 & 0 & 0 \end{array}\right] - \lambda \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right],$$

we can see that $\operatorname{rank}\left(\begin{bmatrix} A_B & B_B \end{bmatrix} - \mu_1 \begin{bmatrix} I_2 & 0 \end{bmatrix}\right) = 1$ and the diagonal entry μ_1 corresponds to the uncontrollable mode.

In the following, we only mention the rules that are applicable on each canonical structure. There is actually only one rule in each step that is allowed. Notably, the graph never splits into two or more branches in this example, as will happen in the orbit case.

Once again we can apply rule (2); now to the structure $L_1 \oplus 2L_0 \oplus J_1(\mu_1)$:



which gives the KCF $3L_0 \oplus J_1(\mu_1) \oplus J_1(\mu_2)$ corresponding to a system with two uncontrollable modes at μ_1 and μ_2. The corresponding BCF is

$$\begin{bmatrix} A_B & B_B \end{bmatrix} - \lambda \begin{bmatrix} I_2 & 0 \end{bmatrix} = \left[\begin{array}{cc|ccc} \mu_1 & 0 & 0 & 0 & 0 \\ 0 & \mu_2 & 0 & 0 & 0 \end{array}\right] - \lambda \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right],$$

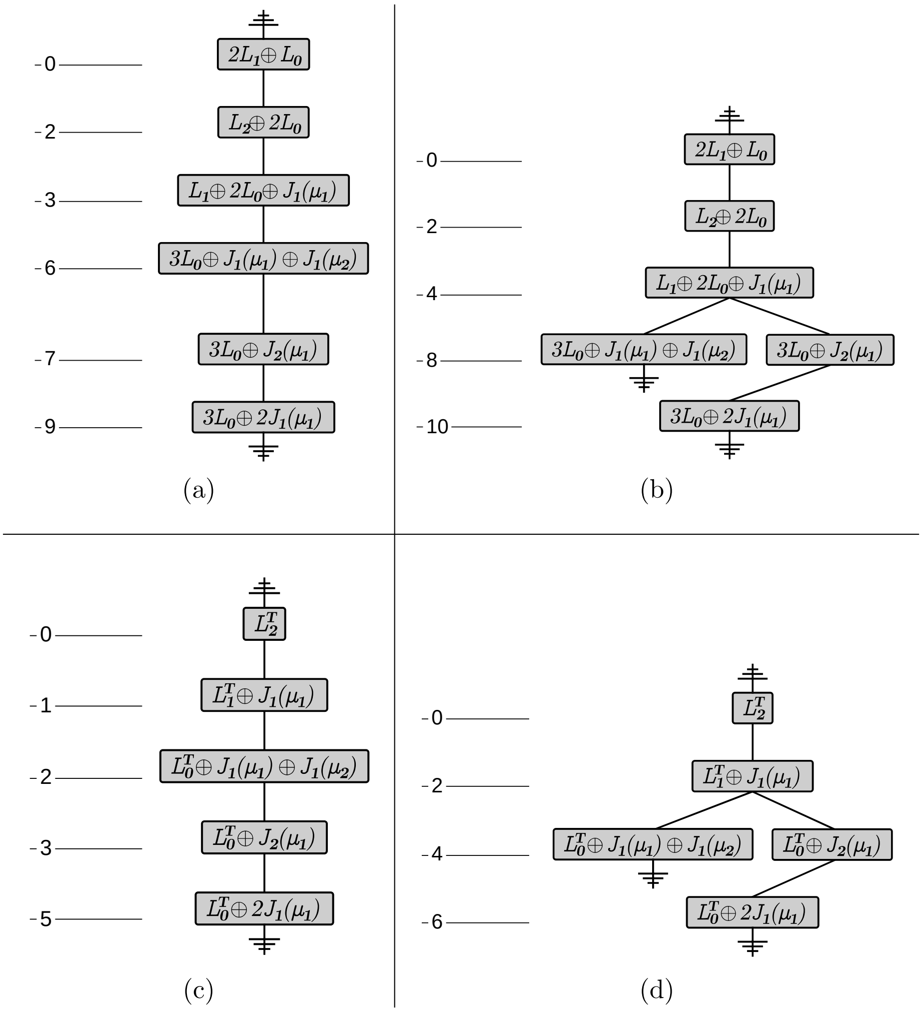
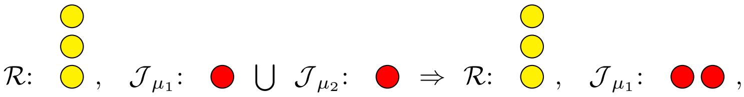


FIGURE 10: The stratification of: (**a**) (A, B)-bundles $(n = 2,\ m = 3)$, (**b**) (A, B)-orbits $(n = 2,\ m = 3)$, (**c**) (A, C)-bundles $(n = 2,\ p = 1)$, and (**d**) (A, C)-orbits $(n = 2,\ p = 1)$.

where the two uncontrollable modes are the diagonal elements of A_B.

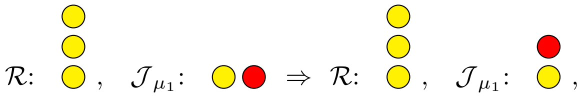
The fourth rule can be applied to $3L_0 \oplus J_1(\mu_1) \oplus J_1(\mu_2)$:



which gives the KCF $3L_0 \oplus J_2(\mu_1)$ with one uncontrollable mode of multiplicity two and with the corresponding BCF:

$$\begin{bmatrix} A_B & B_B \end{bmatrix} - \lambda \begin{bmatrix} I_2 & 0 \end{bmatrix} = \left[\begin{array}{cc:ccc} \mu_1 & 1 & 0 & 0 & 0 \\ 0 & \mu_1 & 0 & 0 & 0 \end{array}\right] - \lambda \left[\begin{array}{cc:ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right].$$

Finally, we can apply rule (3) to $3L_0 \oplus J_2(\mu_1)$:



which gives the KCF $3L_0 \oplus 2J_1(\mu_1)$ with two uncontrollable multiple modes and with the corresponding BCF:

$$\begin{bmatrix} A_B & B_B \end{bmatrix} - \lambda \begin{bmatrix} I_2 & 0 \end{bmatrix} = \left[\begin{array}{cc:ccc} \mu_1 & 0 & 0 & 0 & 0 \\ 0 & \mu_1 & 0 & 0 & 0 \end{array}\right] - \lambda \left[\begin{array}{cc:ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right].$$

By examining the closure hierarchy we get qualitative information about systems under small perturbations. From the stratification in graph (a) of Figure 10 we can see that the two most generic cases are completely controllable (they have no uncontrollable modes), the structure with codimension 3 has one uncontrollable mode, and so forth. By also presenting the upper and lower bounds for the distance from a given controllable structure to the nearest uncontrollable, we also have a quantitative measure on how sensitive the controllable system is for small changes. This quantitative information is available in StratiGraph through the *Matrix Canonical Structure Toolbox* for MATLAB [71], in which StratiGraph [70, 74] has been incorporated (see also [72]).

Finally, let us return to Example 4 where the controllable and observable second order state-space system (2.20) has the controllability pair $L_2 \oplus 2L_0$ and the observability pair L_2^T. As we can see in graph (a) of Figure 10 the controllability pair is not the most generic case, so it can by a small perturbation become the structure $2L_1 \oplus L_0$, which is also controllable. However, the more degenerate case $L_1 \oplus 2L_0 \oplus J_1(\mu_1)$ is uncontrollable with one uncontrollable mode. By computing a lower bound to this system, which is the "nearest" (in complex arithmetic) uncontrollable system in the closure hierarchy, we get a measure on the sensitivity of controllability for the state-space system (2.20).

The observability pair on the other hand is already the most generic case. But as we showed in Example 4 the structure could be very close to the more degenerate case $L_1^T \oplus J_1(\mu_1)$, with the unobservable mode μ_1, if the state-space system (2.20) has γ close to zero. And in graph (c) we can see that all less generic structures are also unobservable.

6 Some concluding remarks

In this paper, we have given an introduction to stratification of orbits and bundles with applications in systems and control. The necessary background theory has been presented, both from a mathematical as well as from an applications point of view using a unifying terminology and notation. The background theory and the stratification theory have throughout the paper been illustrated by examples.

The close relation between the Kronecker canonical form and the generalized Brunovsky canonical form is well known. In Section 2.7, the explicit expressions for the permutation matrices which transform a matrix pencil in KCF to GBCF are derived. Algorithms to determine these two permutation matrices are also presented in the same section.

In Section 5.3, the closure conditions for orbits and bundles of matrix pairs are derived and the cover conditions are presented. In line with previous work on matrices and matrix pencils, we have presented the stratification rules for matrix pairs, both the controllability pair (A, B) and the observability pair (A, C).

The natural continuation of Section 5.3 is to derive both the closure and cover conditions for orbits and bundles of matrix quadruples and matrix triples. Other systems which are of interest are state-space system with fixed structure and generalized state-space systems and subsystems there of. All these systems are of great practical interest and arise in several applications.

Acknowledgement

The author is grateful to Bo Kågström for all constructive comments regarding the structure of the paper and its content as well as suggestions for the literature study on different topics, and Erik Elmroth who has been an invaluable help when formulating the proofs of Algorithms 1 and 2, and the proof of Theorems 5.9. The author would also like to thank Daniel Kressner and Pedher Johansson who have taken their time to read and give their comments on an earlier version of the paper.

References

[1] P. Antsaklis and A. Michel. *Linear Systems*. Birkhäuser, Boston, 2nd edition, 2006. ISBN 978-0-8176-4434-5.

[2] V. I. Arnold. On matrices depending on parameters. *Russian Math. Surveys*, 26:29–43, 1971.

[3] V. I. Arnold. *Geometrical methods in the theory of ordinary differential equations*. Springer-Verlag, New York, 2nd printing, 2nd edition, 1997. ISBN 0-387-96649-8.

[4] T. Beelen and P. Van Dooren. An improved algorithm for the computation of Kronecker's canonical form of a singular pencil. *Linear Algebra Appl.*, 105:9–65, 1988.

[5] T. Beelen and P. Van Dooren. Computational aspects of the Jordan canonical form. In M. Cox and S. Hammarling, editors, *Reliable numerical computations*, pages 57–72. Clarendon Press, Oxford, 1990.

[6] J. M. Berg and H. G. Kwatny. A canonical parameterization of the Kronecker form of a matrix pencil. *Automatica*, 31(5):669–680, 1995.

[7] J. M. Berg and H. G. Kwatny. Unfolding the zero structure of a linear control system. *Linear Algebra Appl.*, 258:19–39, 1997.

[8] D. Boley. *Computing the controllability/observability of a linear time-invariant dynamic system: A numerical approach*. PhD thesis, Computer Science Department, Stanford University, Stanford, CA, June 1981.

[9] D. Boley. Estimating the sensitivity of the algebraic structure of pencils with simple eigenvalue estimates. *SIAM J. Matrix Anal. Appl.*, 11(4):632–643, 1990.

[10] D. Boley. The algebraic structure of pencils and block Toeplitz matrices. *Linear Algebra Appl.*, 279:255–279, 1998.

[11] K. Bongartz. On degenerations and extensions of finite dimensional modules. *Adv. Math.*, 121:245–287, 1996.

[12] P. Brunovsky. A classification of linear controllable systems. *Kybernetika*, 3(6):173–188, 1970.

[13] A. Bunse-Gerstner, R. Byers, V. Mehrmann, and N. K. Nichols. Feedback design for regularizing descriptor systems. *Linear Algebra Appl.*, 299:119–151, 1999.

[14] J. V. Burke, A. S. Lewis, and M. L. Overton. Pseudospectral components and the distance to uncontrollability. *SIAM J. Matrix Anal. Appl.*, 26(2):350–361, 2004.

[15] J. Clotet, M. I. García-Planas, and M. D. Magret. Estimating distances from quadruples satisfying stability properties to quadruples not satisfying them. *Linear Algebra Appl.*, 332–334:541–567, 2001.

[16] D. H. Collingwood and W. M. McGovern. *Nilpotent orbits in semisimple Lie algebras*. Van Nostrand Reinhold, New York, 1993.

[17] L. Dai. *Singular Control Systems*, volume 118 of *Lecture Notes in Control and Information Science*. Springer-Verlag, Berlin, 1989.

[18] B. Datta. *Numerical methods for linear control systems.* Elsevier, New York, 2004. ISBN 0-12-203590-9.

[19] I. De Hoyos. Points of continuity of the Kronecker canonical form. *SIAM J. Matrix Anal. Appl.*, 11:278–300, 1990.

[20] F. De Terán and F. Dopico. Low rank perturbation of Kronecker structures without full rank. *SIAM J. Matrix Anal. Appl.*, 29(2):496–529, 2007.

[21] F. De Terán and F. Dopico. A note on generic Kronecker orbits of matrix pencils with fixed rank. To appear in SIAM Journal on Matrix Analysis and Applications, 2008.

[22] C. DeConcini, D. Eisenbud, and C. Procesi. Young diagrams and determinantal varieties. *Invent. Math.*, 56:129–165, 1980.

[23] J. Demmel and A. Edelman. The dimension of matrices (matrix pencils) with given Jordan (Kronecker) canonical forms. *Linear Algebra Appl.*, 230:61–87, 1995.

[24] J. Demmel and B. Kågström. Computing stable eigendecompositions of matrix pencils. *Linear Algebra Appl.*, 88/89:139–186, 1987.

[25] J. Demmel and B. Kågström. Accurate solutions of ill-posed problems in control theory. *SIAM J. Matrix Anal. Appl.*, 9(1):126–145, January 1988.

[26] J. Demmel and B. Kågström. The generalized Schur decomposition of an arbitrary pencil $A - \lambda B$: Robust software with error bounds and applications. Part I: Theory and algorithms. *ACM Trans. Math. Software*, 19(2):160–174, June 1993.

[27] J. Demmel and B. Kågström. The generalized Schur decomposition of an arbitrary pencil $A - \lambda B$: Robust software with error bounds and applications. Part II: Software and applications. *ACM Trans. Math. Software*, 19(2):175–201, June 1993.

[28] H. Den Boer and Ph. A. Thijsse. Semi-stability of sums of partial multiplicities under additive perturbation. *Integral Equations Operator Theory*, 3:23–42, 1980.

[29] Y. Ebihara. Computing the distance to uncontrollability via LMIs: Lower bound computation with exactness verification. *Systems & Control Lett.*, 57:763–771, 2008.

[30] A. Edelman, E. Elmroth, and B. Kågström. A geometric approach to perturbation theory of matrices and matrix pencils. Part I: Versal deformations. *SIAM J. Matrix Anal. Appl.*, 18:653–692, 1997.

[31] A. Edelman, E. Elmroth, and B. Kågström. A geometric approach to perturbation theory of matrices and matrix pencils. Part II: A stratification-enhanced staircase algorithm. *SIAM J. Matrix Anal. Appl.*, 20:667–669, 1999.

[32] R. Eising. Between controllable and uncontrollable. *Systems Control Lett.*, 4:263–264, 1984.

[33] E. Elmroth, P. Johansson, S. Johansson, and B. Kågström. Orbit and bundle stratification of controllability and observability matrix pairs in StratiGraph. In B. De Moor et.al., editor, *Proc. Sixteenth International Symposium on Mathematical Theory of Networks and Systems (MTNS2004)*, Leuven, Belgium, July 2004. On CD.

[34] E. Elmroth, P. Johansson, and B. Kågström. Computation and presentation of graph displaying closure hierarchies of Jordan and Kronecker structures. *Numer. Linear Algebra Appl.*, 8(6–7):381–399, 2001.

[35] E. Elmroth, P. Johansson, and B. Kågström. Bounds for the distance between nearby Jordan and Kronecker structures in a closure hierarchy. *Journal of Mathematical Sciences*, 114(6):1765–1779, 2003.

[36] E. Elmroth, S. Johansson, and B. Kågström. Stratification of controllability and observability pairs — Theory and use in applications. Technical Report UMINF 08-03, Department of Computing Science, Umeå University, Sweden, 2008. Accepted for publication in *SIAM J. Matrix Anal. Appl.*

[37] E. Elmroth and B. Kågström. The set of 2-by-3 matrix pencils — Kronecker structures and their transitions under perturbations. *SIAM J. Matrix Anal. Appl.*, 17(1):1–34, 1996.

[38] J. Ferrer, M. I. García, and F. Puerta. Regularity of the Brunovsky-Kronecker stratification. *SIAM J. Matrix Anal. Appl.*, 21(3):724–742, 2000.

[39] J. Ferrer, Mª I. García, and F. Puerta. Brunowsky local form of a holomorphic family of pairs of matrices. *Linear Algebra Appl.*, 253:175–198, 1997.

[40] J. Ferrer and Mª I. García-Planas. Structural stability of quadruples of matrices. *Linear Algebra Appl.*, 241–243:279–290, 1996.

[41] J. Ferrer and F. Puerta. Versal deformations of invariant subspaces. *Linear Algebra Appl.*, 332–334:569–582, 2001.

[42] G. D. Jr. Forney. Minimal bases of rational vector spaces with applications to multivariable linear systems. *SIAM J. Control Optim.*, 13(3):493–520, 1975.

[43] F. Gantmacher. *The theory of matrices, Vol. I and II (transl.).* Chelsea, New York, 1959.

[44] M. I. García-Planas and M. D. Magret. Stratification of linear systems. Bifurcation diagrams for families of linear systems. *Linear Algebra Appl.*, 297:23–56, 1999.

[45] M. I. García-Planas and M. D. Magret. A criterion for structural stability of quadruples of matrices defining generalized dynamical systems. *Systems Control Lett.*, 41(3):157–166, 2000.

[46] M. I. García-Planas and A. A. Mailybaev. Reduction to versal deformations of matrix pencils and matrix pairs with application to control theory. *SIAM J. Matrix Anal. Appl.*, 24(4):943–962, 2003.

[47] M. I. García-Planas and V. V. Sergeichuk. Simplest miniversal deformations of matrices, matrix pencils, and contragredient matrix pencils. *Linear Algebra Appl.*, 302–303:45–61, 1999.

[48] M. I. García-Planas and V. V. Sergeichuk. Generic families of matrix pencils and their bifurcation diagrams. *Linear Algebra Appl.*, 332–334:165–179, 2001.

[49] Mª I. García-Planas. Kronecker stratification of the space of quadruples of matrices. *SIAM J. Matrix Anal. Appl.*, 19(4):872–885, 1998.

[50] Mª I. García-Planas and M. D. Magret. Miniversal deformations of linear systems under the full group action. *Systems Control Lett.*, 35(5):279–286, 1998.

[51] Mª I. García-Planas and Mª D. Magret. Deformation and stability of triples of matrices. *Linear Algebra Appl.*, 254:159–192, 1997.

[52] H. Glüsing-Lüerssen and D. Hinrichsen. A Jordan control canonical form for singular systems. *Int. J. Control*, 48(5):1769–1785, 1988.

[53] I. Gohberg, P. Lancaster, and L. Rodman. *Invariant subspaces of matrices with applications*. Wiley, 1986. ISBN 0-471-84260-5.

[54] G. Golub and J. H. Wilkinson. Ill-conditioned eigensystems and the computation of the Jordan canonical form. *SIAM Review*, 18(4):578–619, 1976.

[55] G. H. Golub and C. F. Van Loan. *Matrix Computations*. John Hopkins University Press, Baltimore, MD, 3rd edition, 1996.

[56] G. Goodwin, S. Graebe, and M. Salgado. *Control System Design*. Prentice Hall, 2001. ISBN 0-13-958653-9.

[57] J-M. Gracia and I. De Hoyos. Puntos de continuidad de formas canónicas de matrices. In *the Homage Book of Prof. Luis de Albuquerque of Coimbra*. Coimbra, 1987.

[58] J-M. Gracia and I. De Hoyos. Nearest pair with more nonconstant invariant factors and pseudospectrum. *Linear Algebra Appl.*, 298:143–158, 1999.

[59] J-M. Gracia, I. De Hoyos, and I. Zaballa. Perturbation of linear control systems. *Linear Algebra Appl.*, 121:353–383, 1989.

[60] M. Gu. Finding well-conditioned similarities to block-diagonalize non-symmetric matrices is NP-hard. *J. Complexity*, 11:377–391, 1995.

[61] M. Gu. New methods for estimating the distance to uncontrollability. *SIAM J. Matrix Anal. Appl.*, 21(3):989–1003, 2000.

[62] M. Gu, E. Mengi, M. L. Overton, J. Xia, and J. Zhu. Fast methods for estimating the distance to uncontrollability. *SIAM J. Matrix Anal. Appl.*, 28(2):477–502, 2006.

[63] U. Helmke and M. A. Shayman. A canonical form for controllable singular systems. *Systems & Control Lett.*, 12:111–122, 1989.

[64] D. Hinrichsen and J. O'Halloran. A complete characterization of orbit closures of controllable singular systems under restricted system equivalence. *SIAM J. Control Optim.*, 28(3):602–623, 1990.

[65] D. Hinrichsen and J. O'Halloran. Orbit closures of singular matrix pencils. *J. of Pure and Appl. Alg.*, 81:117–137, 1992.

[66] D. Hinrichsen and J. O'Halloran. A pencil approach to high gain feedback and generalized state space systems. *Kybernetika*, 31:109–139, 1995.

[67] D. Hinrichsen and J. O'Halloran. Limits of generalized state space systems under proportional and derivative feedback. *Mathematics of Control, Signals and Systems*, 10:97–124, 1997.

[68] J. Humphreys. *Linear Algebraic Groups.* Graduate Texts in Mathematics. Springer-Verlag, New York, 1975.

[69] S. Iwata and R. Shimizu. Combinatorial analysis of singular matrix pencils. *SIAM J. Matrix Anal. Appl.*, 29(1):245–259, 2007.

[70] P. Johansson. StratiGraph User's Guide. Technical Report UMINF 03.21, Department of Computing Science, Umeå University, Sweden, 2003.

[71] P. Johansson. Matrix Canonical Structure Toolbox. Technical Report UMINF 06.15, Department of Computing Science, Umeå University, Sweden, 2006.

[72] P. Johansson. *Software tools for matrix canonical computations and web-based software library environments.* PhD thesis, Department of Computing Science, Umeå University, Sweden, November 2006.

[73] P. Johansson. StratiGraph Developer's Guide. Technical Report UMINF 06.14, Department of Computing Science, Umeå University, Sweden, 2006.

[74] P. Johansson. StratiGraph homepage. Department of Computing Science, Umeå University, Sweden, January 2009.
http://www.cs.umu.se/research/nla/singular_pairs/stratigraph/.

[75] S. Johansson. Canonical forms and stratification of orbits and bundles of system pencils. Technical Report UMINF 05.16, Department of Computing Science, Umeå University, Sweden, 2005.

[76] S. Johansson. *Stratification of matrix pencils in systems and control: Theory and algorithms.* Licentiate Thesis, Report UMINF 05.17, Department of Computing Science, Umeå University, Sweden, 2005.

[77] B. Kågström. RGSVD — An algorithm for computing the Kronecker canonical form and reducing subspaces of singular $A - \lambda B$ pencils. *SIAM J. Sci. Statist. Comput.*, 7(1):185–211, 1986.

[78] B. Kågström. Singular matrix pencils. In Z. Bai, J. Demmel, A. Dongarra, J. Ruhe, and H. van der Vorst, editors, *Templates for the solution of algebraic eigenvalue problems: A practical guide*, pages 260–277. SIAM, Philadelphia, 2000.

[79] B. Kågström and A. Ruhe. ALGORITHM 560: An algorithm for the numerical computation of the Jordan normal form of a complex matrix [F2]. *ACM Trans. Math. Softw.*, 6(3):437–443, 1980.

[80] B. Kågström and A. Ruhe. An algorithm for the numerical computation of the Jordan normal form of a complex matrix. *ACM Trans. Math. Software*, 6(3):389–419, 1980.

[81] R. E. Kalman. Kronecker invariants and feedback. In Weiss, editor, *Ordinary Differential Equations*, Washington, 1971. Proc. Conference on ordinary differential equations.

[82] L. Kronecker. Algebraische reduktion der schaaren bilinearer formen. *Sitzungsber. Akad. d. Wiss. Berlin*, pages 763–776, 1890.

[83] V. N. Kublanovskaya. On a method of solving the complete eigenvalue problem for a degenerate matrix (in russian). *Zh. Vychisl. Mat. Fiz.*, 6:611–620, 1966. (USSR Comput. Math. Phys., 6(4):1–16, 1968).

[84] V. N. Kublanovskaya. An approach to solving the spectral problem of $A - \lambda B$. In B. Kågström and A. Ruhe, editors, *Matrix Pencils*, volume 973 of *Lecture Notes in Mathematics*, pages 17–29. Springer-Verlag, Berlin, 1983.

[85] V. N. Kublanovskaya. AB-algorithm and its modifications for the spectral problem of linear pencils of matrices. *Numer. Math.*, 43:329–342, 1984.

[86] J. J. Loiseau, K. Özǫaldiran, M. Malabre, and N. Karcanias. Feedback canonical forms of singular systems. *Kybernetika*, 27(4):289–305, 1991.

[87] A. S. Markus and E. É. Parilis. The change of the Jordan structure of a matrix under small perturbations. *Linear Algebra Appl.*, 54:139–152, 1983.

[88] E. Mengi. *Measures for robust stability and controllability.* PhD thesis, Courant Institute of Mathematical Sciences, New York, NY, 2006.

[89] P. Misra, P. Van Dooren, and A. Varga. Computation of structural invariants of generalized state space systems. *Automatica*, 30:1921–1936, 1994.

[90] B. P. Molinari. Structural invariants of linear multivariable systems. *Int. J. Control*, 28(4):493–510, 1978.

[91] A. S. Morse. Structural invariants of linear multivariable systems. *SIAM J. Control*, 11(3):446–465, 1973.

[92] C. Oară and P. Van Dooren. An improved algorithm for the computation of structural invariants of a system pencil and related geometric aspects. *Systems & Control Lett.*, 30:39–48, 1997.

[93] K. Özçaldrian and V. Eldem. A complete classification of minimal realizations of nonsingular and strict proper transfers under similarity transformations. *Automatica*, 33(5):835–850, 1997.

[94] C. Paige. Properties of numerical algorithms related to computing controllability. *IEEE Trans. Autom. Contr.*, AC-26(1):130–138, 1981.

[95] R. V. Patel, A. J. Laub, and P. Van Dooren, editors. *Numerical linear algebra techniques for systems and control.* Reprint Book Series. IEEE Press, New York, 1994.

[96] D. D. Pervouchine. Hierarchy of closures of matrix pencils. *Journal of Lie Theory*, 14:443–479, 2004.

[97] P. Hr. Petkov, N. D. Christov, and M. M. Konstantinov. *Computational methods for linear control systems.* Prentice Hall, Hertfordshire, UK, 1991. ISBN 0-13-161803-2.

[98] A. Pokrzywa. On perturbations and the equivalence orbit of a matrix pencil. *Linear Algebra Appl.*, 82:99–121, 1986.

[99] F. Puerta, X. Puerta, and S. Tarragona. Versal deformations in orbit spaces. *Linear Algebra Appl.*, 379:329–343, 2004.

[100] H. H. Rosenbrock. *State-space and multivariable theory.* Wiley, New York, NY, 1970.

[101] H. H. Rosenbrock. Structural properties of linear dynamical systems. *Int. J. Control*, 20:191–202, 1974.

[102] A. Ruhe. An algorithm for numerical determination of the structure of a general matrix. *BIT*, 10:196–216, 1970.

[103] A. Tannenbaum. *Invariance and system theory: Algebraic and geometric aspects.* Lecture Notes in Math. 845. Springer-Verlag, New York, 1981.

[104] J. S. Thorp. The singular pencil of a linear dynamical system. *Int. J. Control*, 18(3):577–596, 1973.

[105] D. Vafiadis and N. Karcanias. Canonical forms for descriptor systems under restricted system equivalence. *Automatica*, 33(5):955–958, 1997.

[106] P. Van Dooren. The computation of Kronecker's canonical form of a singular pencil. *Linear Algebra Appl.*, 27:103–141, 1979.

[107] P. Van Dooren. The generalized eigenstructure problem in linear system theory. *IEEE Trans. Autom. Contr.*, AC-26(1):111–129, 1981.

[108] P. Van Dooren. Reducing subspaces: Definitions, properties and algorithms. In B. Kågström and A. Ruhe, editors, *Matrix Pencils, Proc. Pite Havsbad, 1982*, volume 973 of *Lecture Notes in Mathematics*, pages 58–73. Springer-Verlag, Berlin, 1983.

[109] P. Van Dooren and M. Verhaegen. On the use of unitary state-space transformations. In *Special Issue of Contemporary Mathematics in Linear Algebra and Its Role in Systems Theory*, volume 47. Amer. Math. Soc., Providence, R.I., 1985.

[110] A. Varga. Computation of Kronecker-like forms of a system pencil: Applications, algorithms and software. Technical Report TR R181-95, Inst. Robotics and System Dynamics, DLR-Oberpfaffenhofen, January 1995.

[111] A. Varga. Computation of Kronecker-like forms of a system pencil: Applications, algorithms and software. In *Proc. of IEEE International Symposium on Computer Aided Control System Design, CACSD'96*, pages 77–82, Dearborn, MI, 1996.

[112] G. Verghese, P. Van Dooren, and T. Kailath. Properties of the system matrix of a generalized state-space system. *Int. J. Control*, 30(2):235–243, October 1979.

[113] W. Waterhouse. The codimension of singular matrix pairs. *Linear Algebra Appl.*, 57:227–245, 1984.

[114] J. Wilkinson. *The algebraic eigenvalue problem.* Oxford University Press, Oxford, 1965.

[115] J. Willems. Topological classification and structural stability of linear systems. *J. Differential Equations*, 35:306–318, 1980.

[116] W. M. Wonham. *Linear multivariable control theory: A geometric approach.* Applications of mathematics. Springer-Verlag, Berlin, third edition, 1985.

[117] I. Zaballa. Matrices with prescribed rows and invariant factors. *Linear Algebra Appl.*, 87:113–146, 1987.

[118] Z. Zhou, M. A. Shayman, and T-J. Tarn. Singular systems: A new approach in the time domain. *IEEE Trans. Autom. Contr.*, AC-32:42–50, 1987.

A Codimensions of orbits and bundles

Presented below are the explicit expressions for computing the codimension of the state-space system (or independent subsystems of)

$$\dot{x}(t) = Ax(t) + Bu(t),$$
$$y(t) = Cx(t) + Du(t),$$

where $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{n\times m}$, $C \in \mathbb{C}^{p\times n}$ and $D \in \mathbb{C}^{p\times m}$, and the general matrix pencil $G - \lambda H$, where $G,\ H \in \mathbb{C}^{m_{\mathrm{p}}\times n_{\mathrm{p}}}$, with the following invariants:

- The column minimal indices $\epsilon = (\epsilon_1, \ldots, \epsilon_{r_1}, \epsilon_{r_1+1}, \ldots, \epsilon_{r_0})$, where $\epsilon_i \geq 1$ for $i = 1, \ldots, r_1$ and $\epsilon_i = 0$ for $i = r_1 + 1, \ldots, r_0$.
- The row minimal indices $\eta = (\eta_1, \ldots, \eta_{l_1}, \eta_{l_1+1}, \ldots, \eta_{l_0})$, where $\eta_i \geq 1$ for $i = 1, \ldots, l_1$ and $\eta_i = 0$ for $i = l_1 + 1, \ldots, l_0$.
- The Segre characteristics $h_{\mu_i} = (h_1^{(i)}, \ldots, h_{g_i}^{(i)})$, for the finite eigenvalue μ_i, $i = 1, \ldots, q$.
- The Segre characteristics $s = (s_1, \ldots, s_t, s_{t+1}, \ldots, s_{g_\infty})$, for the infinite eigenvalue where $s_i \geq 2$ for $i = 1, \ldots, t$ and $s_i = 1$ for $i = t + 1, \ldots, g_\infty$, i.e., t is the number of N_i blocks of size $i \geq 2$

The codimension of an orbit/bundle can explicitly be determined from the above invariants. In the following, we summarize how the codimension is computed in the orbit case. For all systems the codimension of the bundle is given as:

$$\mathrm{cod}(\mathcal{B}(*)) = \mathrm{cod}(\mathcal{O}(*)) - (\text{number of distinct eigenvalues}).$$

Codimension of the orbit of a matrix A [23]

$$\mathrm{cod}(A) = \sum_{i=1}^{q}\sum_{j=1}^{g_i}(2j-1)h_j^{(i)}. \tag{A.46}$$

Comments: (A.46) comes from the sizes of the Jordan blocks for the finite eigenvalues.

Codimension of the orbit of a general matrix pencil $G - \lambda H$ [23]

$$\begin{aligned}
\mathrm{cod}(G - \lambda H) = &\sum_{\epsilon_i > \epsilon_j}(\epsilon_i - \epsilon_j - 1) &\text{(A.47)}\\
&+ \sum_{\eta_i > \eta_j}(\eta_i - \eta_j - 1) &\text{(A.48)}\\
&+ \sum_{\epsilon_i, \eta_j}(\epsilon_i + \eta_j + 2) &\text{(A.49)}\\
&+ (r_0 + l_0)\left(\sum_{i=1}^{q}\sum_{j=1}^{g_i} h_j^{(i)} + \sum_{j=1}^{g_\infty} s_j\right) &\text{(A.50)}\\
&+ \sum_{i=1}^{q}\sum_{j=1}^{g_i}(2j-1)h_j^{(i)} + \sum_{j=1}^{g_\infty}(2j-1)s_j. &\text{(A.51)}
\end{aligned}$$

Comments: (A.47) and (A.48) come from the interaction between the L blocks and the L^T blocks, respectively. (A.49) comes from the interaction between the right and left singular blocks and is the summation over all pairs of L_{ϵ_i} and $L^T_{\eta_j}$ blocks. (A.50) is the product of the number of right singular blocks and the total size of the regular part, and (A.51) comes from the sizes of the Jordan blocks (as in (A.46)) for the finite and infinite eigenvalues.

Codimension of the orbit of a controllability pair (A, B) [39]

$$\operatorname{cod}(A,B) = \sum_{\epsilon_i > \epsilon_j} (\epsilon_i - \epsilon_j - 1) \tag{A.52}$$

$$+ r_0 \sum_{i=1}^{q} \sum_{j=1}^{g_i} h_j^{(i)} \tag{A.53}$$

$$+ \sum_{i=1}^{q} \sum_{j=1}^{g_i} (2j-1) h_j^{(i)}. \tag{A.54}$$

Comments: (A.52) comes from the interaction between the L blocks, (A.53) is the product of the number of right singular blocks and the total size of the regular part, and (A.54) comes from the sizes of the Jordan blocks for the finite eigenvalues.

The controllability system pencil $\begin{bmatrix} A - \lambda I_n & B \end{bmatrix}$ has full row-rank and cannot have L^T blocks or infinite eigenvalues.

Codimension of the orbit of an observability pair (A, C)

$$\operatorname{cod}(A,C) = \sum_{\eta_i > \eta_j} (\eta_i - \eta_j - 1) \tag{A.55}$$

$$+ l_0 \sum_{i=1}^{q} \sum_{j=1}^{g_i} h_j^{(i)} \tag{A.56}$$

$$+ \sum_{i=1}^{q} \sum_{j=1}^{g_i} (2j-1) h_j^{(i)}. \tag{A.57}$$

Comments: (A.55) comes from the interaction between the L^T blocks, (A.56) is the product of the number of left singular blocks and the total size of the regular part, and (A.57) comes from the sizes of the Jordan blocks for the finite eigenvalues.

The observability system pencil $\begin{bmatrix} A - \lambda I_n \\ C \end{bmatrix}$ has full column-rank and cannot have L blocks or infinite eigenvalues. The terms (A.55)–(A.57) follow by duality from the results for the controllability pair.

Codimension of the orbit of a triple (A, B, C) [51]

$$\begin{aligned}
\operatorname{cod}(A, B, C) = & \sum_{\epsilon_i > \epsilon_j} (\epsilon_i - \epsilon_j - 1) && \text{(A.58)} \\
& + \sum_{\eta_i > \eta_j} (\eta_i - \eta_j - 1) && \text{(A.59)} \\
& + \sum_{\epsilon_i, \eta_j} (\epsilon_i + \eta_j) && \text{(A.60)} \\
& + (r_0 + l_0) \sum_{i=1}^{q} \sum_{j=1}^{g_i} h_j^{(i)} && \text{(A.61)} \\
& + \sum_{i=1}^{q} \sum_{j=1}^{g_i} (2j-1) h_j^{(i)} && \text{(A.62)} \\
& + \sum_{i=1}^{t} (2i-1)(s_i - 2) && \text{(A.63)} \\
& + (m - r_1 - t) \sum_{i=1}^{t} (s_i - 2) && \text{(A.64)} \\
& + (p - l_1 - t) \sum_{i=1}^{t} (s_i - 2) && \text{(A.65)} \\
& + \begin{cases} \sum_{i=1}^{t} (s_i - 2), & \text{if } r_1 > 0, \\ 0, & \text{otherwise} \end{cases} && \text{(A.66)} \\
& + \begin{cases} \sum_{i=1}^{t} (s_i - 2), & \text{if } l_1 > 0, \\ 0, & \text{otherwise.} \end{cases} && \text{(A.67)}
\end{aligned}$$

Comments: In the following, A_ϵ, A_η, A_∞, and A_μ refer to blocks in GBCF (see Section 2.6).

(A.58) and (A.59) come from the interaction between the L blocks and the L^T blocks, respectively. (A.60) comes from the interaction between the right and left singular blocks and is the summation over all pairs of L_{ϵ_i} and $L^T_{\eta_j}$ blocks in A_ϵ and A_η, respectively. (A.61) is the product of the number of right singular blocks and the size of the block A_μ. (A.62) comes from the sizes of the Jordan blocks for the finite eigenvalues, and (A.63) from the sizes of the Jordan blocks for the infinite eigenvalue in A_∞. (A.64) and (A.65) are the products of the number of L_0 and L_0^T blocks, respectively, and the size of A_∞ minus t (t is the number of N_i blocks of size $i \geq 2$). (A.66) and (A.67) add for each existing block A_ϵ and A_η, the size of A_∞ minus t.

Codimension of the orbit of a quadruple (A, B, C, D) [49]

$$\operatorname{cod}(A,B,C,D) = \sum_{\epsilon_i > \epsilon_j} (\epsilon_i - \epsilon_j - 1) \tag{A.68}$$

$$+ \sum_{\eta_i > \eta_j} (\eta_i - \eta_j - 1) \tag{A.69}$$

$$+ \sum_{\epsilon_i, \eta_j} (\epsilon_i + \eta_j) \tag{A.70}$$

$$+ (r_0 + l_0) \sum_{i=1}^{q} \sum_{j=1}^{g_i} h_j^{(i)} \tag{A.71}$$

$$+ \sum_{i=1}^{q} \sum_{j=1}^{g_i} (2j-1) h_j^{(i)} \tag{A.72}$$

$$+ \sum_{i=1}^{t} (2i-1)(s_i - 2) \tag{A.73}$$

$$+ (m - r_1 - g_\infty) \sum_{i=1}^{t} (s_i - 2) \tag{A.74}$$

$$+ (p - l_1 - g_\infty) \sum_{i=1}^{t} (s_i - 2) \tag{A.75}$$

$$+ \begin{cases} \sum_{i=1}^{t} (s_i - 2), & \text{if } r_1 > 0, \\ 0, & \text{otherwise} \end{cases} \tag{A.76}$$

$$+ \begin{cases} \sum_{i=1}^{t} (s_i - 2), & \text{if } l_1 > 0, \\ 0, & \text{otherwise} \end{cases} \tag{A.77}$$

$$+ (r_0 + t)(l_0 + t). \tag{A.78}$$

Comments: In the following, A_ϵ, A_η, A_∞, A_μ, D_∞, and D_B refer to blocks in GBCF.

(A.68) and (A.69) come from the interaction between the L blocks and the L^T blocks, respectively. (A.70) comes from the interaction between the right and left singular blocks and is the summation over all pairs of L_{ϵ_i} and $L_{\eta_j}^T$ blocks in A_ϵ and A_η, respectively. (A.71) is the product of the number of right singular blocks and the size of the block A_μ. (A.72) comes from the sizes of the Jordan blocks for the finite eigenvalues, and (A.73) from the sizes of the Jordan blocks for the infinite eigenvalue in A_∞. (A.74) and (A.75) are the products of the number of L_0 and L_0^T blocks, respectively, and the size of A_∞ minus t (t is the number of N_i blocks of size $i \geq 2$). (A.76) and (A.77) add for each existing block A_ϵ and A_η, the size of A_∞ minus t. (A.78) is the size of D_B minus its rank (size of D_∞).

B Stratification rules of orbits and bundles

In this appendix, we summarize the stratification rules for orbits and bundles of matrices, matrix pencils, and matrix pairs. From the integer partitions representing the canonical structure (information) of an orbit or a bundle, the stratification rules find covering and covered orbits or bundles, respectively. The following structure integer partitions are defined for each system, where each partition has a corresponding set of coins (see Section 2.4 for definitions):

- $\mathcal{J}_{\mu_i}$ for a matrix A, $\mu_i \in \mathbb{C}$.
- $\mathcal{R}$, $\mathcal{L}$ and $\mathcal{J}_{\mu_i}$ for a matrix pencil $G - \lambda H$, $\mu_i \in \overline{\mathbb{C}}$.
- $\mathcal{R}$ and $\mathcal{J}_{\mu_i}$ for a controllability pair (A, B), $\mu_i \in \mathbb{C}$.
- $\mathcal{L}$ and $\mathcal{J}_{\mu_i}$ for an observability pair (A, C), $\mu_i \in \mathbb{C}$.

Stratification rules of orbits and bundles of a matrix A

TABLE 2: Given the structure integer partitions $\mathcal{J}_{\mu_i}$ of A, one of the following *if-and-only-if* rules finds $\widetilde{A}$ fulfilling orbit or bundle covering relations with A [2, 31].

A. $\mathcal{O}(A)$ covers $\mathcal{O}(\widetilde{A})$: (1) Minimum *leftward* coin move in any $\mathcal{J}_{\mu_i}$.	**B.** $\mathcal{O}(A)$ is covered by $\mathcal{O}(\widetilde{A})$ (1) Minimum *rightward* coin move in any $\mathcal{J}_{\mu_i}$.
C. $\mathcal{B}(A)$ covers $\mathcal{B}(\widetilde{A})$: (1) Minimum *leftward* coin move in any $\mathcal{J}_{\mu_i}$. (2) Let any pair of eigenvalues coalesce, i.e., take the union of their sets of coins.	**D.** $\mathcal{B}(A)$ is covered by $\mathcal{B}(\widetilde{A})$: (1) Minimum *rightward* coin move in any $\mathcal{J}_{\mu_i}$. (2) For any $\mathcal{J}_{\mu_i}$, divide the set of coins into two new sets so that their union is $\mathcal{J}_{\mu_i}$.

Comments: For orbits (cases **A** and **B**), the number of eigenvalues and the total size of all blocks associated with the same eigenvalue are the same for all orbits in the closure hierarchy. This in contrast to bundles (cases **C** and **D**) where eigenvalues can coalesce and split apart, respectively.

Stratification rules of orbits and bundles of a matrix pencil $G - \lambda H$

TABLE 3: Given the structure integer partitions $\mathcal{L}$, $\mathcal{R}$ and $\mathcal{J}_{\mu_i}$ of $G - \lambda H$, where $\mu_i \in \overline{\mathbb{C}}$, one of the following *if-and-only-if* rules finds $\widetilde{G} - \lambda \widetilde{H}$ fulfilling orbit or bundle covering relations with $G - \lambda H$ [31].

A. $\mathcal{O}(G - \lambda H)$ covers $\mathcal{O}(\widetilde{G} - \lambda \widetilde{H})$:

(1) Minimum *rightward* coin move in $\mathcal{R}$ (or $\mathcal{L}$).

(2) If the rightmost column in $\mathcal{R}$ (or $\mathcal{L}$) is one single coin, move that coin to a new rightmost column of some $\mathcal{J}_{\mu_i}$ (which may be empty initially).

(3) Minimum *leftward* coin move in any $\mathcal{J}_{\mu_i}$.

(4) Let k denote the total number of coins in all of the longest (= lowest) rows from all of the $\mathcal{J}_{\mu_i}$. Remove these k coins, add one more coin to the set, and distribute $k+1$ coins to r_p, $p = 0, \dots, t$ and l_q, $q = 0, \dots, k-t-1$ such that at least all nonzero columns of $\mathcal{R}$ and $\mathcal{L}$ are given coins.

Rules 1 and 2 are not allowed to do coin moves that affect r_0 (or l_0).

B. $\mathcal{O}(G - \lambda H)$ is covered by $\mathcal{O}(\widetilde{G} - \lambda \widetilde{H})$:

(1) Minimum *leftward* coin move in $\mathcal{R}$ (or $\mathcal{L}$), without affecting r_0 (or l_0).

(2) If the rightmost column in some $\mathcal{J}_{\mu_i}$ consists of one coin only, move that coin to a new rightmost column in $\mathcal{R}$ (or $\mathcal{L}$), where $\mathcal{R}$ (or $\mathcal{L}$) is previously non-empty.

(3) Minimum *rightward* coin move in any $\mathcal{J}_{\mu_i}$.

(4) Remove one coin from each column of $\mathcal{R}$ and $\mathcal{L}$. Subtract one coin from this set and distribute the remaining coins on all $\mathcal{J}_{\mu_i}$ as follows. First, all nonzero columns in each set for all eigenvalues are given one coin each. Remaining coins are assigned to new (rightmost) columns of existing $\mathcal{J}_{\mu_i}$ or on new sets (for new eigenvalues).

C. $\mathcal{B}(G - \lambda H)$ covers $\mathcal{B}(\widetilde{G} - \lambda \widetilde{H})$:

(1) Same as rule 1 above.

(2) Same as rule 2 above, except it is only allowed to start a new set corresponding to a new eigenvalue (i.e., no appending to nonempty sets).

(3) Same as rule 3 above.

(4) Same as rule 4 above, but apply only if there exists only one set of coins corresponding to one eigenvalue, or if all sets corresponding to each eigenvalue have at least two rows of coins.

(5) Let any pair of eigenvalues coalesce, i.e., take the union of their sets of coins.

D. $\mathcal{B}(G - \lambda H)$ is covered by $\mathcal{B}(\widetilde{G} - \lambda \widetilde{H})$:

(1) Same as rule 1 above.

(2) Same as rule 2 above, except that $\mathcal{J}_{\mu_i}$ must consist of one coin only.

(3) Same as rule 3 above.

(4) Same as rule 4 above, except that a new set for a new eigenvalue may only be created if there exist no $\mathcal{J}_{\mu_i}$. If a new set is created, all coins should be assigned to it and create one row.

(5) For any $\mathcal{J}_{\mu_i}$, divide the set of coins into two new partitions so that their union is $\mathcal{J}_{\mu_i}$.

Comments: The restriction for rules **A**.(1) and **A**.(2) implies that the number of left and right singular blocks remain fixed, while rule (4) adds one new block of each kind and rule (3) corresponds to the nilpotent case. Rule (4) cannot be applied if the total number of nonzero columns in $\mathcal{R}$ and $\mathcal{L}$ are more than $k+1$. If the rule can be applied, at least one coin must be assigned to $\mathcal{R}$ and $\mathcal{L}$, respectively. Expressed in KCF, the restriction of rule **C**.4 means that it can only be applied if there is just one eigenvalue or if all eigenvalues have at least two Jordan blocks.

Notably, the eigenvalue μ_i corresponding to the structure integer partition $\mathcal{J}_{\mu_i}$ belongs to the extended complex plane, i.e., $\mu_i \in \mathbb{C} \cup \{\infty\}$.

Stratification rules of orbits and bundles of a controllability pair (A, B)

Table 4: Given the structure integer partitions $\mathcal{R}$ and $\mathcal{J}_{\mu_i}$ of (A, B), one of the following *if-and-only-if* rules finds $(\widetilde{A}, \widetilde{B})$ fulfilling orbit or bundle covering relations with (A, B) [36].

A. $\mathcal{O}(A, B)$ covers $\mathcal{O}(\widetilde{A}, \widetilde{B})$ (1) Minimum rightward coin move in $\mathcal{R}$. (2) If the rightmost column in $\mathcal{R}$ is one single coin, move that coin to a new rightmost column of some $\mathcal{J}_{\mu_i}$ (which may be empty initially). (3) Minimum leftward coin move in any $\mathcal{J}_{\mu_i}$. Rules 1 and 2 are not allowed to do coin moves that affect r_0.	**B.** $\mathcal{O}(A, B)$ is covered by $\mathcal{O}(\widetilde{A}, \widetilde{B})$ (1) Minimum leftward coin move in $\mathcal{R}$, without affecting r_0. (2) If the rightmost column in some $\mathcal{J}_{\mu_i}$ consists of one coin only, move that coin to a new rightmost column in $\mathcal{R}$. (3) Minimum rightward coin move in any $\mathcal{J}_{\mu_i}$.
C. $\mathcal{B}(A, B)$ covers $\mathcal{B}(\widetilde{A}, \widetilde{B})$ (1) Same as rule 1 above. (2) Same as rule 2 above, except it is only allowed to start a new set corresponding to a new eigenvalue (i.e., no appending to nonempty sets). (3) Same as rule 3 above. (4) Let any pair of eigenvalues coalesce, i.e., take the union of their sets of coins.	**D.** $\mathcal{B}(A, B)$ is covered by $\mathcal{B}(\widetilde{A}, \widetilde{B})$ (1) Same as rule 1 above. (2) Same as rule 2 above, except that $\mathcal{J}_{\mu_i}$ must consist of one coin only. (3) Same as rule 3 above. (4) For any $\mathcal{J}_{\mu_i}$, divide the set of coins into two new sets so that their union is $\mathcal{J}_{\mu_i}$.

Comments: The rules for the matrix pair (A, B) differ from the rules for a general matrix pencil in that rule (4) in Table 3 (both for orbits and bundles) cannot be applied to the matrix pair (A, B), since there cannot exist L^T blocks in (A, B). Moreover, rules (1) and (2) only apply to the structure integer partition $\mathcal{R}$.

Stratification rules of orbits and bundles of a observability pair (A, C)

TABLE 5: Given the structure integer partitions $\mathcal{L}$ and $\mathcal{J}_{\mu_i}$ of (A, C), one of the following *if-and-only-if* rules finds $(\widetilde{A}, \widetilde{C})$ fulfilling orbit or bundle covering relations with (A, C) [36].

A. $\mathcal{O}(A, C)$ covers $\mathcal{O}(\widetilde{A}, \widetilde{C})$:

(1) Minimum rightward coin move in $\mathcal{L}$.

(2) If the rightmost column in $\mathcal{L}$ is one single coin, move that coin to a new rightmost column of some $\mathcal{J}_{\mu_i}$ (which may be empty initially).

(3) Minimum leftward coin move in any $\mathcal{J}_{\mu_i}$.

Rules 1 and 2 are not allowed to do coin moves that affect l_0.

B. $\mathcal{O}(A, C)$ is covered by $\mathcal{O}(\widetilde{A}, \widetilde{C})$:

(1) Minimum leftward coin move in $\mathcal{L}$, without affecting l_0.

(2) If the rightmost column in some $\mathcal{J}_{\mu_i}$ consists of one coin only, move that coin to a new rightmost column in $\mathcal{L}$.

(3) Minimum rightward coin move in any $\mathcal{J}_{\mu_i}$.

C. $\mathcal{B}(A, C)$ covers $\mathcal{B}(\widetilde{A}, \widetilde{C})$:

(1) Same as rule 1 above.

(2) Same as rule 2 above, except it is only allowed to start a new set corresponding to a new eigenvalue (i.e., no appending to nonempty sets).

(3) Same as rule 3 above.

(4) Let any pair of eigenvalues coalesce, i.e., take the union of their sets of coins.

D. $\mathcal{B}(A, C)$ is covered by $\mathcal{B}(\widetilde{A}, \widetilde{C})$:

(1) Same as rule 1 above.

(2) Same as rule 2 above, except that $\mathcal{J}_{\mu_i}$ must consist of one coin only.

(3) Same as rule 3 above.

(4) For any $\mathcal{J}_{\mu_i}$, divide the set of coins into two new sets so that their union is $\mathcal{J}_{\mu_i}$.

Comments: The rules for the matrix pair (A, C) differ from the rules for a general matrix pencil in that rule (4) in Table 3 (both for orbits and bundles) cannot be applied to the matrix pair (A, C), since there cannot exist L blocks in (A, C). Moreover, rules (1) and (2) only apply to the structure integer partition $\mathcal{L}$.

Notably, the rules for the matrix pair (A, C) are dual to those for the matrix pair (A, B).

C Notation

$\inf\{\mathcal{A}\}$	The greatest lower bound of a set $\mathcal{A}$.
$\sup\{\mathcal{A}\}$	The least upper bound of a set $\mathcal{A}$.
$\mathcal{A} \supseteq \mathcal{B}$	The set $\mathcal{B}$ is a subset of $\mathcal{A}$, i.e., every member of $\mathcal{B}$ is a member of $\mathcal{A}$.
$\mathcal{A} \supset \mathcal{B}$	The set $\mathcal{B}$ is a proper subset of $\mathcal{A}$, i.e., $\mathcal{A} \supseteq \mathcal{B}$ and $\mathcal{A} \neq \mathcal{B}$.
κ	$\kappa = (\kappa_1, \kappa_2, \ldots)$ is an integer partition with $\kappa_1 \geq \kappa_2 \geq \cdots \geq 0$. Also ν and τ are used.
$\sum \kappa$	The sum $\kappa_1 + \kappa_2 + \cdots$ of κ.
$\kappa + m$	$(\kappa_1 + m, \kappa_2 + m, \ldots)$ where m is a scalar.
$\mathrm{conj}(\kappa)$	The conjugate partition of κ.
$\kappa \cup \nu$	The union of κ and ν.
$\kappa \setminus \nu$	The difference between κ and ν.
$\kappa \geq \nu$	$\kappa_1 + \cdots + \kappa_i \geq \nu_1 + \cdots + \nu_i$ for all $i = 1, 2, \ldots$
$\kappa > \nu$	κ dominates ν, i.e., $\kappa \geq \nu$ and $\kappa \neq \nu$.
$\mathbb{R}$	The field of real numbers.
$\mathbb{C}$	The field of complex numbers.
$\overline{\mathbb{C}}$	$\mathbb{C} \cup \{\infty\}$, i.e., the extended complex plane.
$\mathbb{C}^{m \times n}$	The set of complex matrices of order $m \times n$.
A	A square matrix of size $n \times n$. I or I_n is the identity matrix.
A^T	The transpose of A.
A^H	The conjugate transpose of A.
$Gl_n(\mathbb{C})$	The linear group of order n over $\mathbb{C}$. If $A \in Gl(\mathbb{C})$ then the $n \times n$ matrix A is nonsingular.
$\mathrm{vec}(A)$	An ordered stack of the columns of a matrix A from left to right.
$\mathrm{null}(A)$	Null space (kernel) of the space spanned by the columns of A.
$\mathrm{ran}(A)$	Range (image) of the space spanned by the columns of A.
$\mathrm{diag}(A_1, \ldots, A_b)$	A block diagonal matrix with diagonal blocks A_i.
$A \otimes B$	The Kronecker product of two matrices A and B whose (i, j)-th block element is $a_{ij}B$.
$A \equiv A_1 \oplus A_2 \oplus \cdots$	Direct sum of matrices, $A = \mathrm{diag}(A_1, A_2, \ldots)$.
(E, A, B, C, D)	Matrix tuple representing an LTI system associated with $\left\{\begin{smallmatrix} E\dot{x} & = & Ax(t)+Bu(t) \\ y & = & Cx(t)+Du(t) \end{smallmatrix}\right.$. Independent subsystems are represented by subsets of the tuple, e.g. (E, A, B, C), (A, B), (A, C), etc.
$G - \lambda H$	A general matrix pencil of size $m_{\mathrm{p}} \times n_{\mathrm{p}}$.
$\mathbf{S}(\lambda)$	The system pencil $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] - \lambda \left[\begin{smallmatrix} E & 0 \\ 0 & 0 \end{smallmatrix}\right]$ of size $(n + p) \times (n + m)$ corresponding to the tuple (E, A, B, C, D).
$\mathbf{S}_{\mathrm{C}}(\lambda)$	The controllability system pencil $\left[\begin{smallmatrix} A & B \end{smallmatrix}\right] - \lambda \left[\begin{smallmatrix} E & 0 \end{smallmatrix}\right]$, where in most cases $E = I$.
$\mathbf{S}_{\mathrm{O}}(\lambda)$	The observability system pencil $\left[\begin{smallmatrix} A \\ C \end{smallmatrix}\right] - \lambda \left[\begin{smallmatrix} E \\ 0 \end{smallmatrix}\right]$, where in most cases $E = I$.
$\mathbf{C}(A, B)$	The controllability matrix.
$\mathbf{O}(A, C)$	The observability matrix.
$\mathbf{C}_{\mathcal{S}}(A, B)$	The controllable subspace of (A, B).
$\overline{\mathbf{O}}_{\mathcal{S}}(A, C)$	The unobservable subspace of (A, C).
Ω	Abbreviation used in the following for a matrix A, matrix pencil $G - \lambda H$, or a system pencil $\mathbf{S}(\lambda)$.

μ_i	Eigenvalue of Ω (also α and β are used). Can also be represented by the pair of eigenvalues (α_i, β_i), where $\mu_i = \alpha_i/\beta_i$ if $\beta_i \neq 0$ else μ_i is the infinite eigenvalue.
$\mathcal{O}(\Omega)$	The orbit of Ω, i.e. the set of similar matrices or equivalent matrix or system pencils to Ω (canonical structure and eigenvalues fixed).
$\overline{\mathcal{O}}(\Omega)$	The closure of an orbit.
$\mathcal{B}(\Omega)$	The bundle of Ω, $\cup_{\mu_i}\mathcal{O}(\Omega)$. Eigenvalues not specified.
$\overline{\mathcal{B}}(\Omega)$	The closure of a bundle.
$\mathrm{nrk}\,(\Omega)$	The normal rank of Ω, i.e., the order of Ω's greatest minor different from polynomial zero.
$\tan(\Omega)$	The tangent space of $\mathcal{O}(\Omega)$ at Ω.
$\mathrm{nor}(\Omega)$	The normal space of $\mathcal{O}(\Omega)$ at Ω, i.e., the orthogonal complement to the tangent space.
$\dim(\Omega)$	Dimension of $\mathcal{O}(\Omega)$.
$\mathrm{cod}(\Omega)$	Codimension of $\mathcal{O}(\Omega)$, where $\dim(\Omega) + \mathrm{cod}(\Omega)$ is equal to the dimension of the complete space Ω, e.g. matrices belongs to a n^2-dimensional space and matrix pencils to a $2m_{\mathrm{p}}n_{\mathrm{p}}$-dimensional space.
$D_j(A)$	The greatest common divisors of all the minors of order j of the matrix A.
d_{μ_i}	$d_{\mu_i} = (d_0^{(i)}, \ldots, d_n^{(i)})$ is the integer partition representing the multiplicity $d_j^{(i)}$ of $(\lambda - \mu_i)$ in $D_j(A)$.
$P_j(A)$	The invariant factors of A.
q	Number of distinct finite eigenvalues.
g_i	The geometric multiplicity of the finite eigenvalue μ_i.
g_∞	The geometric multiplicity of the infinite eigenvalue.
r_0	Number of column minimal indices.
r_1	Number of column minimal indices greater than zero.
l_0	Number of row minimal indices.
l_1	Number of row minimal indices greater than zero.
h_{μ_i}	$h_{\mu_i} = (h_1^{(i)}, \ldots, h_{g_i}^{(i)})$ is the integer partition representing the Segre characteristics for the finite eigenvalue μ_i.
s	$s = (s_1, \ldots, s_{g_\infty})$ is the integer partition representing the Segre characteristics for the infinite eigenvalue.
ϵ	$\epsilon = (\epsilon_1, \ldots, \epsilon_{r_0})$ is the integer partition representing the column (left) minimal indices. The conjugate $r = (r_1, \ldots, r_{\epsilon_1})$ is the r-numbers.
η	$\eta = (\eta_1, \ldots, \eta_{l_0})$ is the integer partition representing the row (right) minimal indices. The conjugate $l = (l_1, \ldots, l_{\eta_1})$ is the l-numbers.
$\mathcal{J}_{\mu_i}$	$\mathcal{J}_{\mu_i} = (j_1, j_2, \ldots)$ is the integer partition representing the Weyr characteristics for the finite eigenvalue μ_i.
$\mathcal{N}$	$\mathcal{N} = (n_1, n_2, \ldots)$ is the integer partition representing the Weyr characteristics for the infinite eigenvalue.
$\mathcal{R}$	$\mathcal{R} = (r_0, r_1, \ldots) = (r_0) \cup \mathrm{conj}(\epsilon)$ is the integer partition representing the right singular structure.
$\mathcal{L}$	$\mathcal{L} = (l_0, l_1, \ldots) = (l_0) \cup \mathrm{conj}(\eta)$ is the integer partition representing the left singular structure.
$J_k(\mu_i)$	Jordan block of size $k \times k$ associated with the eigenvalue μ_i.

N_k	Jordan block of size $k \times k$ associated with the infinite eigenvalue.
L_k	Singular block of size $k \times (k+1)$ associated with a column (right) minimal index k.
L_k^T	Singular block of size $(k+1) \times k$ associated with a row (left) minimal index k.
JCF	Jordan canonical form; $PAP^{-1} = \mathrm{diag}(J(\mu_1), \ldots, J(\mu_q))$.
KCF	Kronecker canonical form; $U(G - \lambda H)V^{-1} = \mathrm{diag}(L, J, N, L^T)$.
BCF	Brunovsky canonical form; $P \begin{bmatrix} A-\lambda I & B \end{bmatrix} \begin{bmatrix} P^{-1} & 0 \\ R & Q^{-1} \end{bmatrix} = \begin{bmatrix} A_\epsilon & 0 & B_\epsilon & 0 \\ 0 & A_\mu & 0 & 0 \end{bmatrix}$, and $\begin{bmatrix} P & S \\ 0 & T \end{bmatrix} \begin{bmatrix} A-\lambda I \\ C \end{bmatrix} P^{-1} = \begin{bmatrix} A_\eta & 0 \\ 0 & A_\mu \\ C_\eta & 0 \\ 0 & 0 \end{bmatrix}$.
GBCF	Generalized Brunovsky canonical form; $\begin{bmatrix} P & S \\ 0 & T \end{bmatrix} \begin{bmatrix} A-\lambda I & B \\ C & D \end{bmatrix} \begin{bmatrix} P^{-1} & 0 \\ R & Q^{-1} \end{bmatrix} = \begin{bmatrix} A_\epsilon & 0 & 0 & 0 & B_\epsilon & 0 & 0 & 0 \\ 0 & A_\eta & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & A_\infty & 0 & 0 & B_\infty & 0 & 0 \\ 0 & 0 & 0 & A_\mu & 0 & 0 & 0 & 0 \\ 0 & C_\eta & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & C_\infty & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & D_\infty & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & D_0 \end{bmatrix}$.