# Linear Differential-Algebraic Equations of Higher-Order and the Regularity or Singularity of Matrix Polynomials


vorgelegt von

**Dipl.-Math. Chunchao Shi**


von der Fakultät II - Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
- Dr. rer. nat. -

genehmigte Dissertation

Promotionsausschuss:

| | |
|---|---|
| Vorsitzender: | Prof. Dr. Alexander Bobenko |
| Berichter: | Prof. Dr. Volker Mehrmann |
| Berichter: | PD Dr. Christian Mehl |
| Gutachter: | Prof. Dr. Ralph Byers |

Tag der wissenschaftlichen Aussprache: 12.07.2004


Berlin 2004

D 83

## Eidesstattliche Erklärung

Hiermit erkläre ich, daß ich die vorliegende Dissertation selbständig verfaßt habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt worden sind.

Berlin, den 07.05.2004 Chunchao Shi

# Erklärung

a) Kein Teil der eingereichten Dissertation wurde bisher veröffentlicht.

b) Ich habe weder jetzt noch in der Vergangenheit eine Anmeldung einer Promotionsabsicht bei einer anderen Hochschule oder anderen Fakultät beantragt.

Berlin, den 07.05.2004 Chunchao Shi

## ACKNOWLEDGEMENT

I am deeply indebted to my supervisor, Prof. Dr. Volker Mehrmann, for giving me the opportunity to complete this dissertation; but for his innumerable and generous consideration and help from the very beginning of my application onward, this work would not have been existent. Moreover, I would like to thank him for his introducing me to the research field, providing information on key references, critically reading the manuscript, and for his many helpful suggestions during this work.

I remember with deep gratitude the innumerable and magnanimous help that I have received from Dr. Christian Mehl from the beginning of my coming to Berlin onward. In particular, I would like to thank him for his critically reading the manuscript and giving me many helpful suggestions during the early stage of this thesis.

My deep gratitude also goes to Dr. Michael Karow who helped me translate the abstract and write the declaration, and gave me helpful suggestions during the preparation of this thesis.

I would like to thank Prof. Dr. Ralph Byers for his critically reading the manuscript and for his many helpful suggestions and comments.

I am very grateful to my colleagues who have helped me a lot with their kindness and patience during the years; especially, my sincere thanks go to Dr. Michael Lorenz (at TU-Chemnitz), Mrs. Wei Cai and Dr. Jungong Xue (at TU-Chemnitz), Mr. Andreas Steinbrecher, Mr. Christian Marwitz, and Mr. Michael Schmidt for their great and generous help in the aspects of both living and working.

Finally, I would like to acknowledge the financial supports I have received from "Forschungsthema VF 8621-215201-55" in Chemnitz, and "Projekte DFG-Gz. ME 790/13-1, ME 790/13-2, und ME 790/14-1" in Berlin during the different periods of time.

# List of Notations and Conventions

| | |
|---|---|
| $\mathbb{N}$ | set of natural numbers |
| $\mathbb{N}_0$ | $\mathbb{N} \cup \{0\}$ |
| $\mathbb{C}$ | field of complex numbers |
| $\mathbb{C}^n$ | space of all n-dimensional column vectors with components in $\mathbb{C}$ |
| $\mathbb{C}^{m\times n}$ | space of complex matrices of size $m \times n$ |
| $\mathcal{C}^q([t_0, t_1], \mathbb{C}^n)$ | set of all $q$-times continuously differentiable vector-valued functions mapping from the real interval $[t_0, t_1]$ to $\mathbb{C}^n$, where $q \in \mathbb{N}_0$ |
| $\mathcal{C}^q([t_0, t_1], \mathbb{C}^{m\times n})$ | set of all $q$-times continuously differentiable matrix-valued functions mapping from the real interval $[t_0, t_1]$ to $\mathbb{C}^{m\times n}$, where $q \in \mathbb{N}_0$ |
| $\overline{\lambda}$ | complex conjugate of $\lambda \in \mathbb{C}$ |
| $\arg(\cdot)$ | argument of a complex number |
| $[a_{i,j}]_{i,j=1}^{m,n}$ | matrix of dimension $m \times n$ |
| $I$ or $I_n$ | identity matrix of size $n \times n$ |
| $e_i$ | $i$-th unit vector |
| $A(i, j)$ | element of matrix $A$ lying on the $i$th row and the $j$th column of $A$ |
| $A(:, i)$ | $i$-th column of matrix $A$ |
| $A(:, i : (i + k))$ | submatrix of $A$ including its columns from the $i$-th to the $(i + k)$-th |
| $A(i, :)$ | $i$-th row of matrix $A$ |
| $A(i : (i + k), :)$ | submatrix of $A$ including its rows from the $i$-th to the $(i + k)$-th |
| $A^T$ | transpose of matrix $A$ |
| $A^H$ | conjugate transpose of matrix $A$ |
| $A^{-1}$ | inverse of matrix $A$ |
| $A^{-T}$ | inverse of the transpose of matrix $A$ |
| $A^{-H}$ | inverse conjugate transpose of matrix $A$ |
| $\lambda_{\max}(\cdot)$ | largest eigenvalue of a Hermitian matrix |
| $\sigma_{\min}(\cdot)$ | smallest singular value of a matrix |
| $\operatorname{diag}[A_1, \ldots, A_m]$ | $\begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_m \end{bmatrix}$ |
| $\mathcal{R}(\cdot)$ | column space of a matrix |
| $\mathcal{N}(\cdot)$ | null space of a matrix |
| $\operatorname{rank}(\cdot)$ | rank of a matrix, a matrix-valued function, or a matrix polynomial |
| $\det(\cdot)$ | determinant of a square matrix |

| | |
|---|---|
| $\mathrm{trace}(\cdot)$ | trace of a matrix |
| $\mathrm{span}\{x_1, \ldots, x_m\}$ | subspace spanned by vectors $x_1, \ldots, x_m \in \mathbb{C}^n$ |
| $\deg(\cdot)$ | degree of a polynomial |
| $\dim(\cdot)$ | dimension of a subspace |
| $\mathcal{X} \dotplus \mathcal{Y}$, $\mathcal{X}, \mathcal{Y} \subseteq \mathbb{C}^n$ | $\{x + y : x \in \mathcal{X}, y \in \mathcal{Y}\}$ |
| $A\mathcal{X}$, $A \in \mathbb{C}^{m \times n}$, $\mathcal{X} \subseteq \mathbb{C}^n$ | $\{Ax : x \in \mathcal{X}\}$ |
| $:=$ | equals by definition |
| $\square$ | end of proof |
| $\lozenge$ | end of an example or remark or algorithm |

## ZUSAMMENFASSUNG

Die Arbeit liefert einen Beitrag zur theoretischen Analyse linearer differentiell-algebraischer Gleichungen (DAEs) höherer Ordnung sowie der Regularität und Singularität von Matrixpolynomen.

Für Systeme von linearen DAEs höherer Ordnung mit variablen und konstanten Koeffizienten werden Invarianten und kondensierte Formen unter angemessenen Äquivalenztransformationen angegeben. Ausgehend von den kondensierten Formen kann das ursprüngliche DAE-System induktiv durch Differentiations- und Eliminationsschritte in ein strangeness-freies System transformiert werden, aus dem das Lösungsverhalten (u.a die Konsistenz der Anfangsbedingungen und die Eindeutigkeit der Lösung) direkt ablesbar ist.

Für quadratische DAE-Systeme mit konstanten Koeffizienten wird gezeigt, dass genau dann zu jeder konsistenten Anfangsbedingung und jeder rechten Seite $f(t) \in \mathcal{C}^{\mu}([t_0, t_1], \mathbb{C})$ eine eindeutige Lösung existiert, wenn das zugehörige Matrixpolynom regulär ist. Dabei ist $\mu$ der Strangeness-Index des Systems.

Es werden einige notwendige und hinreichende Bedingungen für die Zeilen-und Spaltenregularität und -singularität allgemeiner rechteckiger Matrixpolynome angegeben. Eine geometrische Charakterisierung singulärer Matrixbüschel wird ebenfalls hergeleitet. Darüber hinaus wird ein Algorithmus vorgestellt, durch den man mittels Rang-Informationen über die Koeffizientenmatrizen und Determinantenberechnungen bestimmen kann, ob ein gegebenes quadratisches Matrixpolynom regulär ist.

Ein weiteres Thema der Arbeit ist die Bestimmung des Abstands eines regulären Matrixpolynoms von der Menge der singulären Matrixpolynome. Es wird gezeigt, dass dieses Problem äquivalent zu der Aufgabe ist, in einer gewissen strukturierten Menge von Matrizen die nächstgelegene Matrix mit niedrigerem Rang zu finden. Für Matrixbüschel wird eine Charakterisierung des Abstands zur Singularität mit Hilfe von Matrixsingulärwerten angegeben. Schließlich werden einige untere Schranken für den Abstand zur Singularität hergeleitet.

ABSTRACT

This thesis contributes to the theoretical analysis of linear differential-algebraic equations (DAEs) of higher order as well as of the regularity and singularity of matrix polynomials.

Some invariants and condensed forms under appropriate equivalent transformations are given for systems of linear higher-order DAEs with constant and variable coefficients. Inductively, based on condensed forms the original DAE system can be transformed by differentiation-and-elimination steps into an equivalent strangeness-free system, from which the solution behaviour (including consistency of initial conditions and unique solvability) of the original DAE system and related initial value problem can be directly read off. It is shown that the following equivalence holds for a DAE system with strangeness-index $\mu$ and square and constant coefficients. For any consistent initial condition and any right-hand side $f(t) \in \mathcal{C}^{\mu}([t_0, t_1], \mathbb{C}^n)$ the associated initial value problem has a unique solution if and only if the matrix polynomial associated with the system is regular.

Some necessary and sufficient conditions for column- and row- regularity and singularity of rectangular matrix polynomials are derived. A geometrical characterization of singular matrix pencils is also given. Furthermore, an algorithm is presented which - using rank information about the coefficients matrices and via computing determinants - decides whether a given matrix polynomial is regular.

Another subject of the thesis is the determination of the distance of a regular matrix polynomial to the set of singular matrix polynomials. It is shown that this nearness problem is equivalent to a rank-deficiency problem for a certain class of structured and constrained perturbations. In addition, a characterization, in terms of the singular values of matrices, of the distance to singularity for matrix pencils is obtained. Finally, some lower bounds for the distance of a matrix polynomial to singularity are established.

# Contents

# Chapter 1

# Introduction

There are two parts in this thesis. In its first part, consisting of Chapters 2 and 3, we shall study linear $l$th-order differential-algebraic equations (DAEs) with constant coefficients

$$A_l x^{(l)}(t) + A_{l-1}x^{(l-1)}(t) + \cdots + A_0 x(t) = f(t),\ t \in [t_0, t_1] \quad \left( x^{(k)}(t) = \frac{d^k}{dt^k}x(t) \right) \tag{1.1}$$

and linear $l$th-order DAEs with variable coefficients

$$A_l(t)x^{(l)}(t) + A_{l-1}(t)x^{(l-1)}(t) + \cdots + A_0(t)x(t) = f(t),\ t \in [t_0, t_1], \tag{1.2}$$

where $A_i \in \mathbb{C}^{m\times n}$, $i = 0, 1, \ldots, l$, $l \in \mathbb{N}_0$, $A_l \neq 0$, $t$ is a real variable on the interval $[t_0, t_1]$, $A_i(t) \in \mathcal{C}([t_0, t_1], \mathbb{C}^{m\times n})$, $i = 0, 1, \ldots, l$, $A_l(t) \not\equiv 0$, $x(t)$ is an unknown vector-valued function in $\mathcal{C}([t_0, t_1], \mathbb{C}^n)$, and the right-hand side $f(t)$ is a given vector-valued function in $\mathcal{C}^\mu([t_0, t_1], \mathbb{C}^m)$. Here $\mathcal{C}^\mu([t_0, t_1], \mathbb{C}^{m\times n})$, $\mu \in \mathbb{N}_0$, denotes the set of all $\mu$-times continuously differentiable matrix-valued functions mapping from the real interval $[t_0, t_1]$ to the complex vector space $\mathbb{C}^{m\times n}$.

As the name "DAE" indicates, a system of DAEs is a system that consists of ordinary differential equations (ODEs) coupled with purely algebraic equations; in other words, DAEs are *everywhere singular implicit* ODEs (cf., for example, E. Griepentrog, M. Hanke and R. März [16]). Based on this notion, in this thesis we always call (1.1) and (1.2) systems of DAEs if $m \neq n$, and if $m = n$ it is always assumed that the leading coefficient matrix $A_l$ in the system (1.1) and the leading coefficient matrix-valued function $A_l(t)$ in the system (1.2) are singular, namely,

$$\operatorname{rank}(A_l) < n, \quad \text{and} \quad \operatorname{rank}(A_l(t)) < n,\ t \in [t_0, t_1].$$

Here, the *rank* of a matrix-valued function $A(t)$ on the interval $[t_0, t_1]$ is defined as

$$\operatorname{rank}(A(t)) = \max_{t_0 \le \nu \le t_1} \{\operatorname{rank}(A(\nu))\}. \tag{1.3}$$

Often, we will refer to linear DAEs with order greater than one simply as linear *higher-order* systems.

Systems of linear *first-order* DAEs with constant coefficients

$$A_1\dot{x}(t) + A_0x(t) = f(t),\ t \in [t_0, t_1], \tag{1.4}$$

where $\dot{x}(t)$ denotes the derivative of $x$ with respect to $t$, and systems of linear *first-order* DAEs with variable coefficients

$$A_1(t)\dot{x}(t) + A_0(t)x(t) = f(t),\ t \in [t_0, t_1], \tag{1.5}$$

as well as general nonlinear *first-order* DAEs

$$F(t, x(t), \dot{x}(t)) = 0,\ t \in [t_0, t_1], \tag{1.6}$$

where $F$ and $x(t)$ are vector-valued, play a key role in the modelling and simulation of constrained dynamical systems in numerous applications. Such systems have been intensively studied, theoretically as well as numerically, in the past two decades. For systematic and comprehensive exposition of important aspects regarding the theory and the numerical treatment of first-order DAEs, see, for example, the monographs of S. L. Campbell [7, 8] (1980, 1982), E. Griepentrog and R. März [18] (1986), K. E. Brenan, S. L. Campbell, and L. R. Petzold [3] (1996), P. Kunkel and V. Mehrmann [34] (in manuscript), and the references therein.

However, the systems (1.1) and (1.2) of linear *higher-order* DAEs also arise naturally and frequently in many mathematical models. Take, for example, a model for controlled multibody systems in R. Schüpphaus [51] ( p. 9) which can be formulated in the following system of linear second-order DAEs with constant coefficients:

$$\begin{bmatrix} M & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ J & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \ddot{z}(t) \\ \lambda(t) \\ \mu(t) \\ \nu(t) \\ \xi(t) \end{bmatrix} + \begin{bmatrix} 0 & P & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & G & 0 & 0 & 0 \\ 0 & L & 0 & 0 & 0 \\ 0 & Y & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{z}(t) \\ \lambda(t) \\ \mu(t) \\ \nu(t) \\ \xi(t) \end{bmatrix}$$
$$+ \begin{bmatrix} Q & -F^T & -G^T & -J^T & -\hat{Z} \\ F & 0 & 0 & 0 & 0 \\ H & 0 & 0 & 0 & 0 \\ K & 0 & 0 & 0 & 0 \\ X & 0 & 0 & 0 & Z \end{bmatrix} \begin{bmatrix} z(t) \\ \lambda(t) \\ \mu(t) \\ \nu(t) \\ \xi(t) \end{bmatrix} = \begin{bmatrix} Su(t) \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

where $M, P, Q \in \mathbb{C}^{u\times u}$, $J, L, K \in \mathbb{C}^{s\times u}$, $G, H \in \mathbb{C}^{q\times u}$, $Y, X, \hat{Z} \in \mathbb{C}^{v\times u}$, $F \in \mathbb{C}^{p\times u}$, $Z \in \mathbb{C}^{v\times v}$, $S \in \mathbb{C}^{u\times r}$, $z(t) \in \mathcal{C}([t_0, t_1], \mathbb{C}^u)$, $\lambda(t) \in \mathcal{C}([t_0, t_1], \mathbb{C}^p)$, $\mu(t) \in \mathcal{C}([t_0, t_1], \mathbb{C}^q)$,

$\nu(t) \in \mathcal{C}([t_0, t_1], \mathbb{C}^s)$, $\xi(t) \in \mathcal{C}([t_0, t_1], \mathbb{C}^v)$, $u(t) \in \mathcal{C}([t_0, t_1], \mathbb{C}^r)$, and $\ddot{(\cdot)}(t)$ denotes the second derivative of $(\cdot)$ with respect to $t$.

Usually, as in the well-known classical theory of ordinary differential equations, the method employed to treat systems (1.1) and (1.2) of higher-order DAEs is to transform them into first-order systems by introducing the derivative, the second derivative, ..., the $(l-1)$th derivative of the unknown vector-valued function as a part of a new enlarged unknown vector-valued function, and then to solve the first-order systems of DAEs associated. Nonetheless, if the degree of differentiability of the right-hand side $f(t)$ in the higher-order systems is limited, such transformation may be *nonequivalent* in the sense that there may not exist any continuous solution to the first-order system after transformation, whereas there exist continuous solutions to the original higher-order system. In Section 2.2, Chapter 2 we will give a definition of so-called *strangeness-index* and present an example to demonstrate this nonequivalence in terms of strangeness-index which can also be regarded as one of the key aspects in which DAEs differ from ODEs. The reason for the nonequivalence is due to the fact that systems of higher-order DAEs may essentially consist of not only ordinary differential equations in a classical sense, but also purely algebraic equations and further *strange parts* which arise from the couplings between differential and algebraic equations. Therefore, to get continuous solutions to the systems (1.1) and (1.2), introducing the derivative, the second derivative, ..., the $(l-1)$th derivative of the unknown vector-valued function $x(t)$ as a part of a new enlarged unknown may require more times continuous differentiability of the right-hand side $f(t)$ than that required in the original higher-order systems.

Observing the above nonequivalence, we see that it is not thoroughly satisfactory to convert higher-order systems of DAEs to first-order systems in order to solve them. Thus, the need to directly treat higher-order systems of DAEs provides a major motivation for our study. It is the aim of the first part of this thesis to directly investigate the mathematical structure of linear higher-order systems of DAEs and to lay a theoretical foundation for a better understanding of such systems.

The results of Chapters 2 and 3 are obtained mainly by a procedure of changing coordinates under equivalent transformations, differentiating the strange part of unknowns, and then eliminating through insertions the coupled strange part of the systems. Such techniques have been introduced and used by P. Kunkel and V. Mehrmann [28, 29], [34] (Chapters 2 and 3) to deal with linear *first-order* systems of DAEs, especially those with variable coefficients. The work in Chapters 2 and 3 is very close in spirit to the work done by P. Kunkel and V. Mehrmann, and the theory developed here is a natural extension of that of linear *first-order* systems to the systems of linear *higher-order* DAEs.

In outline, in Chapters 2 and 3 we first develop condensed forms, under appropriate equivalent transformations, for linear higher-order systems, whereupon we decouple the system concerned into ordinary-differential-equation part, 'strange' coupled differential-algebraic-equation part, and algebraic-equation part. Then we eliminate the strange part by differentiations and insertions, and repeat this process of decoupling and eliminating, until finally we transform the system into a so-called *strangeness-index zero* or *strangeness-free* normal form of the system of DAEs which has an equivalent solution set to that of the original system. Hence, based on the final normal form we can investigate the solution behaviour of the original system, and obtain results on solvability, uniqueness of solutions of the system, consistency of initial conditions (possibly given together with the system), and existence and uniqueness of solutions of the initial value problem possibly associated with the system. In such context, we will see that the major difference between the constant coefficient case and the variable coefficient case is that in the latter case, the variable coefficients must be sufficiently smooth and satisfy a set of *regularity conditions*, so that we can get the condensed form and pass the system through the inductive process to obtain the final normal form; (cf. Sections 3.1 and 3.2); whereas in the former case, the constant coefficients naturally satisfy such restrictive conditions.

Since the idea here is quite definite, for convenience and brevity of expression, our work in Chapters 2 and 3 will be mainly concentrated on the systems of linear *second-order* DAEs with constant and variable coefficients. The key results obtained for second-order systems can be extended without difficulty to linear high-order systems.

As we shall see in Sections 2.4 and 2.5, in the case of a constant coefficient system of DAEs, the solution behaviour of an initial value problem for the system of DAEs is closely related to the properties of regularity and singularity of the matrix polynomial associated with the system. This close relatedness provides us one of the major motivations to study *regularity* and *singularity* of matrix polynomials. This study will be conducted in the second part of this thesis, namely Chapter 4.

In Chapter 4 we shall study, from the point of view of the theory of matrices, regularity and singularity of $m \times n$ matrix polynomials of degree $l$

$$A(\lambda) = \sum_{i=0}^{l} \lambda^i A_i = \lambda^l A_l + \lambda^{l-1} A_{l-1} + \cdots + \lambda A_1 + A_0, \tag{1.7}$$

where $\lambda \in \mathbb{C}$ and the matrices $A_i \in \mathbb{C}^{m \times n}$, $i = 1, \ldots, l$. Here, we call a matrix polynomial $A(\lambda)$ *column-singular* (or *row-singular*, respectively) if $\operatorname{rank}(A(\lambda)) < n$ (or $\operatorname{rank}(A(\lambda)) < m$, respectively), otherwise it is *column-regular* (or *row-regular*, respectively).

Apart from the subject matter of DAEs mentioned as above, in the case of $m = n$, the study of polynomial eigenvalue problems (PEPs) provides another major motivation for our investigation (for theoretical and numerical analysis of PEPs, see, for example, [35, 17, 44, 46, 26, 41, 21, 56, 13, 57, 22, 23, 24, 9], and the references therein). In the monograph of Gohberg, Lancaster, and Rodman [17], a spectral theory for *regular* matrix polynomials has been developed. Nonetheless, for *singular* matrix polynomials, especially those of *degrees* greater than or equal to 2, the general theoretical analysis has been largely ignored. For this reason and the numerical concerns related, the second part of this thesis is aimed at theoretically analyzing characterizations of the column- and row- regularity and singularity of matrix polynomials, detecting the regularity or singularity of a given square matrix polynomial, and investigating the nearness to singularity problem for square matrix polynomials.

In Section 4.2, we shall prove sufficient and necessary conditions for the *singularity* of matrix polynomials, which means, put simply in the column case, that a matrix polynomial $A(\lambda)$ is column-singular if and only if there exists a vector polynomial $x(\lambda)$, which is not identically equal to zero, such that $A(\lambda)x(\lambda) = 0$. More essential, we shall also give an attainable upper bound on the least possible degree of such nonzero vector polynomials $x(\lambda)$. This main result of Section 4.2 will lead to other sufficient and necessary conditions, in terms of the matrix coefficients, for the *singularity* and *regularity* of matrix polynomials. In particular, as direct applications of such sufficient and necessary conditions, a geometrical characterization of singular matrix pencils and a canonical form, under equivalence transformations (defined by (2.17) in Section 2.3), for $2 \times 2$ *singular* quadratic matrix polynomials are presented.

Since, in many DAE problems and in all polynomial eigenvalue problems, *square* matrix polynomials are involved, we restrict the study conducted in Sections 4.3 and 4.4 to the square case. As we shall see in Section 2.5, in the setting of DAE problems (1.1), if the matrix polynomial associated is singular, then either the homogeneous initial value problem associated with (1.1) has a nontrivial solution, or there exist arbitrary smooth inhomogeneities $f(t)$ for which the system (1.1) is not solvable. While, in the setting of polynomial eigenvalue problems, if the matrix polynomial associated is singular, then it is immediate that every complex number can be regarded as an eigenvalue. Even if the matrix polynomial associated with a DAE problems or a polynomial eigenvalue problem is regular but *nearly* singular, a high sensitivity of the solutions of these problems to perturbations of the matrix coefficients may be expected. Therefore, in view of such singular or *nearly* singular phenomena, to detect the regularity or singularity of square matrix polynomials and to tackle the related nearness to singularity problem become very important from both the theoretical and numerical viewpoint.

Section 4.3 is devoted to detecting whether or not a given square matrix polynomial $A(\lambda)$ is regular. We shall present a natural approach to detection via rank information of the matrix coefficients and, if necessary, a finite number of computing determinants of matrices, where the finite number is an attainable upper bound on the number of nonzero roots of $\det(A(\lambda))$, if we suppose that $A(\lambda)$ is regular.

Section 4.4 deals with the nearness to singularity problem for square and regular matrix polynomials. First, we give an example to demonstrate a possibly high sensitivity of a regular but *nearly* singular polynomial eigenvalue problem if its matrix coefficients are perturbed. Then, we shall give a definition of nearness in terms of the spectral and Frobenius matrix norms and some properties of the distance to the nearest singular square matrix polynomials. Based on the sufficient and necessary conditions of the regularity of matrix polynomials proved in Section 4.2, we shall also present a general theoretical characterization of the nearest distance, which shows that the nearness problem is in fact a perturbation-structured and constrained rank-deficiency problem. In addition, on the basis of the result obtained in [4], we shall give a sharper characterization, in terms of singular values of matrices, of the nearness to singularity for matrix pencils, which coincides with a geometrical characterization for singular matrix pencils obtained in Section 4.2. Subsection 4.4.3 contains two special cases of matrix polynomials, for which we give explicit formulae for the nearest distance to singularity. At the end of Section 4.4, two types of lower bounds on the nearest distance to singularity for general regular matrix polynomials are also presented, which are generalizations of the results for matrix pencils obtained in [4].

Finally, in Chapter 5 we draw some conclusions and give an outlook for future work and investigations.

# Chapter 2

# Linear Higher-Order DAEs with Constant Coefficients

## 2.1 Introduction

In this chapter, we consider systems of linear $l$th-order ($l \geq 2$) differential-algebraic equations with constant coefficients of the form

$$A_l x^{(l)}(t) + A_{l-1} x^{(l-1)}(t) + \cdots + A_0 x(t) = f(t), \ t \in [t_0, t_1], \tag{2.1}$$

where $A_i \in \mathbb{C}^{m \times n}$, $i = 0, 1, \ldots, l$, $A_l \neq 0$, $f(t) \in \mathcal{C}^{\mu}([t_0, t_1], \mathbb{C}^m)$, possibly together with initial conditions

$$x(t_0) = x_0, \ \ldots, \ x^{(l-2)}(t_0) = x_0^{[l-2]}, \ x^{(l-1)}(t_0) = x_0^{[l-1]}, \quad x_0, \ldots, x_0^{[l-2]}, x_0^{[l-1]} \in \mathbb{C}^n. \tag{2.2}$$

Here, the nonnegative integer $\mu$ is the *strangeness-index* of the system (2.1), i.e., to get continuous solutions of the (2.1), the right-hand side $f(t)$ has to be $\mu$-times continuously differentiable (later, in Section 2.2 we shall give an explicit definition of the strangeness-index).

First, let us clarify the concepts of *solution* of the system (2.1), *solution* of the initial value problem (2.1)-(2.2), and *consistency* of the initial conditions (2.2).

**Definition 2.1** *A vector-valued function $x(t) := [x_1(t), \ldots, x_n(t)]^T \in \mathcal{C}([t_0, t_1], \mathbb{C}^n)$ is called* SOLUTION OF (2.1) *if $\sum_{k=1}^{n} A_i(j,k) x_k^{(i)}(t)$, $i = 0, \ldots, l$, $j = 1, \ldots, m$, exist and for $j = 1, \ldots, m$ the following equations are satisfied:*

$$\sum_{k=1}^{n} A_l(j,k) x_k^{(l)}(t) + \sum_{k=1}^{n} A_{l-1}(j,k) x_k^{(l-1)}(t) + \cdots + \sum_{k=1}^{n} A_0(j,k) x_k(t) = f_j(t),$$

*where $A_i(j,k)$ denotes the element of the matrix $A_i$ lying on the $j$th row and the $k$th column of $A_i$ and $f(t) := [f_1(t), \ldots, f_m(t)]^T$.*

*A vector-valued function $x(t) \in \mathcal{C}([t_0, t_1], \mathbb{C}^n)$ is called* SOLUTION OF THE INITIAL VALUE PROBLEM (2.1)-(2.2) *if it is a solution of* (2.1) *and, furthermore, satisfies* (2.2).

*Initial conditions* (2.2) *are called* CONSISTENT *with the system* (2.1) *if the associated initial value problem* (2.1)-(2.2) *has at least one solution.*

It should be noted that, since the system (2.1) of DAEs possibly contains purely algebraic equations, we are here interested in the weakest possible solution space $\mathcal{C}([t_0, t_1], \mathbb{C}^n)$, rather than $\mathcal{C}^l([t_0, t_1], \mathbb{C}^n)$. The differential operators $d^l/dt^l, d^{l-1}/dt^{l-1}, \ldots, d/dt$ in the system (2.1) are so far only symbols, which do not definitely mean that the unknown vector-valued function $x(t)$ should be $i$-times continuously differentiable, $i = l, l-1, \ldots, 1$. Later, in Section 2.2 we will see an example to demonstrate this point.

Based upon these concepts, we are naturally interested in the following questions:

1. Does the behaviour of the system (2.1) differ from that of a system of first-order DAEs into which (2.1) may be transformed in the same way as in the classical theory of ODEs?

2. Does the system (2.1) always have solutions? If it has, how many solutions do exist? Under which conditions does it have unique solutions?

3. If the system (2.1) has solutions, how smooth is the right-hand side $f(t)$ required to be?

4. Which conditions are required of consistent initial conditions?

5. Under which conditions does the initial value problem (2.1)-(2.2) have unique solutions?

In the following sections we shall answer the above questions one by one. In Section 2.2 we present an example to show the difference that may occur, in terms of *strangeness-index*, between the higher-order system (2.1) and a system of first-order DAEs into which the original system is converted. In Section 2.3 we shall give a condensed form, under *strong equivalence transformations*, for matrix triples that are associated with systems of second-order DAEs. Then, in Section 2.4, based on the condensed form, we partially read off the the properties of the corresponding system of second-order DAEs, and by *differentiation-and-elimination steps* reduce the system to a simpler but equivalent system. After an inductive procedure of this kind of reduction, we shall present a final equivalent *strangeness-free* system by which we can answer the questions posed in the above. Finally, in Section 2.5, the main results of second-order systems obtained in Section 2.4 are extended to general higher-order systems, and moreover, the connection between the solution behaviour of a system of DAEs and *regularity* or *singularity* of the matrix polynomial associated with the system is presented.

## 2.2 An Example

It is well known that one of the key aspects in which a system of DAEs differs from a system of ODEs is that, to get the solutions of DAEs, only continuity of the right-hand side $f(t)$ may not be sufficient and therefore higher derivatives of $f(t)$ may be required. Later, in Section 2.4, we will clearly see the reason for this difference. Furthermore, as we have mentioned in Chapter 1, another different point between higher-order DAEs and higher-order ODEs is that, in order to get continuous solutions, different degrees of differentiation of the right-hand side of the system of higher-order DAEs may be required than of the system of first-order DAEs into which the original higher-order system is converted; whereas in the case of ODEs the solution behaviour keeps completely invariant after such kind of conversion. To show this difference, we need the following definition of the sufficient and necessary degrees of differentiation required of the right-hand side of systems of DAEs, which is introduced in [28, 30, 34].

**Definition 2.2** *Provided that the system (2.1) has solutions, the minimum number $\mu$ of times that all or part of the right-hand side $f(t)$ in the system (2.1) must be differentiated in order to determine any solution $x(t)$ as a continuous function of $t$ is the* STRANGENESS-INDEX *of the system (2.1) of DAEs.*

Obviously, according to Definition 2.2, both a system of ODEs and a system of purely algebraic equations have a zero strangeness-index.

In the following, we present an example of an initial value problem for linear second-order DAEs to demonstrate the possible difference of strangeness index of the original system from that of the converted first-order system of DAEs.

**Example 2.3** We investigate the initial value problem for the linear second-order constant coefficient DAEs

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \ddot{x}(t) + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \dot{x}(t) + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x(t) = f(t), \quad t \in [t_0, t_1] \tag{2.3}$$

where $x(t) = [x_1(t), x_2(t)]^T$, and $f(t) = [f_1(t), f_2(t)]^T$ is sufficiently smooth, together with the initial conditions

$$x(t_0) = x_0, \quad \dot{x}(t_0) = x_0^{[1]}, \tag{2.4}$$

where $x_0 = [x_{01}, x_{02}]^T \in \mathbb{C}^2$, $x_0^{[1]} = [x_{01}^{[1]}, x_{02}^{[1]}]^T \in \mathbb{C}^2$. A short computation shows that system (2.3) has the unique solution

$$\begin{cases} x_1(t) = f_2(t), \\ x_2(t) = f_1(t) - \dot{f}_2(t) - \ddot{f}_2(t). \end{cases} \tag{2.5}$$

Moreover, (2.5) is the unique solution of the initial value problem (2.3)-(2.4) if the initial conditions (2.4) are consistent, namely,

$$\begin{cases} x_{01} &= f_2(t_0), \\ x_{02} &= f_1(t_0) - \dot{f}_2(t_0) - \ddot{f}_2(t_0), \\ x_{01}^{[1]} &= \dot{f}_2(t_0), \\ x_{02}^{[1]} &= \dot{f}_1(t_0) - \ddot{f}_2(t_0) - \left.\frac{d^3 f_2(t)}{dt^3}\right|_{t_0+}. \end{cases} \tag{2.6}$$

If we let

$$v(t) = [v_1(t), v_2(t)]^T = [\dot{x}_1(t), \dot{x}_2(t)]^T, \; y(t) = [v_1(t), v_2(t), x_1(t), x_2(t)]^T,$$

then we have the following initial-value problem for the linear first-order DAEs

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \dot{y}(t) + \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} y(t) = \begin{bmatrix} f_1(t) \\ f_2(t) \\ 0 \\ 0 \end{bmatrix}, \tag{2.7}$$

together with the initial condition

$$y(t_0) = [x_{01}^{[1]}, x_{02}^{[1]}, x_{01}, x_{02}]^T. \tag{2.8}$$

It is immediate that the system (2.7) of first-order DAEs has the unique solution

$$\begin{cases} x_1(t) = f_2(t), \\ x_2(t) = f_1(t) - \dot{f}_2(t) - \ddot{f}_2(t), \\ v_1(t) = \dot{f}_2(t), \\ v_2(t) = \dot{f}_1(t) - \ddot{f}_2(t) - f_2^{(3)}(t). \end{cases} \tag{2.9}$$

In this form, (2.9) is the unique solution of the initial value problem (2.7)-(2.8) if the initial condition (2.8) is consistent, i.e.,

$$\begin{cases} x_{01} &= f_2(t_0), \\ x_{02} &= f_1(t_0) - \dot{f}_2(t_0) - \ddot{f}_2(t_0), \\ x_{01}^{[1]} &= \dot{f}_2(t_0), \\ x_{01}^{[1]} &= \dot{f}_1(t_0) - \ddot{f}_2(t_0) - f_2^{(3)}(t_0). \end{cases} \tag{2.10}$$

**Remark 2.4** Example 2.3 shows that the second-order system (2.3) has a unique continuous solution (2.5) if and only if the right-hand side satisfies $f(t) \in \mathcal{C}^2([t_0, t_1], \mathbb{C}^2)$, whereas the converted first-order system (2.7) has a unique continuous solution if and only if $f(t) \in \mathcal{C}^3([t_0, t_1], \mathbb{C}^2)$; or in other words, the *strangeness-index* of the converted first-order system (2.7) is larger by one than that of the original second-order system (2.3). For a general system of $l$-th-order DAEs, it is not difficult to find similar examples.

Not unobviously, the reason for the above difference in terms of strangeness-index is that, in the converted first-order system the enlarged set of unknown functions includes not only the unknowns of the original higher-order system but also their derivatives, and that to get the solutions of such "new" unknowns, higher degree of the smoothness of the right-hand side function $f(t)$ may be required. Therefore, unlike the classical theory of ODEs (see, for example, E. A. Coddington and N. Levinson [10]), a direct transformation of a system of higher-order DAEs into an associated system of first-order DAEs is not always equivalent in the sense that higher degree of the smoothness of the right-hand side $f(t)$ may be involved in the solutions of the latter. $\Diamond$

It should be noted that Example 2.3 also shows that, to obtain continuous solutions of a system of DAEs, some parts of the right-hand side $f(t)$ may be required to be more differentiable than other parts which may be only required to be continuous; for a detailed investigation, we refer to, for example, [2, 37, 38]. Nonetheless, in order to simplify algebraic forms of a system of DAEs, we usually apply algebraic equivalence transformation to its matrix coefficients. For this reason and to avoid becoming too technical, we always consider the differentiability of the right-hand side vector-valued function $f(t)$ as a whole, and do not distinguish the degrees of smoothness required of its components from each other.

Now that a direct reduction of a system of higher-order DAEs to a first-order system may be nonequivalent in terms of the strangeness-index, we have to find another approach to enable us to investigate more deeply the behaviour of systems of higher-order DAEs. In the following sections, we will see that through purely algebraic techniques, like the treatment of systems of linear first-order DAEs with constant coefficients, we can get a thorough understanding of the behaviour of systems of linear higher-order constant coefficient DAEs.

## 2.3 Condensed Form for Matrix Triples

As we have mentioned in Chapter 1, for convenience of notation and expression, in this section we shall work mainly with systems of linear second-order DAEs with constant coefficients

$$M\ddot{x}(t) + C\dot{x}(t) + Kx(t) = f(t),\ t \in [t_0, t_1], \tag{2.11}$$

with $M, C, K \in \mathbb{C}^{m\times n},\ f(t) \in \mathcal{C}^{\mu}([t_0, t_1], \mathbb{C}^m)$, possibly together with initial conditions

$$x(t_0) = x_0,\ \dot{x}(t_0) = x_0^{[1]}, \quad x_0, x_0^{[1]} \in \mathbb{C}^n. \tag{2.12}$$

It is well-known that the nature of the solutions of the system of linear first-order constant coefficient DAEs

$$E\dot{x}(t) = Ax(t) + f(t),\ t \in [t_0, t_1],$$

with $E, A \in \mathbb{C}^{m\times n}$ and $f(t) \in \mathcal{C}^{\mu}([t_0, t_1], \mathbb{C}^m)$, can be determined by the properties of the corresponding matrix pencil $\lambda E - A$. Furthermore, the algebraic properties of the matrix pencil $\lambda E - A$ can be well understood through studying the canonical forms for the set of matrix pencils

$$\lambda(PEQ) - (PAQ), \tag{2.13}$$

where $P \in \mathbb{C}^{m\times m}, Q \in \mathbb{C}^{n\times n}$ are any nonsingular matrices; see, for example, [3] (Section 2.3) and [34] (Section 2.1). In particular, among those canonical forms for (2.13) are the well-known *Weierstrass canonical form* for regular matrix pencils ([55], Chapter VI) and the *Kronecker canonical form* for general singular matrix pencils ([15], Chapter XII), from which one can directly read off the solution properties of the corresponding DAEs.

Similarly, as we will see later in this chapter, the behaviour of solutions of the system (2.11), as well as the initial value problem (2.11)-(2.12), depends on the properties of the quadratic matrix polynomial

$$A(\lambda) = \lambda^2 M + \lambda C + K. \tag{2.14}$$

If we let $x(t) = Qy(t)$, and premultiply (2.11) by $P$, where $P \in \mathbb{C}^{m\times m}, Q \in \mathbb{C}^{n\times n}$ are nonsingular matrices, we obtain an *equivalent* system of DAEs

$$(PMQ)\ddot{y}(t) + (PCQ)\dot{y}(t) + (PKQ)y(t) = Pf(t), \tag{2.15}$$

and a new corresponding quadratic matrix polynomial

$$\hat{A}(\lambda) = \lambda^2 \hat{M} + \lambda \hat{C} + \hat{K} := \lambda^2(PMQ) + \lambda(PCQ) + (PKQ). \tag{2.16}$$

Here, by *equivalence* we mean not only that the relation $x(t) = Qy(t)$ (or $y(t) = Q^{-1}x(t)$) gives a one-to-one correspondence between the two corresponding solution sets of the system (2.11) and the system (2.15), but also that, in order to get continuous solutions of the systems (2.15) and (2.11), the smoothness conditions required of $Pf(t)$ in (2.15) are equal to those required of the right-hand side $f(t)$ in (2.11).

However, it is also well-known that it is an open problem to find a canonical form for quadratic matrix polynomials (2.16), let alone higher-degree matrix polynomials, from which we can *directly* read off the solution properties of the corresponding system of DAEs. Nonetheless, inspired by the work of [28, 29] (though the papers mainly deal with linear first-order DAEs with *variable coefficients*), we shall in this section give an equivalent condensed form for quadratic matrix polynomials (2.14) through purely algebraic manipulations and coordinate changes. Based on the condensed form we can *partially* decouple the system into three parts, namely, an ordinary-differential-equation part, an algebraic part and a coupling part, and therefore pave the way for the further treatment of the system in the following section.

Sometimes, we will use the notation $(A_l, \ldots, A_1, A_0)$ of a matrix $(l+1)$-tuple instead of the matrix polynomial $\lambda^l A_l + \cdots + \lambda A_1 + A_0$ of $l$th degree which is associated with the general $l$th-order system (2.1) of DAEs. By the following definition, we make the concept of *equivalence* between two general matrix $(l+1)$-tuples clear.

**Definition 2.5** *Two $(l+1)$-tuples $(A_l, \ldots, A_1, A_0)$ and $(B_l, \ldots, B_1, B_0)$, $A_i, B_i \in \mathbb{C}^{m\times n}$, $i = 0, 1, \ldots, l$, $l \in \mathbb{N}_0$, of matrices are called* (STRONGLY) EQUIVALENT *if there are nonsingular matrices $P \in \mathbb{C}^{m\times m}$ and $Q \in \mathbb{C}^{n\times n}$ such that*

$$B_i = PA_iQ, \quad i = 0, 1, \ldots, l. \tag{2.17}$$

*If this is the case, we write $(A_l, \ldots, A_1, A_0) \sim (B_l, \ldots, B_1, B_0)$.*

It is obvious that relation (2.17) is an equivalence relation, in other words, it is reflexive, symmetric, and transitive. In the remainder of this section we shall look for a condensed form for matrix triples under the equivalence relation (2.17). Before embarking on this, let us first review the canonical form for a matrix and a condensed form for a matrix pair under the equivalence relation (2.17).

The result on the canonical form for a single matrix under equivalence relation (2.17) is well-known:

**Lemma 2.6** *([36], p. 51) Let $A \in \mathbb{C}^{m\times n}$. Then there are nonsingular matrices $P := \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} \in \mathbb{C}^{m\times m}$ and $Q := [Q_1, Q_2] \in \mathbb{C}^{n\times n}$ such that*

$$PAQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}, \tag{2.18}$$

*where $P_1 \in \mathbb{C}^{r\times m}$, $Q_1 \in \mathbb{C}^{n\times r}$. Moreover, we have*

$$r = \operatorname{rank}(A),\ \mathcal{N}(A) = \mathcal{R}(Q_2),\ \mathcal{N}(A^T) = \mathcal{R}(P_2^T), \tag{2.19}$$

*where $\mathcal{N}(\cdot)$ denotes the null space of a matrix, and $\mathcal{R}(\cdot)$ the column space of a matrix.*

The condensed form for a matrix pair $(E, A)$ under equivalence relation (2.17) has been implicitly given in [28].

**Lemma 2.7** *Let $E, A \in \mathbb{C}^{m\times n}$, and let*

$$\begin{array}{lll} (a) & Z_1 \in \mathbb{C}^{m\times(m-r)} & \textit{be a matrix whose columns form a basis for } \mathcal{N}(E^T) \\ (b) & Z_2 \in \mathbb{C}^{n\times(n-r)} & \textit{be a matrix whose columns form a basis for } \mathcal{N}(E). \end{array} \tag{2.20}$$

*Then, the matrix pair $(E, A)$ is equivalent to a matrix pair of the form*

$$\left( \begin{bmatrix} I_s & 0 & 0 & 0 \\ 0 & I_d & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & A_{12} & 0 & A_{14} \\ 0 & A_{22} & 0 & A_{24} \\ 0 & 0 & I_a & 0 \\ I_s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right) \begin{array}{l} s \\ d \\ a \\ s \\ v \end{array}, \tag{2.21}$$

*where $s, d, a, v \in \mathbb{N}_0$, $A_{14} \in \mathbb{C}^{s\times u}$, $u \in \mathbb{N}_0$, and the quantities (in the following we use the convention* $\operatorname{rank}(0) = 0$*)*

$$\begin{array}{ll} (a) & r = \operatorname{rank}(E) \\ (b) & a = \operatorname{rank}(Z_1^T A Z_2) \\ (c) & s = \operatorname{rank}(Z_1^T A) - a \\ (d) & d = r - s \\ (e) & v = m - r - a - s \\ (f) & u = n - r - a \end{array} \tag{2.22}$$

*are invariant under equivalence relation* (2.17)*.*

For completeness, we give a proof of this lemma.

**Proof of Lemma 2.7**. In the following, the word "new" on top of the equivalence operator denotes that the subscripts of the entries are adapted to the new block structure of the matrices. Using Lemma 2.6, we obtain the following sequence of equivalent matrix pairs.

$$(E, A) \sim \left( \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \right) \overset{\text{new}}{\sim} \left( \begin{bmatrix} I_r & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & I_a & 0 \\ A_{31} & 0 & 0 \end{bmatrix} \right)$$

$$\overset{\text{new}}{\sim} \left( \begin{bmatrix} I_r & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ 0 & I_a & 0 \\ A_{31} & 0 & 0 \end{bmatrix} \right) \sim \left( \begin{bmatrix} I_r & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} A_{11} & 0 & A_{13} \\ 0 & I_a & 0 \\ A_{31} & 0 & 0 \end{bmatrix} \right)$$

$$\overset{\text{new}}{\sim}\left(\begin{bmatrix} P_{11} & P_{12} & 0 & 0 \\ P_{21} & P_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} A_{11} & A_{12} & 0 & A_{14} \\ A_{21} & A_{22} & 0 & A_{24} \\ 0 & 0 & I_a & 0 \\ I_s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right)$$

$$\left(\text{where the matrix } \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \text{ is nonsingular}\right)$$

$$\overset{\text{new}}{\sim}\left(\begin{bmatrix} I_s & 0 & 0 & 0 \\ 0 & I_d & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & A_{12} & 0 & A_{14} \\ 0 & A_{22} & 0 & A_{24} \\ 0 & 0 & I_a & 0 \\ I_s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right).$$

It remains to show that such quantities $r, s, d, a, v, u$ are well-defined by (2.22) and invariant under the equivalence relation (2.17). In the case of $r = \text{rank}(E)$, this is clear. For the other quantities, indeed, we only need to show two quantities $a$ and $s$ are well-defined and invariant under equivalence relation (2.17). Since we have proved (2.21), let $P := \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} \in \mathbb{C}^{m\times m}$ and $Q := [Q_1, Q_2] \in \mathbb{C}^{n\times n}$ be nonsingular matrices, where $P_1 \in \mathbb{C}^{r\times m}$, $Q_1 \in \mathbb{C}^{n\times r}$, such that

$$\begin{bmatrix} P_1 \\ P_2 \end{bmatrix} E[Q_1, Q_2] = \begin{bmatrix} I_s & 0 & 0 & 0 \\ 0 & I_d & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} A[Q_1, Q_2] = \begin{bmatrix} 0 & A_{12} & 0 & A_{14} \\ 0 & A_{22} & 0 & A_{24} \\ 0 & 0 & I_a & 0 \\ I_s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{2.23}$$

By Lemma 2.6, we have

$$\mathcal{N}(E^T) = \mathcal{R}(P_2^T), \quad \mathcal{N}(E) = \mathcal{R}(Q_2), \tag{2.24}$$

namely, the columns of $P_2^T$ span $\mathcal{N}(E^T)$, and the columns of $Q_2$ span $\mathcal{N}(E)$. From (2.23) it immediately follows that

$$P_2AQ = \begin{bmatrix} 0 & 0 & I_a & 0 & 0 \\ I_s & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad P_2AQ_2 = \begin{bmatrix} I_a & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{2.25}$$

Hence, by (2.25), we have

$$a = \text{rank}\,(P_2AQ_2), \quad s = \text{rank}\,(P_2AQ) - a = \text{rank}\,(P_2A) - a. \tag{2.26}$$

From (2.20) and (2.24) it follows that there exist nonsingular matrices $T_1 \in \mathbb{C}^{(m-r)\times(m-r)}$ and $T_2 \in \mathbb{C}^{(n-r)\times(n-r)}$ such that

$$P_2^T = Z_1 T_1, \quad Q_2 = Z_2 T_2. \tag{2.27}$$

Then, from (2.26) and (2.27) it follows that

$$\begin{aligned} a &= \operatorname{rank}(P_2 A Q_2) = \operatorname{rank}\left(T_1^T Z_1^T A Z_2 T_2\right) = \operatorname{rank}(Z_1^T A Z_2), \quad \text{and} \\ s &= \operatorname{rank}(P_2 A) - a = \operatorname{rank}\left(T_1^T Z_1^T A\right) - a = \operatorname{rank}(Z_1^T A) - a. \end{aligned}$$

Thus, $a$ and $s$ are indeed well-defined by (2.22) and therefore so are the quantities $d$, $v$ and $u$.

At last, we shall prove that $a$ and $s$ are invariant under the equivalence relation. Let $(E_i, A_i)$, $i = 1, 2$, be equivalent, and let $Z_1^{(i)}, Z_2^{(i)}$ be bases associated with $(E_i, A_i)$, $i = 1, 2$, i.e., let

$(a)$ $Z_1^{(i)}$ be a matrix whose columns form a basis for $\mathcal{N}(E_i^T)$
$(b)$ $Z_2^{(i)}$ be a matrix whose columns form a basis for $\mathcal{N}(E_i)$.

Since there exist nonsingular matrices $P \in \mathbb{C}^{m\times m}$ and $Q \in \mathbb{C}^{n\times n}$ such that $E_1 = PE_2Q$ and $A_1 = PA_2Q$, from $E_1^T Z_1^{(1)} = 0$ and $E_1 Z_2^{(1)} = 0$ it follows that

$$Q^T E_2^T P^T Z_1^{(1)} = 0, \quad PE_2QZ_2^{(1)} = 0,$$

and therefore

$$E_2^T P^T Z_1^{(1)} = 0, \quad E_2QZ_2^{(1)} = 0.$$

Thus, the columns of $P^T Z_1^{(1)}$ from a basis for $\mathcal{N}(E_2^T)$ and the columns of $QZ_2^{(1)}$ from a basis for $\mathcal{N}(E_2)$. Therefore, there exist nonsingular matrices $\hat{T}_1 \in \mathbb{C}^{(m-r)\times(m-r)}$ and $\hat{T}_2 \in \mathbb{C}^{(n-r)\times(n-r)}$ such that

$$P^T Z_1^{(1)} = Z_1^{(2)} \hat{T}_1, \quad QZ_2^{(1)} = Z_2^{(2)} \hat{T}_2.$$

Then, we can complete the proof of the invariance of $a$ and $s$ with the fact that

$$\begin{aligned} \operatorname{rank}({Z_1^{(1)}}^T A_1) &= \operatorname{rank}({Z_1^{(2)}}^T P^{-1} P A_2 Q) = \operatorname{rank}({Z_1^{(2)}}^T A_2 Q) = \operatorname{rank}({Z_1^{(2)}}^T A_2), \\ \operatorname{rank}({Z_1^{(1)}}^T A_1 Z_2^{(1)}) &= \operatorname{rank}({Z_1^{(2)}}^T P^{-1} P A_2 Q Q^{-1} Z_2^{(2)} \hat{T}_2) = \operatorname{rank}({Z_1^{(2)}}^T A_2 Z_2^{(2)}). \end{aligned}$$

□

Using the condensed form (2.21) for matrix pairs and similar algebraic techniques utilized in the proof of Lemma 2.7, we can then derive a condensed form for matrix triples, which is the main result of this section.

**Lemma 2.8** *Let $M, C, K \in \mathbb{C}^{m \times n}$. Then, $(M, C, K)$ is equivalent to a matrix triple $(\hat{M}, \hat{C}, \hat{K})$ of the following form*

$$
\left( \begin{bmatrix}
I_{s^{(MCK)}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & I_{s^{(MC)}} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & I_{s^{(MK)}} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & I_{d^{(2)}} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}, \begin{bmatrix}
0 & 0 & C & C & 0 & 0 & C & C \\
0 & 0 & C & C & 0 & 0 & C & C \\
0 & 0 & C & C & 0 & 0 & C & C \\
0 & 0 & C & C & 0 & 0 & C & C \\
0 & 0 & 0 & C & I_{s^{(CK)}} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & I_{d^{(1)}} & 0 & 0 \\
I_{s^{(MCK)}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & I_{s^{(MC)}} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}, \begin{bmatrix}
0 & K & 0 & K & 0 & K & 0 & K \\
0 & K & 0 & K & 0 & K & 0 & K \\
0 & K & 0 & K & 0 & K & 0 & K \\
0 & K & 0 & K & 0 & K & 0 & K \\
0 & K & 0 & K & 0 & K & 0 & K \\
0 & K & 0 & K & 0 & K & 0 & K \\
0 & K & 0 & K & 0 & K & 0 & K \\
0 & K & 0 & K & 0 & K & 0 & K \\
0 & 0 & 0 & 0 & 0 & 0 & I_a & 0 \\
0 & 0 & 0 & 0 & I_{s^{(CK)}} & 0 & 0 & 0 \\
0 & 0 & I_{s^{(MK)}} & 0 & 0 & 0 & 0 & 0 \\
I_{s^{(MCK)}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix} \right) \begin{array}{l}
s^{(MCK)} \\ s^{(MC)} \\ s^{(MK)} \\ d^{(2)} \\ s^{(CK)} \\ d^{(1)} \\ s^{(MCK)} \\ s^{(MC)} \\ a \\ s^{(CK)} \\ s^{(MK)} \\ s^{(MCK)} \\ v
\end{array}, \tag{2.28}
$$

*where the quantities $s^{(MCK)}$, $s^{(MC)}$, $s^{(MK)}$, $s^{(CK)}$, $d^{(2)}$, $d^{(1)}$, $a$ and $v$ are nonnegative integers.*

It should be noted that, for convenience of expression, in this lemma and in the following proof we drop the subscripts of the elements of block matrices unless they are needed for clarification.

**Proof of Lemma 2.8**. The proof in the following will proceed from the condensed form for matrix pair $(M, C)$ obtained in Lemma 2.7. And, the canonical form (2.18) for matrices shown in Lemma 2.6 will be repeatedly employed in the course of the proof.

$$
\begin{aligned}
&(M, C, K) \\
&\sim \left( \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & C \\ 0 & C & 0 & C \\ 0 & 0 & I & 0 \\ I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & K \\ K & K & K & K \\ K & K & K & K \\ K & K & K & K \\ K & K & K & K \end{bmatrix} \right) \\
&\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & C & C \\ 0 & C & 0 & C & C \\ 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & K & K \\ K & K & K & K & K \\ K & K & K & K & K \\ K & K & K & K & K \\ K & K & K & I & 0 \\ K & K & K & 0 & 0 \end{bmatrix} \right)
\end{aligned}
$$

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & C & C \\ 0 & C & 0 & C & C \\ 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & K & I & 0 \\ K & K & K & 0 & 0 \end{bmatrix} \right)$$

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} C & C & C & C & C \\ C & C & C & C & C \\ 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & K & 0 & K \\ 0 & 0 & 0 & I & 0 \\ K & K & K & 0 & 0 \end{bmatrix} \right)$$

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & C & C \\ 0 & C & 0 & C & C \\ 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & K & 0 & K \\ 0 & 0 & 0 & I & 0 \\ K & K & K & 0 & 0 \end{bmatrix} \right)$$

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & 0 & C & C \\ 0 & C & 0 & 0 & C & C \\ 0 & 0 & P_{11} & P_{12} & 0 & 0 \\ 0 & 0 & P_{21} & P_{22} & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & K & 0 & K \\ K & K & K & K & 0 & K \\ K & K & K & K & 0 & K \\ K & K & K & K & 0 & K \\ K & K & K & K & 0 & K \\ 0 & 0 & 0 & 0 & I & 0 \\ K & K & I & 0 & 0 & 0 \\ K & K & 0 & 0 & 0 & 0 \end{bmatrix} \right)$$

(where the matrix $\begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$ is nonsingular)

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & 0 & C & C \\ 0 & C & 0 & 0 & C & C \\ 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & I & 0 \\ K & K & I & 0 & 0 & 0 \\ K & K & 0 & 0 & 0 & 0 \end{bmatrix} \right)$$

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & 0 & C & C \\ 0 & C & 0 & 0 & C & C \\ C & C & I & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & I & 0 & 0 & 0 \\ K & K & 0 & 0 & 0 & 0 \end{bmatrix} \right)$$

$$
\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & 0 & C & C \\ 0 & C & 0 & 0 & C & C \\ 0 & C & I & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & I & 0 & 0 & 0 \\ K & K & 0 & 0 & 0 & 0 \end{bmatrix} \right)
$$

$$
\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & Q_{11} & Q_{12} & 0 & 0 & 0 & 0 \\ 0 & Q_{21} & Q_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & C & 0 & 0 & C & C \\ 0 & C & C & 0 & 0 & C & C \\ 0 & C & C & 0 & 0 & C & C \\ 0 & C & C & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & 0 & K & 0 & K \\ K & K & K & 0 & K & 0 & K \\ K & K & K & 0 & K & 0 & K \\ K & K & K & 0 & K & 0 & K \\ K & K & K & 0 & K & 0 & K \\ K & K & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 \\ K & I & 0 & 0 & 0 & 0 & 0 \\ K & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right)
$$

(where the matrix $\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}$ is nonsingular)

$$
\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & C & 0 & 0 & C & C \\ 0 & C & C & 0 & 0 & C & C \\ 0 & C & C & 0 & 0 & C & C \\ 0 & C & C & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 \\ K & I & 0 & 0 & 0 & 0 & 0 \\ K & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right)
$$

$$
\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 & 0 \\ M & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} C & C & C & 0 & 0 & C & C \\ C & C & C & 0 & 0 & C & C \\ C & C & C & 0 & 0 & C & C \\ C & C & C & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 \\ K & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right)
$$

$$
\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & C & 0 & 0 & C & C \\ 0 & C & C & 0 & 0 & C & C \\ 0 & C & C & 0 & 0 & C & C \\ 0 & C & C & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 \\ K & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right)
$$

$$\sim \left( \begin{bmatrix} R_{11} & R_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ R_{21} & R_{22} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I & 0 & 0 \\ R_{11} & R_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ R_{21} & R_{22} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & 0 & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right)$$

(where the matrix $\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$ is nonsingular)

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & 0 & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right)$$

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right)$$

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & 0 & C & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & K & 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right)$$

$$
\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & 0 & C & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right). \qquad \square
$$

Each of the quantities in Lemma 2.8 has an expression in terms of dimension of column spaces or rank of matrices, and is invariant under equivalence relation (2.17), as the next lemma shows.

**Lemma 2.9** *Let $M, C, K \in \mathbb{C}^{m\times n}$ and*

$$
\begin{array}{lll}
(a) & Z_1 & \textit{be a matrix whose columns form a basis for } \mathcal{N}(M^T) \\
(b) & Z_2 & \textit{be a matrix whose columns form a basis for } \mathcal{N}(M) \\
(c) & Z_3 & \textit{be a matrix whose columns form a basis for } \mathcal{N}(M^T)\cap\mathcal{N}(C^T) \\
(d) & Z_4 & \textit{be a matrix whose columns form a basis for } \mathcal{N}(M)\cap\mathcal{N}(Z_1^T C)
\end{array}
\tag{2.29}
$$

*Then the quantities*

$$
\begin{array}{lll}
(a) & r = \operatorname{rank}(M) & \text{(rank of } M\text{)} \\
(b) & a = \operatorname{rank}(Z_3^T K Z_4) & \text{(algebraic part)} \\
(c) & s^{(MCK)} = \dim\left(\mathcal{R}(M^T)\cap\mathcal{R}(C^T Z_1)\cap\mathcal{R}(K^T Z_3)\right) & \text{(strangeness due to } M, C, K\text{)} \\
(d) & s^{(CK)} = \operatorname{rank}(Z_3^T K Z_2) - a & \text{(strangeness due to } C, K\text{)} \\
(e) & d^{(1)} = \operatorname{rank}(Z_1^T C Z_2) - s^{(CK)} & \text{(1st-order differential part)} \\
(f) & s^{(MC)} = \operatorname{rank}(Z_1^T C) - s^{(MCK)} - s^{(CK)} - d^{(1)} & \text{(strangeness due to } M, C\text{)} \\
(g) & s^{(MK)} = \operatorname{rank}(Z_3^T K) - a - s^{(MCK)} - s^{(CK)} & \text{(strangeness due to } M, K\text{)} \\
(h) & d^{(2)} = r - s^{(MCK)} - s^{(MC)} - s^{(MK)} & \text{(2nd-order differential part)} \\
(i) & v = m - r - 2s^{(CK)} - d^{(1)} - 2s^{(MCK)} - s^{(MC)} - a - s^{(MK)} & \text{(vanishing equations)} \\
(j) & u = n - r - s^{(CK)} - d^{(1)} - a & \text{(undetermined part)}
\end{array}
\tag{2.30}
$$

*are invariant under the strong equivalence relation* (2.17) *and* $(M, C, K)$ *is (strongly) equivalent to the condensed form* (2.28).

It should be pointed out that the meanings of the invariants indicated in the parentheses in (2.30) stem from the context of the system (2.11) of DAEs, which we will explain in the next section. Lemma 2.9 will be used in the next chapter.

**Proof of Lemma 2.9**. The proof can be carried out along the same lines of the proof of Lemma 2.7.

*Step 1.* First, we show that the quantities in (2.30) are well-defined with respect to the choices of the bases in (2.29). We take $a = \operatorname{rank}(Z_3^T K Z_4)$ as an example. Every change of basis can be represented by

$$\tilde{Z}_3 = Z_3 Q_1, \quad \tilde{Z}_4 = Z_4 Q_2$$

with nonsingular matrices $Q_1$, $Q_2$. From

$$\operatorname{rank}(\tilde{Z}_3^T K \tilde{Z}_4) = \operatorname{rank}(Q_1^T Z_3^T K Z_4 Q_2) = \operatorname{rank}(Z_3^T K Z_4),$$

it then follows that $\operatorname{rank}(Z_1^T C Z_2)$ is well-defined. Similarly, we can prove that the other quantities in (2.30) are also well–defined.

*Step 2.* Next, we show that the quantities in (2.30) are invariant under the equivalence relation (2.17). Here, we just take $s^{(MCK)}$ as an example. Let $(M, C, K)$ and $(\tilde{M}, \tilde{C}, \tilde{K})$ be equivalent, namely, there are nonsingular matrices $P$ and $Q$ such that

$$\tilde{M} = PMQ, \quad \tilde{C} = PCQ, \quad \tilde{K} = PKQ. \tag{2.31}$$

Let the columns of $\tilde{Z}_1$ form a basis for $\mathcal{N}(\tilde{M}^T)$, and let the columns of $\tilde{Z}_3$ form a basis for $\mathcal{N}(\tilde{M}^T) \cap \mathcal{N}(\tilde{C}^T)$. Then, from (2.31) it follows that the columns of $Z_1 := P^T \tilde{Z}_1$ form a basis for $\mathcal{N}(M^T)$, and the columns of $Z_3 := P^T \tilde{Z}_3$ form a basis for $\mathcal{N}(M^T) \cap \mathcal{N}(C^T)$. Thus, the invariance of $s^{(MCK)}$ follows from

$$\begin{aligned}
\tilde{s}^{(MCK)} &= \dim\left(\mathcal{R}(\tilde{M}^T) \cap \mathcal{R}(\tilde{C}^T \tilde{Z}_1) \cap \mathcal{R}(\tilde{K}^T \tilde{Z}_3)\right) \\
&= \dim\left(\mathcal{R}(Q^T M^T P^T) \cap \mathcal{R}(Q^T C^T P^T \tilde{Z}_1) \cap \mathcal{R}(Q^T K^T P^T \tilde{Z}_3)\right) \\
&= \dim\left(\mathcal{R}(M^T P^T) \cap \mathcal{R}(C^T P^T \tilde{Z}_1) \cap \mathcal{R}(K^T P^T \tilde{Z}_3)\right) \\
&= \dim\left(\mathcal{R}(M^T) \cap \mathcal{R}(C^T Z_1) \cap \mathcal{R}(K^T Z_3)\right) \\
&= s^{(MCK)}.
\end{aligned}$$

Similarly, the invariance of the other quantities in (2.30) can be proved.

*Step 3.* Finally, we shall show that the quantities in the equivalent form (2.28) of $(M, C, K)$ are identical with those defined in (2.30). Let $P \in \mathbb{C}^{m \times m}$, $Q \in \mathbb{C}^{n \times n}$ be nonsingular matrices such that

$$(\hat{M}, \hat{C}, \hat{K}) = (PMQ, PCQ, PKQ)$$

where $(\hat{M}, \hat{C}, \hat{K})$ is of the form (2.28). Furthermore, let $P$ and $Q$ be partitioned as $P := \left[P_1^T, P_2^T, \ldots, P_{13}^T\right]^T$ and $Q := [Q_1, Q_2, \ldots, Q_8]$ conformally with the row structure and column structure of the block matrices in (2.28), respectively. Then, by (2.28), we have

$$\begin{aligned}
\left[P_5^T, \ldots, P_{13}^T\right]^T M &= 0, \\
M\,[Q_5, \ldots, Q_8] &= 0, \\
\left[P_9^T, \ldots, P_{13}^T\right]^T C &= 0, \\
\left[P_5^T, \ldots, P_{13}^T\right]^T C\,[Q_7, Q_8] &= 0,
\end{aligned}$$

namely, the columns of $P^T(:,5:13) := \left[P_5^T, \dots, P_{13}^T\right]$ form a basis for $\mathcal{N}(M^T)$, the columns of $Q(:,5:8) := [Q_5, \dots, Q_8]$ form a basis for $\mathcal{N}(M)$, the columns of $P^T(:,9:13) := \left[P_9^T, \dots, P_{13}^T\right]$ form a basis for $\mathcal{N}(M^T) \cap \mathcal{N}(C^T)$, and the columns of $[Q_7, Q_8]$ form a basis for $\mathcal{N}(M) \cap \mathcal{N}((P^T(:,5:13))^T C)$. Observing that, by (2.28),

$$\left(P^T(:,9:13)\right)^T K\,[Q_7, Q_8] = \begin{bmatrix} P_9 \\ \vdots \\ P_{13} \end{bmatrix} K\,[Q_7, Q_8] = \begin{bmatrix} I_a & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

we have $a = \operatorname{rank}\left(\left(P^T(:,9:13)\right)^T K\,[Q_7, Q_8]\right)$ which is equal to $a = \operatorname{rank}(Z_3^T K Z_4)$ defined in (2.30) (since $a$ is well-defined). Similarly, we can prove that the other quantities in the equivalent form (2.28) are equal to those defined in (2.30). $\square$

**Remark 2.10** Using the same techniques developed in Lemmas 2.7, 2.8, and 2.9, we can see that for $(l+1)$-tuples $(A_l, \dots, A_1, A_0)$ of matrices of size $m \times n$, there also exists a similar kind of condensed form via strong equivalence transformations, with which a set of invariant quantities are associated. For the sake of convenience of expression, we do not here explicitly present the condensed form for $(l+1)$-tuples of matrices. $\diamondsuit$

Thus, we have prepared the way for further analyzing the systems (2.11) and (2.1) of DAEs, which will be presented in the next two sections.

## 2.4 Linear Second-Order DAEs with Constant Coefficients

In this section, we discuss the system (2.11) of DAEs, and answer those questions raised at the beginning of this chapter.

Let us start by writing down the system of differential-algebraic equations

$$\hat{M}\ddot{y}(t) + \hat{C}\dot{y}(t) + \hat{K}y(t) = \hat{f}(t) \tag{2.32}$$

where

$$\hat{M} = PMQ,\ \hat{C} = PCQ,\ \hat{K} = PKQ,\ x(t) = Qy(t),\ \hat{f}(t) = Pf(t) \tag{2.33}$$

with $P, Q$ nonsingular matrices, and the matrix triple $(\hat{M}, \hat{C}, \hat{K})$ is of the condensed form (2.28), as follows.

$$\begin{array}{ll}
(a) & \ddot{y}_1(t) + \sum_{i=3,4,7,8} C_{1,i}\dot{y}_i(t) + \sum_{i=2,4,6,8} K_{1,i}y_i(t) = \hat{f}_1(t) \\
(b) & \ddot{y}_2(t) + \sum_{i=3,4,7,8} C_{2,i}\dot{y}_i(t) + \sum_{i=2,4,6,8} K_{2,i}y_i(t) = \hat{f}_2(t) \\
(c) & \ddot{y}_3(t) + \sum_{i=3,4,7,8} C_{3,i}\dot{y}_i(t) + \sum_{i=2,4,6,8} \hat{K}K_{3,i}y_i(t) = \hat{f}_3(t) \\
(d) & \ddot{y}_4(t) + \sum_{i=3,4,7,8} C_{4,i}\dot{y}_i(t) + \sum_{i=2,4,6,8} K_{4,i}y_i(t) = \hat{f}_4(t) \\
(e) & C_{5,4}\dot{y}_4(t) + \dot{y}_5(t) + \sum_{i=2,4,6,8} K_{5,i}y_i(t) = \hat{f}_5(t) \\
(f) & \dot{y}_6(t) + \sum_{i=2,4,6,8} K_{6,i}y_i(t) = \hat{f}_6(t) \\
(g) & \dot{y}_1(t) + \sum_{i=2,4,6,8} K_{7,i}y_i(t) = \hat{f}_7(t) \\
(h) & \dot{y}_2(t) + \sum_{i=2,4,6,8} K_{8,i}y_i(t) = \hat{f}_8(t) \\
(i) & y_7(t) = \hat{f}_9(t) \\
(j) & y_5(t) = \hat{f}_{10}(t) \\
(k) & y_3(t) = \hat{f}_{11}(t) \\
(l) & y_1(t) = \hat{f}_{12}(t) \\
(m) & 0 = \hat{f}_{13}(t).
\end{array} \tag{2.34}$$

Immediately, we recognize a consistency condition (2.34-m) for the inhomogeneity (vanishing equations) and a possible free condition of choice in $y_8(t)$ (undetermined unknowns). In addition, (2.34-i) may include purely algebraic equations (algebraic part), and, (2.34-f) (1st-order differential part) and (2.34-d) (2nd-order differential part) look like first-order differential equations and second-order differential equations, respectively. What is more essential to DAEs is the coupling (strangeness due to $M, C, K$) between the algebraic equations (2.34-l) and the differential equations (2.34-g) and (2.34-a), the coupling (strangeness due to $M, K$) between the algebraic equations (2.34-k) and the differential equations (2.34-c), the coupling (strangeness due to $C, K$) between the algebraic equations (2.34-j) and the differential equations (2.34-e), the coupling (strangeness due to $M, C$) between the differential equations (2.34-h) and the differential equations (2.34-b), and the possible coupling between the algebraic equations (2.34-i) and the differential equations (2.34-a) - (2.34-d).

Our direct objective now is the reduction of the system (2.34) to a simpler but equivalent system by means of decoupling those equations coupled to each other in the system (2.34). Here, by equivalence we mean that, given any sufficiently and necessarily smooth right-hand side $f(t)$ (i.e. $f(t)$ is $\mu$-times continuously differentiable, where $\mu$ is the strangeness-index of the system (2.11)), there is a one-to-one correspondence between the solution sets of the two systems via a nonsingular matrix. As a natural extension of the theory of [28], [34] (Chapter III) for linear first-order variable coefficient DAEs, the technique of decoupling consists in several differentiation-and-elimination steps. In detail, with respect to the system (2.34), we

1. differentiate equation (2.34-l) and insert it in (2.34-g) to eliminate $\dot{y}_1(t)$;

2. differentiate equation (2.34-l) twice and insert it in (2.34-a) to eliminate $\ddot{y}_1(t)$;
3. differentiate equation (2.34-k) twice and insert it in (2.34-c) to eliminate $\ddot{y}_3(t)$;
4. differentiate equation (2.34-j) and insert it in (2.34-e) to eliminate $\dot{y}_5(t)$;
5. differentiate equation (2.34-h) and insert it in (2.34-b) to eliminate $\ddot{y}_2(t)$; and
6. differentiate the whole of or some parts of equation (2.34-i) and insert the derivatives in (2.34-a) - (2.34-d) to eliminate possibly existent $\dot{y}_7(t)$.

The above differentiation-and-elimination steps correspond to transforming the system (2.32) into an equivalent second-order system of DAEs

$$M^{\langle 1\rangle}\ddot{y}(t) + C^{\langle 1\rangle}\dot{y}(t) + K^{\langle 1\rangle}y(t) = f^{\langle 1\rangle}(t) \tag{2.35}$$

with $(M^{\langle 1\rangle}, C^{\langle 1\rangle}, K^{\langle 1\rangle}; f^{\langle 1\rangle})$ being of the following form

$$\left(\begin{bmatrix} 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&I_{d^{(2)}}&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0 \end{bmatrix}, \begin{bmatrix} 0&0&C&C&0&0&0&C\\ 0&C&C&C&0&C&0&C\\ 0&0&C&C&0&0&0&C\\ 0&0&C&C&0&0&0&C\\ 0&0&0&C&0&0&0&0\\ 0&0&0&0&0&I_{d^{(1)}}&0&0\\ 0&0&0&0&0&0&0&0\\ 0&I_{s^{(MC)}}&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0 \end{bmatrix},\right.$$

$$\left.\begin{bmatrix} 0&K&0&K&0&K&0&K\\ 0&K&0&K&0&K&0&K\\ 0&K&0&K&0&K&0&K\\ 0&K&0&K&0&K&0&K\\ 0&K&0&K&0&K&0&K\\ 0&K&0&K&0&K&0&K\\ 0&K&0&K&0&K&0&K\\ 0&K&0&K&0&K&0&K\\ 0&0&0&0&0&0&I_a&0\\ 0&0&0&0&I_{s^{(CK)}}&0&0&0\\ 0&0&I_{s^{(MK)}}&0&0&0&0&0\\ I_{s^{(MCK)}}&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0 \end{bmatrix}; \begin{bmatrix} \hat{f}_1(t)-\ddot{\hat{f}}_{12}(t)\\ \hat{f}_2(t)-\dot{\hat{f}}_8(t)\\ \hat{f}_3(t)-\ddot{\hat{f}}_{11}(t)\\ \hat{f}_4(t)\\ \hat{f}_5(t)-\dot{\hat{f}}_{10}(t)\\ \hat{f}_6(t)\\ \hat{f}_7(t)-\dot{\hat{f}}_{12}(t)\\ \hat{f}_8(t)\\ \hat{f}_9(t)\\ \hat{f}_{10}(t)\\ \hat{f}_{11}(t)\\ \hat{f}_{12}(t)\\ \hat{f}_{13}(t) \end{bmatrix}\right) \begin{array}{l} s^{(MCK)}\\ s^{(MC)}\\ s^{(MK)}\\ d^{(2)}\\ s^{(CK)}\\ d^{(1)}\\ s^{(MCK)}\\ s^{(MC)}\\ a\\ s^{(CK)}\\ s^{(MK)}\\ s^{(MCK)}\\ v \end{array}. \tag{2.36}$$

It is clear and should be stressed that the above procedure of differentiation and elimination only involves the sufficient and necessary order derivatives of the right-hand side $\hat{f}(t)$. Moreover, after the transformation from the system (2.32) to the system (2.35), the solution sets of the two systems are the same.

Then, a natural question arises, namely, what is the relation between the quadratic matrix polynomial associated with the system (2.32) and its counterpart which is associated with the equivalent system (2.35)? The following lemma gives an answer to this question.

**Lemma 2.11** *Let $\hat{A}(\lambda) = \lambda^2\hat{M} + \lambda\hat{C} + \hat{K}$, and let $A^{\langle 1\rangle}(\lambda) = \lambda^2 M^{\langle 1\rangle} + \lambda C^{\langle 1\rangle} + K^{\langle 1\rangle}$, where $\hat{M}, \hat{C}$ and $\hat{K}$ are as in (2.32), and $M^{\langle 1\rangle}, C^{\langle 1\rangle}$ and $K^{\langle 1\rangle}$ are as in (2.35). Then,*

$$A^{\langle 1\rangle}(\lambda) = E(\lambda)\hat{A}(\lambda), \tag{2.37}$$

*where $E(\lambda)$ is a unimodular matrix polynomial, i.e., the determinant of $E(\lambda)$ is a nonzero constant.*

**Proof**. Observe that, in terms of elementary row operations for matrix polynomials (cf. Subsection 4.2.2, p.63), the differentiation-and-elimination steps 1-6 (on pages 24-25) with respect to $\hat{A}(\lambda)$ correspond to premultiplying $\hat{A}(\lambda)$ by elementary matrix polynomial $E_1(\lambda)$, premultiplying $E_1(\lambda)\hat{A}(\lambda)$ by elementary matrix polynomial $E_2(\lambda)$, premultiplying $E_2(\lambda)E_1(\lambda)\hat{A}(\lambda)$ by elementary matrix polynomial $E_3(\lambda)$, premultiplying $E_3(\lambda)\cdots E_1(\lambda)\hat{A}(\lambda)$ by elementary matrix polynomial $E_4(\lambda)$, premultiplying $E_4(\lambda)\cdots E_1(\lambda)\hat{A}(\lambda)$ by elementary matrix polynomial $E_5(\lambda)$, and premultiplying $E_5(\lambda)\cdots E_1(\lambda)\hat{A}(\lambda)$ by elementary matrix polynomial $E_6(\lambda)$, respectively, where

$$E_1(\lambda) = \begin{bmatrix} I & & & & & \\ & \ddots & & & & \\ & & I & \cdots & -\lambda I & \\ & & & \ddots & \vdots & \\ & & & & I & \\ & & & & & I \end{bmatrix} \begin{array}{l} s^{(MCK)} \\ \vdots \\ s^{(MCK)} \\ \vdots \\ s^{(MCK)} \\ v \end{array}, E_2(\lambda) = \begin{bmatrix} I & \cdots & \cdots & & -\lambda^2 I & \\ & \ddots & & & \vdots & \\ & & I & & & \\ & & & \ddots & \vdots & \\ & & & & I & \\ & & & & & I \end{bmatrix} \begin{array}{l} s^{(MCK)} \\ \vdots \\ s^{(MCK)} \\ \vdots \\ s^{(MCK)} \\ v \end{array},$$

$$E_3(\lambda) = \begin{bmatrix} I & & & & & & \\ & \ddots & & & & & \\ & & I & \cdots & -\lambda^2 I & & \\ & & & \ddots & \vdots & & \\ & & & & I & & \\ & & & & & \ddots & \\ & & & & & & I \end{bmatrix} \begin{array}{l} s^{(MCK)} \\ \vdots \\ s^{(MK)} \\ \vdots \\ s^{(MK)} \\ \vdots \\ v \end{array}, E_4(\lambda) = \begin{bmatrix} I & & & & & & \\ & \ddots & & & & & \\ & & I & \cdots & -\lambda I & & \\ & & & \ddots & \vdots & & \\ & & & & I & & \\ & & & & & \ddots & \\ & & & & & & I \end{bmatrix} \begin{array}{l} s^{(MCK)} \\ \vdots \\ s^{(CK)} \\ \vdots \\ s^{(CK)} \\ \vdots \\ v \end{array},$$

$$E_5(\lambda) = \left[\begin{array}{cccccc} I & & & & & \\ & I & \cdots & -\lambda I & & \\ & & \ddots & \vdots & & \\ & & & I & & \\ & & & & \ddots & \\ & & & & & I \end{array}\right] \begin{array}{l} s^{(MCK)} \\ s^{(MC)} \\ \vdots \\ s^{(MC)} \\ \vdots \\ v \end{array}, E_6(\lambda) = \left[\begin{array}{cccccccc} I & \cdots & & & \cdots & -\lambda C_{1,7} & & \\ & I & \cdots & & \cdots & -\lambda C_{2,7} & & \\ & & I & \cdots & \cdots & -\lambda C_{3,7} & & \\ & & & I & \cdots & -\lambda C_{4,7} & & \\ & & & & \ddots & \vdots & & \\ & & & & & I & & \\ & & & & & & \ddots & \\ & & & & & & & I \end{array}\right] \begin{array}{l} s^{(MCK)} \\ s^{(MC)} \\ s^{(MK)} \\ d^{(2)} \\ \vdots \\ a \\ \vdots \\ v \end{array},$$

and $E_i(t)$, $i = 1, \ldots, 6$ are partitioned conformally with the block matrices in (2.28). Let $E(\lambda) = E_6(\lambda) \cdots E_1(\lambda)$. Then we have $A^{\langle 1 \rangle}(\lambda) = E(\lambda)\hat{A}(\lambda)$. Since the determinant of each $E_i(t)$, $i = 1, \ldots, 6$, is a nonzero constant, the determinant of $E(\lambda)$ is also a nonzero constant. $\square$

Let us now turn back to the new matrix triple $(M^{\langle 1 \rangle}, C^{\langle 1 \rangle}, K^{\langle 1 \rangle})$ in (2.36) obtained after the differentiation-and-elimination steps. We can again transform it to the condensed form (2.28), and apply the differentiation-and-elimination steps to pass it to the form (2.36). In this way, therefore, we can conduct an inductive procedure to get a sequence of triples of matrices $(M^{\langle i \rangle}, C^{\langle i \rangle}, K^{\langle i \rangle})$, $i \in \mathbb{N}_0$, where $(M^{\langle 0 \rangle}, C^{\langle 0 \rangle}, K^{\langle 0 \rangle}) = (M, C, K)$ and $(M^{\langle i+1 \rangle}, C^{\langle i+1 \rangle}, K^{\langle i+1 \rangle})$ is derived from $(M^{\langle i \rangle}, C^{\langle i \rangle}, K^{\langle i \rangle})$ by bringing it into the form (2.28) and then applying the differentiation-and-elimination steps.

Comparing $\hat{M}$ in (2.28) with $M^{\langle 1 \rangle}$ in (2.36), we have

$$\begin{aligned} \operatorname{rank}(M^{\langle 1 \rangle}) &= \operatorname{rank}(\hat{M}) - s_{\langle 0 \rangle}^{(MCK)} - s_{\langle 0 \rangle}^{(MK)} - s_{\langle 0 \rangle}^{(MC)} \\ &= \operatorname{rank}(M^{\langle 0 \rangle}) - s_{\langle 0 \rangle}^{(MCK)} - s_{\langle 0 \rangle}^{(MK)} - s_{\langle 0 \rangle}^{(MC)}, \end{aligned} \tag{2.38}$$

where $s_{\langle 0 \rangle}^{(MCK)}$, $s_{\langle 0 \rangle}^{(MK)}$, and $s_{\langle 0 \rangle}^{(MC)}$ denote the strangeness due to $M^{\langle 0 \rangle}, C^{\langle 0 \rangle}, K^{\langle 0 \rangle}$, the strangeness due to $M^{\langle 0 \rangle}, K^{\langle 0 \rangle}$, and the strangeness due to $M^{\langle 0 \rangle}, C^{\langle 0 \rangle}$, respectively. Since after the differentiation-and-elimination step 4 (on page 25), equation (2.34-j) becomes uncoupled purely algebraic equation, it follows that

$$\operatorname{rank}(K) \geq a_{\langle 1 \rangle} \geq \left( a_{\langle 0 \rangle} + s_{\langle 0 \rangle}^{(CK)} \right), \tag{2.39}$$

where $a_{\langle 1 \rangle}$, $a_{\langle 0 \rangle}$, and $s_{\langle 0 \rangle}^{(CK)}$ denote the number of algebraic part of $(M^{\langle 1 \rangle}, C^{\langle 1 \rangle}, K^{\langle 1 \rangle})$, the number of algebraic part of $(M^{\langle 0 \rangle}, C^{\langle 0 \rangle}, K^{\langle 0 \rangle})$, and the strangeness due to $C^{\langle 0 \rangle}, K^{\langle 0 \rangle}$, respectively. Hence, the relations in (2.38) and (2.39) guarantee that after a finite number (say $q$) of steps, the strangeness $s_{\langle q \rangle}^{(MCK)}$ due to $M^{\langle q \rangle}, C^{\langle q \rangle}, K^{\langle q \rangle}$, the strangeness $s_{\langle q \rangle}^{(MK)}$ due to $M^{\langle q \rangle}, K^{\langle q \rangle}$, the strangeness $s_{\langle q \rangle}^{(MC)}$ due to $M^{\langle q \rangle}, C^{\langle q \rangle}$, and the strangeness $s_{\langle q \rangle}^{(CK)}$ due to $C^{\langle q \rangle}, K^{\langle q \rangle}$ must vanish. If this is the case, then we arrive at a final equivalent second-order system of DAEs with a very special structure (we call the system *strangeness-free*). Note that, by the above procedure, there may exist many ways in which the original system can be transformed into an equivalent strangeness-free system. But since in each way the strangeness-free system obtained is equivalent to the original system, here we just take any one of them, and assume that $\tilde{\mu} \in \mathbb{N}_0$ is given by the number of necessary differentiations of $f(t)$ in (2.11) that is required in the given way to obtain an equivalent strangeness-free system. Thus, we have the following essential result of this section.

**Theorem 2.12** *Let $f(t) \in \mathcal{C}^{\tilde{\mu}}([t_0, t_1], \mathbb{C}^m)$, $\tilde{\mu} \in \mathbb{N}_0$. Then, the system (2.11) is equivalent (in the sense that there is a one-to-one correspondence between the solution sets of the two systems via a nonsingular matrix) to a system of second-order differential-algebraic equations $\tilde{M}\ddot{\tilde{x}}(t) + \tilde{C}\dot{\tilde{x}}(t) + \tilde{K}\tilde{x}(t) = \tilde{f}(t)$ of the form*

$$
\begin{array}{ll}
(a) & \ddot{\tilde{x}}_1(t) + \tilde{C}_{1,1}\dot{\tilde{x}}_1(t) + \tilde{C}_{1,4}\dot{\tilde{x}}_4(t) + \tilde{K}_{1,1}\tilde{x}_1(t) + \tilde{K}_{1,2}\tilde{x}_2(t) + \tilde{K}_{1,4}\tilde{x}_4(t) = \tilde{f}_1(t),\\
(b) & \dot{\tilde{x}}_2(t) + \tilde{K}_{2,1}\tilde{x}_1(t) + \tilde{K}_{2,2}\tilde{x}_2(t) + \tilde{K}_{2,4}\tilde{x}_4(t) = \tilde{f}_2(t),\\
(c) & \tilde{x}_3(t) = \tilde{f}_3(t),\\
(d) & 0 = \tilde{f}_4(t),
\end{array}
\tag{2.40}
$$

*where the inhomogeneity $\tilde{f}(t) := \left[\tilde{f}_1^T(t), \ldots, \tilde{f}_4^T(t)\right]^T$ is determined by $f^{(0)}(t)$, ..., $f^{(\tilde{\mu})}(t)$. In particular, $\tilde{d}^{(2)}$, $\tilde{d}^{(1)}$, and $\tilde{a}$ are the number of second-order differential, first-order differential, and algebraic components of the unknown $\tilde{x}(t) := \left[\tilde{x}_1^T(t), \ldots, \tilde{x}_4^T(t)\right]^T$ in (2.40-a), (2.40-b), and (2.40-c) respectively, while $\tilde{u}$ is the dimension of the undetermined vector $\tilde{x}_4(t)$ in (2.40-a) and (2.40-b), and $\tilde{v}$ is the number of conditions in (2.40-d).*

**Proof**. In the given way, inductively transforming $(M, C, K)$ to the condensed form (2.28) in Lemma 2.8 and then converting it by differentiation-and-elimination steps into the form in (2.36) until $s_i^{(MCK)} = s_i^{(MC)} = s_i^{(MK)} = s_i^{(CK)} = 0$ yield a triple $(\tilde{M}, \tilde{C}, \tilde{K})$ of matrices of the form

$$
\left( \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \tilde{C}_{1,1} & 0 & 0 & \tilde{C}_{1,4} \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \tilde{K}_{1,1} & \tilde{K}_{1,2} & 0 & \tilde{K}_{1,4} \\ \tilde{K}_{2,1} & \tilde{K}_{2,2} & 0 & \tilde{K}_{2,4} \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right), \tag{2.41}
$$

with block sizes $\tilde{d}^{(2)}, \tilde{d}^{(1)}, \tilde{a}, \tilde{v}$ for the rows and $\tilde{d}^{(2)}, \tilde{d}^{(1)}, \tilde{a}, \tilde{u}$ for the columns. By (2.33), we know that the transformation from $(M, C, K)$ to $(\hat{M}, \hat{C}, \hat{K})$ in the condensed form (2.28) establishes, via a nonsingular matrix, a one-to-one correspondence between the solution sets of the two corresponding systems of DAEs. We also note that by (2.35), the differentiation-and-elimination steps do not change the solution sets at all. Hence, there exists a nonsingular matrix $\tilde{Q}$ such that for any solution $x(t)$ of the system (2.11) (if existent), there corresponds a solution $\tilde{x}(t)$ of the system (2.40) satisfying

$$x(t) = \tilde{Q}\tilde{x}(t), \tag{2.42}$$

and vice versa. □

Now, we can answer the questions at the beginning of this chapter, which are concerned with the existence and uniqueness of solutions and consistency of initial conditions.

**Corollary 2.13** *Under the assumption of Theorem 2.12, the following statements hold.*

1. *The system (2.11) is solvable if and only if one of the following two cases happens.*
   *(i)* $\tilde{v} = 0$.
   *(ii) If* $\tilde{v} > 0$*, then the* $\tilde{u}$ *functional consistency conditions*
   $$\tilde{f}_4(t) = 0 \tag{2.43}$$
   *are satisfied.*
2. *If the system (2.11) is solvable, then it is uniquely solvable without providing any initial condition if and only if the conditions*
   $$\tilde{d}^{(2)} = \tilde{d}^{(1)} = \tilde{u} = 0 \tag{2.44}$$
   *hold.*
3. *If the system (2.11) is solvable, then initial conditions (2.12) are consistent if and only if one of the following two cases happens.*
   *(i)* $\tilde{a} = 0$.
   *(ii) If* $\tilde{a} > 0$*, then the* $\tilde{a}$ *conditions*
   $$\tilde{x}_3(t_0) = \tilde{f}_3(t_0),\ \dot{\tilde{x}}_3(t_0) = \left.\frac{d\tilde{f}_3(t)}{dt}\right|_{t_0+} \tag{2.45}$$
   *are implied by (2.12).*
4. *If the initial value problem (2.11)-(2.12) is solvable, then it is uniquely solvable if and only if*
   $$\tilde{u} = 0 \tag{2.46}$$
   *holds.*

**Proof**. The results are direct consequences of Theorem 2.12. □

In order to write down explicit solutions of the coupled but strangeness-free second-order and first-order differential equations (2.40-a) and (2.40-b), we need the following result on linear ordinary differential equations from E. A. Coddington and N. Levinson [10] (p. 78).

**Lemma 2.14** *The inhomogeneous system of linear ordinary differential equations*

$$\dot{x}(t) = Ax(t) + b(t), \ t \in [t_0, t_1], \tag{2.47}$$

*together with an initial condition*

$$x(t_0) = x_0 \tag{2.48}$$

*has a unique solution*

$$x(t) = e^{(t-t_0)A}x_0 + \int_{t_0}^{t} e^{(t-s)A}b(s)\ ds. \tag{2.49}$$

**Corollary 2.15** *Under the assumption of Theorem 2.12, the system (2.40) together with initial conditions*

$$\tilde{x}_1(t_0) = \tilde{x}_{1,0}, \ \dot{\tilde{x}}_1(t_0) = \tilde{x}_{1,0}^{[1]}, \ \tilde{x}_2(t_0) = \tilde{x}_{2,0},$$

*is equivalent to the following system*

$$\begin{array}{ll} (a) & \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} \tilde{x}_1 \\ \dot{\tilde{x}}_1 \\ \tilde{x}_2 \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & I \end{bmatrix} e^{(t-t_0)A} \begin{bmatrix} \tilde{x}_{1,0} \\ \tilde{x}_{1,0}^{[1]} \\ \tilde{x}_{2,0} \end{bmatrix} + \int_{t_0}^{t} e^{(t-s)A}b(s)\ ds, \\ (b) & \tilde{x}_3(t) = \tilde{f}_3(t), \\ (c) & 0 = \tilde{f}_4(t), \end{array} \tag{2.50}$$

*where*

$$A = \begin{bmatrix} 0 & I & 0 \\ -\tilde{K}_{1,1} & -\tilde{C}_{1,1} & -\tilde{K}_{1,2} \\ -\tilde{K}_{2,1} & 0 & -\tilde{K}_{2,2} \end{bmatrix}, \begin{bmatrix} \tilde{x}_{1,0} \\ \tilde{x}_{1,0}^{[1]} \\ \tilde{x}_{2,0} \end{bmatrix} = \begin{bmatrix} \tilde{x}_1(t_0) \\ \dot{\tilde{x}}_1(t_0) \\ \tilde{x}_2(t_0) \end{bmatrix}, \tag{2.51}$$

$$b(t) = \begin{bmatrix} 0 \\ \tilde{f}_1(t) - \tilde{C}_{1,4}\dot{\tilde{x}}_4(t) - \tilde{K}_{1,4}\tilde{x}_4(t) \\ \tilde{f}_2(t) - \tilde{K}_{2,4}\tilde{x}_4(t) \end{bmatrix}. \tag{2.52}$$

**Proof**. We transform the system (2.40-a), (2.40-b) of second-order ODEs into an equivalent system of linear first-order ODEs

$$\begin{bmatrix} \tilde{x}_1 \\ \dot{\tilde{x}}_1 \\ \tilde{x}_2 \end{bmatrix}^{\cdot} = A \begin{bmatrix} \tilde{x}_1 \\ \dot{\tilde{x}}_1 \\ \tilde{x}_2 \end{bmatrix} + b(t),$$

where $A, b(t)$ are as in (2.51), (2.52) respectively. Then, by Lemma 2.14, the result is immediate. $\square$

Now, we can turn our attention to the strangeness-index $\mu$ of the system (2.11) which has been introduced in Definition 2.2, and obtain the following result.

**Corollary 2.16** *Under the assumption of Theorem 2.12, assume that the system (2.11) is solvable. Let $\mu \in \mathbb{N}_0$ be the strangeness-index of the system (2.11). Then,*

$$\mu = \tilde{\mu}. \tag{2.53}$$

**Proof**. By Theorem 2.12, there is a one-to-one correspondence (2.42), via a nonsingular matrix, between the solution sets of the two systems (2.11) and (2.40). Therefore, the system (2.11) is solvable if and only if the system (2.40) is solvable. Thus, by Corollaries 2.13 and 2.15, we can see that, provided that initial conditions are necessary and consistent, the right-hand side $f(t)$ in (2.11) must be $\tilde{\mu}$-times continuously differentiable in order to determine $\tilde{x}(t)$ as a continuous function of $t$, and therefore by the one-to-one correspondence (2.42), to determine $x(t)$ as a continuous function of $t$. By Definition 2.2 of strangeness-index, (2.53) immediately follows. $\square$

Of special interest is the case of the system (2.11) of DAEs with which a *regular* quadratic matrix polynomial is associated. Here, we call a matrix polynomial $A(\lambda)$ of size $m \times n$ a *regular* matrix polynomial if $m = n$ and the determinant of $A(\lambda)$ is not identically equal to zero; otherwise, it is called *singular* (for more details, cf. Chapter 4).

As in the case of linear first-order constant coefficient system of DAEs, regularity of the quadratic matrix polynomial associated with the system (2.11) is closely related to the solution behaviour of the system (2.11). Indeed, regularity of the quadratic matrix polynomial is a sufficient and necessary condition for the property that for every inhomogeneity $f(t) \in \mathcal{C}^{\mu}([t_0, t_1], \mathbb{C}^n)$, where $\mu$ is the strangeness-index of the system (2.11), there are initial conditions such that the initial value problem associated with (2.11) has a unique solution. In the following two theorems, we show the sufficiency and the necessity respectively.

**Theorem 2.17** *Let $M, C, K \in \mathbb{C}^{n\times n}$, and let the matrix polynomial $A(\lambda) := \lambda^2 M + \lambda C + K$ be regular. Let $f(t) \in \mathcal{C}^{\mu}([t_0, t_1], \mathbb{C}^n)$, where $\mu$ is the strangeness-index of the system (2.11) associated with $A(\lambda)$. Then, there is a unique solution of the initial value problem (2.11)-(2.12), provided that the given initial conditions (2.12) are necessary and consistent.*

**Proof**. Let $\tilde{A}(\lambda) := \lambda^2 \tilde{M} + \lambda \tilde{C} + \tilde{K}$, where $\tilde{M}, \tilde{C}, \tilde{K} \in \mathbb{C}^{n\times n}$ are associated with the system (2.40). Then, inductively by (2.33) and Lemma 2.11, we have

$$\tilde{A}(\lambda) = E_r(\lambda)P_r E_{r-1}(\lambda)P_{r-1}\cdots E_1(\lambda)P_1 A(\lambda)Q_1 Q_2\cdots Q_r, \tag{2.54}$$

where $P_i$ and $Q_i$, $i = 1, \ldots, r$, are nonsigular matrices, and $E_i(\lambda)$, $i = 1, \ldots, r$, are unimodular matrix polynomials. From (2.54) it follows that $\det(\tilde{A}(\lambda)) = \frac{1}{c}\det(A(\lambda))$, where $c$ is a nonzero constant. Since $\det(A(\lambda)) \not\equiv 0$, we have $\det(\tilde{A}(\lambda)) \not\equiv 0$, in other words, $\tilde{A}(\lambda)$ is regular. This immediately implies that in the system (2.40),

$$\tilde{u} = 0, \quad \tilde{v} = 0.$$

Then, under the condition that the given initial conditions (2.12) are consistent, the existence and uniqueness of solutions of the initial value problem (2.11)-(2.12) directly follows from Corollaries 2.13, 2.15 and 2.16. $\square$

**Theorem 2.18** *Let $M, C, K \in \mathbb{C}^{m\times n}$, and suppose that the matrix polynomial $A(\lambda) := \lambda^2 M + \lambda C + K$ is singular.*

1. *If $\operatorname{rank}(A(\lambda)) < n$ for all $\lambda \in \mathbb{C}$, then the homogeneous initial value problem*

$$M\ddot{x}(t) + C\dot{x}(t) + Kx(t) = 0, \;\; x(t_0) = \dot{x}(t_0) = 0 \tag{2.55}$$

*has a nontrivial solution.*

2. *If $\operatorname{rank}(A(\lambda)) = n$ for some $\lambda \in \mathbb{C}$ and hence $m > n$, then there exist arbitrary smooth inhomogeneities $f(t)$ for which the the corresponding system (2.11) of DAEs is not solvable.*

**Proof**.

1. Suppose that $\operatorname{rank}(A(\lambda)) < n$ for all $\lambda \in \mathbb{C}$. Let $\lambda_i$, $i = 1, \ldots, n+1$, be pairwise different complex numbers. Then, for each $\lambda_i$, there exists a nonzero vector $v_i \in \mathbb{C}^n$ satisfying $(\lambda_i^2 M + \lambda_i C + K)v_i = 0$, and clearly the vectors $v_i$, $i = 1, \ldots, n+1$,

are linearly dependent. Hence, there exist $\alpha_i \in \mathbb{C}$, $i = 1, \ldots, n+1$, not all of them being zero, such that

$$\sum_{i=1}^{n+1} \alpha_i v_i = 0.$$

For the function $x(t)$ defined by

$$x(t) = \sum_{i=1}^{n+1} \alpha_i v_i e^{\lambda_i (t-t_0)},$$

we then have $x(t_0) = 0$ as well as

$$M\ddot{x}(t) + C\dot{x}(t) + Kx(t) = \sum_{i=1}^{n+1} \alpha_i (\lambda_i^2 M + \lambda_i C + K) v_i e^{\lambda_i (t-t_0)} = 0.$$

Since $x(t)$ is not the zero function, it is a nontrivial solution of the homogeneous initial value problem (2.55).

2. Suppose that there is a $\lambda \in \mathbb{C}$ such that $\operatorname{rank}(A(\lambda)) = n$. Because $A(\lambda)$ is assumed to be singular, we have $m > n$. We set

$$x(t) = e^{\lambda t} \tilde{x}(t),$$

and therefore

$$\dot{x}(t) = e^{\lambda t} \left( \dot{\tilde{x}}(t) + \lambda \tilde{x}(t) \right), \ \ddot{x}(t) = e^{\lambda t} \left( \ddot{\tilde{x}}(t) + 2\lambda \dot{\tilde{x}}(t) + \lambda^2 \tilde{x}(t) \right),$$

such that (2.11) is transformed to

$$M \left( \ddot{\tilde{x}}(t) + 2\lambda \dot{\tilde{x}}(t) \right) + C \dot{\tilde{x}}(t) + \left( \lambda^2 M + \lambda C + K \right) \tilde{x}(t) = e^{-\lambda t} f(t).$$

Since $\lambda^2 M + \lambda C + K$ has full column rank, there exists a nonsingular matrix $P \in \mathbb{C}^{m,m}$, such this equation premultiplied by $P$ gives

$$\begin{bmatrix} M_1 \\ M_2 \end{bmatrix} \left( \ddot{\tilde{x}}(t) + 2\lambda \dot{\tilde{x}}(t) \right) + \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \dot{\tilde{x}}(t) + \begin{bmatrix} I \\ 0 \end{bmatrix} \tilde{x}(t) = \begin{bmatrix} f_1(t) \\ f_2(t) \end{bmatrix}.$$

Obviously the matrix polynomial $\xi^2 M_1 + \xi \left( 2\lambda M_1 + C_1 \right) + I$ in $\xi$ is regular. By Theorem 2.17, the initial value problem

$$M_1 \ddot{\tilde{x}}(t) + \left( 2\lambda M_1 + C_1 \right) \dot{\tilde{x}}(t) + \tilde{x}(t) = f_1(t), \ \tilde{x}(t_0) = \tilde{x}_0, \dot{\tilde{x}}(t_0) = \tilde{x}_0^{[1]}$$

has a unique solution for every sufficiently smooth inhomogeneity $f_1(t)$ and for every consistent initial value. But then

$$f_2(t) = M_2 \left( \ddot{\tilde{x}}(t) + 2\lambda \dot{\tilde{x}}(t) \right) + C_2 \dot{\tilde{x}}(t)$$

is a consistency condition for the inhomogeneity $f_2(t)$ that must hold for a solution to exist. This immediately shows that there are arbitrary smooth functions $f(t)$ for which this consistency condition is not satisfied. □

## 2.5 Linear $l$th-Order DAEs with Constant Coefficients ($l \geq 3$)

In this section we shall extend the main results of the last section to the case of general systems (2.1) of linear $l$th-order ($l \geq 3$) DAEs with constant coefficients.

As we mentioned in Section 2.3, by induction one can get a condensed form via strong equivalence transformations for $(l+1)$-tuples $(A_l, \ldots, A_1, A_0)$ of matrices which is similar to (2.21) and (2.28) for matrix pairs and triples respectively. Clearly, with the condensed form a set of invariant quantities are associated. In the context of the corresponding system (2.1), among the invariant quantities, especially, are those which can be used to characterize algebraic part, 1st-order, 2nd-order, ..., and $l$th-order differential parts, and strange parts due to each two, each three, ..., each $l$, and $l+1$ matrices out of $A_l, \ldots, A_1$, and $A_0$, respectively.

Then, based on the condensed form for $(l+1)$-tuples of matrices, one can write down the system of differential-algebraic equations after the strong equivalence transformations. Analogous to the treatment of systems of second-order DAEs in Section 2.4, one can design differentiation-and-elimination steps in order to decouple those equations coupled to each other in the system and to reduce it to a simpler but equivalent system, which can be again transformed to the condensed form. Inductively, by this procedure one can get a sequence of $(l+1)$-tuples of matrices, and after a finite number of steps, it can be expected that all strange parts of the corresponding system will vanish, in other words, in the final the system becomes strangeness-free. Here, for convenience of expression, we only state the essential results which are parallel to Theorem 2.12 and its main consequences in the last section.

Clearly, there may exist many ways in which the original system can be transformed into an equivalent strangeness-free system. But, as we mentioned in the last section, since in each way the strangeness-free system obtained is equivalent to the original system, here we just take any one of them, and assume that $\tilde{\mu} \in \mathbb{N}_0$ is given by the number of necessary differentiations of $f(t)$ in (2.1) that is required in the given way to obtain an equivalent strangeness-free system. Later, as in the case of the second-order systems, we will see that $\tilde{\mu} \in \mathbb{N}_0$ is in fact the strangeness-index of the $l$th-order system (2.1), provided that the system is solvable.

**Theorem 2.19** *Let $f(t) \in \mathcal{C}^{\tilde{\mu}}([t_0, t_1], \mathbb{C}^m)$, $\tilde{\mu} \in \mathbb{N}_0$. Then, the system (2.1) is equivalent (in the sense that there is a one-to-one correspondence between the solution sets of the two systems via a nonsingular matrix) to a system of $l$th-order differential-algebraic equations*

$$A^{\langle l \rangle}\tilde{x}^{(l)}(t) + A^{\langle l-1 \rangle}\tilde{x}^{(l-1)}(t) + \cdots + A^{\langle 0 \rangle}\tilde{x}(t) = \tilde{f}(t)$$

*of the form*

$$\begin{array}{ll}
(1) & \frac{d^l \tilde{x}_1(t)}{dt^l} + \sum_{i=0}^{l-1}\sum_{j=i}^{l-1} A_{1,l-j}^{\langle i \rangle}\frac{d^i \tilde{x}_{l-j}(t)}{dt^i} + \sum_{i=0}^{l-1} A_{1,l+2}^{\langle i \rangle}\frac{d^i \tilde{x}_{l+2}(t)}{dt^i} = \tilde{f}_1(t), \\
 & \\
(2) & \frac{d^{l-1} \tilde{x}_2(t)}{dt^{l-1}} + \sum_{i=0}^{l-2}\sum_{j=i}^{l-2} A_{2,l-1-j}^{\langle i \rangle}\frac{d^i \tilde{x}_{l-1-j}(t)}{dt^i} \\
 & \qquad + \sum_{i=0}^{l-2}\left(A_{2,1}^{\langle i \rangle}\frac{d^i \tilde{x}_1(t)}{dt^i} + A_{2,l+2}^{\langle i \rangle}\frac{d^i \tilde{x}_{l+2}(t)}{dt^i}\right) = \tilde{f}_2(t), \\
\vdots & \qquad \vdots \\
(k) & \frac{d^{l-k+1} \tilde{x}_{l-k+1}(t)}{dt^{l-k+1}} + \sum_{i=0}^{l-k}\sum_{j=i}^{l-k} A_{k,l-k+1-j}^{\langle i \rangle}\frac{d^i \tilde{x}_{l-k+1-j}(t)}{dt^i} \\
 & \qquad + \sum_{i=0}^{l-k}\left(\sum_{j=1}^{k} A_{k,j}^{\langle i \rangle}\frac{d^i \tilde{x}_1(t)}{dt^i} + A_{k,l+2}^{\langle i \rangle}\frac{d^i \tilde{x}_{l+2}(t)}{dt^i}\right) = \tilde{f}_k(t), \ (1 \leq k \leq l) \\
\vdots & \qquad \vdots \\
(l+1) & \tilde{x}_{l+1}(t) = \tilde{f}_{l+1}(t), \\
 & \\
(l+2) & 0 = \tilde{f}_{l+2}(t),
\end{array} \tag{2.56}$$

*where $A_{p,q}^{\langle i \rangle}$, $1 \leq p \leq (l+2)$, $1 \leq q \leq (l+2)$, denotes a submatrix of $A^{\langle i \rangle}$, $i = 0, 1, \ldots, l$, and the inhomogeneity $\tilde{f}(t) := \left[\tilde{f}_1^T(t), \ldots, \tilde{f}_{l+2}^T(t)\right]^T$ is determined by $f^{(0)}(t)$, …, $f^{(\tilde{\mu})}(t)$. In particular, $\tilde{d}^{(l)}$, …, $\tilde{d}^{(1)}$, and $\tilde{a}$ are the number of $l$th-order differential, …, first-order differential, and algebraic components of the unknown $\tilde{x}(t) := \left[\tilde{x}_1^T(t), \ldots, \tilde{x}_{l+2}^T(t)\right]^T$ in (2.56-1), …, and (2.56-(l+1)) respectively, while $\tilde{u}$ is the dimension of the undetermined vector $\tilde{x}_{l+2}(t)$ in (2.56-1), …, (2.56-l), and $\tilde{v}$ is the number of conditions in (2.56-(l+2)).*

**Proof**. The proof is analogous to the proof of Theorem 2.12 and follows by induction. □

The following corollary answers question 2, questions 4 and 5 that were posed at the beginning of this chapter.

**Corollary 2.20** *Under the assumption of Theorem 2.19, the following statements hold.*

1. *The system (2.1) is solvable if and only if one of the following two cases happens.*

   *(i)* $\tilde{v} = 0$.

   *(ii) If* $\tilde{v} > 0$, *then the* $\tilde{u}$ *functional consistency conditions*
$$\tilde{f}_{l+2}(t) = 0 \tag{2.57}$$
   *are satisfied.*

2. *If the system (2.1) is solvable, then it is uniquely solvable without providing any initial condition if and only if the conditions*
$$\tilde{d}^{(l)} = \cdots = \tilde{d}^{(2)} = \tilde{d}^{(1)} = \tilde{u} = 0 \tag{2.58}$$
   *hold.*

3. *If the system (2.1) is solvable, then initial conditions (2.2) are consistent if and only if one of the following two cases happens.*

   *(i)* $\tilde{a} = 0$.

   *(ii) If* $\tilde{a} > 0$, *then the* $\tilde{a}$ *conditions*
$$\begin{aligned} &\tilde{x}_{l+1}(t_0) = \tilde{f}_{l+1}(t_0), \\ &\dot{\tilde{x}}_{l+1}(t_0) = \left.\tfrac{d\tilde{f}_{l+1}(t)}{dt}\right|_{t_0+}, \ldots, \tfrac{d^{l-1}\tilde{x}_{l+1}(t_0)}{dt^{l-1}} = \left.\tfrac{d^{l-1}\tilde{f}_{l+1}(t)}{dt^{l-1}}\right|_{t_0+} \end{aligned} \tag{2.59}$$
   *are implied by (2.2).*

4. *If the initial value problem (2.1)-(2.2) is solvable, then it is uniquely solvable if and only if*
$$\tilde{u} = 0 \tag{2.60}$$
   *holds.*

**Proof**. The results are direct consequences of Theorem 2.19. □

To answer question 3, namely, to determine the strangeness-index of the system (2.1), we have the following corollary.

**Corollary 2.21** *Under the assumption of Theorem 2.19, assume that the system (2.1) is solvable. Let* $\mu \in \mathbb{N}_0$ *be the strangeness-index of the system (2.1). Then,*
$$\tilde{\mu} = \mu. \tag{2.61}$$

**Proof**. The proof is analogous to the proof of Corollary 2.16. $\square$

Finally, let us turn to the special case of the system (2.1) with which a *regular* matrix polynomial $A(\lambda) = \sum_{i=0}^{l} \lambda_i A_i$, $\lambda \in \mathbb{C}$, of degree of $l$ is associated. As we presented in the last section, regularity of the matrix polynomial $A(\lambda) = \sum_{i=0}^{l} \lambda_i A_i$ is closely related to existence and uniqueness of the solutions of the initial value problem which is assocaited with the system (2.1).

**Theorem 2.22** *Let $A_i \in \mathbb{C}^{n\times n}$, $i = 0, 1, \ldots, l$, and let the matrix polynomial $A(\lambda) := \sum_{i=0}^{l} \lambda_i A_i$ be regular. Let $f(t) \in \mathcal{C}^{\mu}([t_0, t_1], \mathbb{C}^n)$, where $\mu$ is the strangeness index of the system (2.1) associated with $A(\lambda)$. Then, there is a unique solution of the initial value problem (2.1)-(2.2), provided that the given initial conditions (2.2) are consistent.*

**Proof**. The proof is analogous to the proof of Theorem 2.17. $\square$

**Theorem 2.23** *Let $A_i \in \mathbb{C}^{m\times n}$, $i = 0, 1, \ldots, l$, and suppose that the matrix polynomial $A(\lambda) := \sum_{i=0}^{l} \lambda_i A_i$ is singular.*

1. *If* $\operatorname{rank}(A(\lambda)) < n$ *for all* $\lambda \in \mathbb{C}$, *then the homogeneous initial value problem*

$$A_l x^{(l)}(t) + A_{l-1} x^{(l-1)}(t) + \cdots + A_0 x(t) = 0, \ x(t_0) = \dot{x}(t_0) = \cdots = x^{(l-1)}(t_0) = 0 \tag{2.62}$$

   *has a nontrivial solution.*

2. *If* $\operatorname{rank}(A(\lambda)) = n$ *for some* $\lambda \in \mathbb{C}$ *and hence* $m > n$, *then there exist arbitrary smooth inhomogeneities* $f(t)$ *for which the the corresponding system (2.1) of DAEs is not solvable.*

**Proof**. The proof is analogous to the proof of Theorem 2.18. $\square$

**Remark 2.24** It should be pointed out that the importance of *regularity* and *singularity* of matrix polynomials in the context of the solution behaviour of systems of differential-algebraic equations, which we have seen from the above theorems, provides one of the major motivations for our later study carried out in the second part of this thesis. $\lozenge$

In the next chapter, we shall generalize the techniques employed and the results obtained in this chapter to the case of higher-order systems of DAEs with variable coefficients.

# Chapter 3

# Linear Higher-Order DAEs with Variable Coefficients

In this chapter, we study linear $l$th-order differential-algebraic equations with variable coefficients

$$A_l(t)x^{(l)}(t) + A_{l-1}(t)x^{(l-1)}(t) + \cdots + A_0(t)x(t) = f(t),\ t \in [t_0, t_1], \tag{3.1}$$

where $A_i(t) \in \mathcal{C}([t_0, t_1], \mathbb{C}^{m\times n})$, $i = 0, 1, \ldots, l$, $A_l(t) \not\equiv 0$, $f(t) \in \mathcal{C}([t_0, t_1], \mathbb{C}^m)$, possibly together with initial conditions

$$x(t_0) = x_0,\ \ldots,\ x^{(l-2)}(t_0) = x_0^{[l-2]},\ x^{(l-1)}(t_0) = x_0^{[l-1]}, \quad x_0, \ldots, x_0^{[l-2]}, x_0^{[l-1]} \in \mathbb{C}^n. \tag{3.2}$$

As in the case of constant coefficients, we shall apply very similar techniques (transforming, differentiating, and inserting) to the system (3.1) with variable coefficients, and obtain parallel results on the system (3.1), and on the initial value problem (3.1)-(3.2).

Analogous to Section 2.3, in Section 3.1 we concentrate on the treatment of linear second-order DAEs with variable coefficients. We shall prove that the quantities developed in Section 2.3 are still invariant under *local equivalence transformations*, and present a condensed form under a set of *regular conditions*. Later, in Section 3.2, based on the results of Section 3.1, we describe the solution behaviour (solvability, uniqueness of solutions and consistency of initial values) of the higher-order system (3.1) and of the initial value problem (3.1)-(3.2).

It should be pointed out that the work in this chapter is carried out along the lines of the work with respect to linear first-order DAEs with variable coefficients in [28, 29, 34]; for a comprehensive exposition, we refer to [34], Chapter 3.

## 3.1 Condensed Form for Triples of Matrix-Valued Functions

In this section, we shall mainly treat systems of linear second-order DAEs with variable coefficients

$$M(t)\ddot{x}(t) + C(t)\dot{x}(t) + K(t)x(t) = f(t), \quad t \in [t_0, t_1], \tag{3.3}$$

where $M(t), C(t), K(t) \in \mathcal{C}([t_0, t_1], \mathbb{C}^{m\times n})$, $f(t) \in \mathcal{C}([t_0, t_1], \mathbb{C}^m)$, possibly together with initial value conditions

$$x(t_0) = x_0, \ \dot{x}(t_0) = x_0^{[1]}, \quad x_0, x_0^{[1]} \in \mathbb{C}^n. \tag{3.4}$$

Using similar techniques as those employed in Section 2.3, our main objective in this section is to develop a set of invariants and a condensed form under *(global) equivalence transformations* for the triple $(M(t), C(t), K(t))$ of matrix-valued functions which satisfy certain *regularity conditions* in terms of the set of invariants.

First, let us make clear the concept of *(global) equivalence relation* between time varying systems of DAEs. Instead of the constant transformations in the case of constant coefficient system (2.11), in the case of variable coefficient system (3.3) we consider the time varying coordinate transformations given by $x(t) = Q(t)y(t)$ and premultiplication by $P(t)$, where $Q(t) \in \mathcal{C}^2([t_0, t_1], \mathbb{C}^{n\times n})$ and $P(t) \in \mathcal{C}([t_0, t_1], \mathbb{C}^{m\times m})$ are pointwise nonsingular matrix-valued functions. These changes of coordinates transform (3.3) to an *equivalent* system of DAEs

$$\begin{aligned} & \tilde{M}(t)\ddot{y}(t) + \tilde{C}(t)\dot{y}(t) + \tilde{K}(t)y(t) \\ := \ & P(t)M(t)Q(t)\ddot{y}(t) + \Big(P(t)C(t)Q(t) + 2P(t)M(t)\dot{Q}(t)\Big)\,\dot{y}(t) \\ & + \Big(P(t)K(t)Q(t) + P(t)C(t)\dot{Q}(t) + P(t)M(t)\ddot{Q}(t)\Big)\,y(t) = P(t)f(t). \end{aligned} \tag{3.5}$$

Obviously, the relation $x(t) = Q(t)y(t)$ (or $y(t) = Q^{-1}(t)x(t)$) gives a one-to-one correspondence between the two corresponding solution sets of the system (3.3) and the system (3.5). If we use the notation of triples $(M(t), C(t), K(t))$ and $(\tilde{M}(t), \tilde{C}(t), \tilde{K}(t))$ to represent the systems (3.3) and (3.5) respectively, then we can write the *equivalent relation* in terms of matrix multiplications:

$$[\tilde{M}(t), \tilde{C}(t), \tilde{K}(t)] = P(t)[M(t), C(t), K(t)] \begin{bmatrix} Q(t) & 2\dot{Q}(t) & \ddot{Q}(t) \\ 0 & Q(t) & \dot{Q}(t) \\ 0 & 0 & Q(t) \end{bmatrix}. \tag{3.6}$$

In the general case of $l$th-order system (3.1), if we make use of the notation of an $(l+1)$-tuple $(A_l(t), \ldots, A_1(t), A_0(t))$ of matrix-valued functions to represent the system (3.1), we have the following definition of *equivalence* of variable coefficient systems via time varying transformations.

**Definition 3.1** *Two $(l+1)$-tuples $(A_l(t), \ldots, A_1(t), A_0(t))$ and $(B_l(t), \ldots, B_1(t), B_0(t))$ of matrix-valued functions with $A_i(t), B_i(t) \in \mathcal{C}([t_0, t_1], \mathbb{C}^{m\times n})$, $i = 0, 1, \ldots, l$, are called* (globally) equivalent *if there are pointwise nonsingular matrix-valued functions $P(t) \in \mathcal{C}([t_0, t_1], \mathbb{C}^{m\times m})$ and $Q(t) \in \mathcal{C}^l([t_0, t_1], \mathbb{C}^{n\times n})$ such that*

$$
\begin{aligned}
&[B_l(t), \ldots, B_0(t)] \\
&= P(t)[A_l(t), \ldots, A_0(t)] \begin{bmatrix} Q(t) & \binom{l}{1}\frac{d}{dt}Q(t) & \cdots & \cdots & \binom{l}{l}\frac{d^l}{dt^l}Q(t) \\ & Q(t) & \binom{l-1}{1}\frac{d}{dt}Q(t) & \cdots & \binom{l-1}{l-1}\frac{d^{l-1}}{dt^{l-1}}Q(t) \\ & & \ddots & \ddots & \vdots \\ & & & Q(t) & \binom{1}{1}\frac{d}{dt}Q(t) \\ & & & & Q(t) \end{bmatrix},
\end{aligned} \tag{3.7}
$$

*where $\binom{j}{i} = j!/(j-i)!i!$ denotes a binomial coefficient, $i, j \in \mathbb{N}$, $i \le j$. If this is the case and the context is clear, we still write $(A_l(t), \ldots, A_1(t), A_0(t)) \sim (B_l(t), \ldots, B_1(t), B_0(t))$.*

As already suggested by the definition, the following proposition shows that relation (3.7) is an equivalence relation.

**Proposition 3.2** *Relation (3.7) introduced in Definition 3.1 is an equivalence relation on the set of $(l+1)$-tuples of matrix-valued functions.*

**Proof**. We shall show relation (3.7) has the three properties required of an equivalence relation.

1. Reflexivity: Let $P(t) = I_m$ and $Q(t) = I_n$. Then, we have $(A_l(t), \ldots, A_1(t), A_0(t)) \sim (A_l(t), \ldots, A_1(t), A_0(t))$.

2. Symmetry: Assume that $(A_l(t), \ldots, A_1(t), A_0(t)) \sim (B_l(t), \ldots, B_1(t), B_0(t))$ with pointwise nonsingular matrix-valued functions $P(t)$ and $Q(t)$ that satisfy (3.7). We shall prove that $(B_l(t), \ldots, B_1(t), B_0(t)) \sim (A_l(t), \ldots, A_1(t), A_0(t))$. Note that, from the identity $Q(t)Q^{-1}(t) = I$ it follows that any order derivative of $Q(t)Q^{-1}(t)$ with respect to $t$ is identically zero. Then, by this fact, it is immediate

to verify that

$$\begin{bmatrix} Q(t) & \binom{l}{1}\frac{d}{dt}Q(t) & \cdots & \cdots & \binom{l}{l}\frac{d^l}{dt^l}Q(t) \\ & Q(t) & \binom{l-1}{1}\frac{d}{dt}Q(t) & \cdots & \binom{l-1}{l-1}\frac{d^{l-1}}{dt^{l-1}}Q(t) \\ & & \ddots & \ddots & \vdots \\ & & & Q(t) & \binom{1}{1}\frac{d}{dt}Q(t) \\ & & & & Q(t) \end{bmatrix}$$
$$\cdot \begin{bmatrix} Q^{-1}(t) & \binom{l}{1}\frac{d}{dt}Q^{-1}(t) & \cdots & \cdots & \binom{l}{l}\frac{d^l}{dt^l}Q^{-1}(t) \\ & Q^{-1}(t) & \binom{l-1}{1}\frac{d}{dt}Q^{-1}(t) & \cdots & \binom{l-1}{l-1}\frac{d^{l-1}}{dt^{l-1}}Q^{-1}(t) \\ & & \ddots & \ddots & \vdots \\ & & & Q^{-1}(t) & \binom{1}{1}\frac{d}{dt}Q^{-1}(t) \\ & & & & Q^{-1}(t) \end{bmatrix} = I. \tag{3.8}$$

Hence, by (3.7) and (3.8), we have

$$[A_l(t), \ldots, A_0(t)] = P^{-1}(t)[B_l(t), \ldots, B_0(t)]$$
$$\cdot \begin{bmatrix} Q^{-1}(t) & \binom{l}{1}\frac{d}{dt}Q^{-1}(t) & \cdots & \cdots & \binom{l}{l}\frac{d^l}{dt^l}Q^{-1}(t) \\ & Q^{-1}(t) & \binom{l-1}{1}\frac{d}{dt}Q^{-1}(t) & \cdots & \binom{l-1}{l-1}\frac{d^{l-1}}{dt^{l-1}}Q^{-1}(t) \\ & & \ddots & \ddots & \vdots \\ & & & Q^{-1}(t) & \binom{1}{1}\frac{d}{dt}Q^{-1}(t) \\ & & & & Q^{-1}(t) \end{bmatrix},$$

namely, $(B_l(t), \ldots, B_1(t), B_0(t)) \sim (A_l(t), \ldots, A_1(t), A_0(t))$.

3. Transitivity: Assume that $(A_l(t), \ldots, A_0(t)) \sim (B_l(t), \ldots, B_0(t))$ with pointwise nonsingular matrix-valued functions $P_1(t)$ and $Q_1(t)$ and that $(B_l(t), \ldots, B_0(t)) \sim (C_l(t), \ldots, C_0(t))$ with pointwise nonsingular matrix-valued functions $P_2(t)$ and $Q_2(t)$, which satisfy (3.7), respectively. We shall prove that $(A_l(t), \ldots, A_0(t)) \sim (C_l(t), \ldots, C_0(t))$. By the product rule and Leibnitz's rule (cf. [6], p. 203) for

differentiation, we can immediately verify that

$$
\begin{aligned}
&\begin{bmatrix}
Q_1(t) & \binom{l}{1}\frac{d}{dt}Q_1(t) & \cdots & \cdots & \binom{l}{l}\frac{d^l}{dt^l}Q_1(t) \\
 & Q_1(t) & \binom{l-1}{1}\frac{d}{dt}Q_1(t) & \cdots & \binom{l-1}{l-1}\frac{d^{l-1}}{dt^{l-1}}Q_1(t) \\
 & & \ddots & \ddots & \vdots \\
 & & & Q_1(t) & \binom{1}{1}\frac{d}{dt}Q_1(t) \\
 & & & & Q_1(t)
\end{bmatrix} \\
\cdot &\begin{bmatrix}
Q_2(t) & \binom{l}{1}\frac{d}{dt}Q_2(t) & \cdots & \cdots & \binom{l}{l}\frac{d^l}{dt^l}Q_2(t) \\
 & Q_2(t) & \binom{l-1}{1}\frac{d}{dt}Q_2(t) & \cdots & \binom{l-1}{l-1}\frac{d^{l-1}}{dt^{l-1}}Q_2(t) \\
 & & \ddots & \ddots & \vdots \\
 & & & Q_2(t) & \binom{1}{1}\frac{d}{dt}Q_2(t) \\
 & & & & Q_2(t)
\end{bmatrix} \\
= &\begin{bmatrix}
Q_1(t)Q_2(t) & \binom{l}{1}\frac{d}{dt}\left(Q_1(t)Q_2(t)\right) & \cdots & \binom{l}{l}\frac{d^l}{dt^l}\left(Q_1(t)Q_2(t)\right) \\
 & Q_1(t)Q_2(t) & \cdots & \binom{l-1}{l-1}\frac{d^{l-1}}{dt^{l-1}}\left(Q_1(t)Q_2(t)\right) \\
 & & \ddots & \vdots \\
 & & \ddots & \binom{1}{1}\frac{d}{dt}\left(Q_1(t)Q_2(t)\right) \\
 & & & Q_1(t)Q_2(t)
\end{bmatrix}.
\end{aligned} \tag{3.9}
$$

Thus, by the assumptions and (3.9), we have

$$
\begin{aligned}
&[C_l(t), \ldots, C_0(t)] = P_1(t)P_2(t)[A_l(t), \ldots, A_0(t)] \\
\cdot &\begin{bmatrix}
Q_1(t)Q_2(t) & \binom{l}{1}\frac{d}{dt}\left(Q_1(t)Q_2(t)\right) & \cdots & \binom{l}{l}\frac{d^l}{dt^l}\left(Q_1(t)Q_2(t)\right) \\
 & Q_1(t)Q_2(t) & \cdots & \binom{l-1}{l-1}\frac{d^{l-1}}{dt^{l-1}}\left(Q_1(t)Q_2(t)\right) \\
 & & \ddots & \vdots \\
 & & \ddots & \binom{1}{1}\frac{d}{dt}\left(Q_1(t)Q_2(t)\right) \\
 & & & Q_1(t)Q_2(t)
\end{bmatrix},
\end{aligned}
$$

namely, $(A_l(t), \ldots, A_1(t), A_0(t)) \sim (C_l(t), \ldots, C_1(t), C_0(t))$. □

In order to introduce a set of *regularity conditions* under which we can get a condensed form via (global) equivalence transformations (3.6) for the triple $(M(t), C(t), K(t))$ in (3.3), we need the concept of *(local) equivalence relation* between two triples of matrices.

Two triples $(M, C, K)$ and $(\tilde{M}, \tilde{C}, \tilde{K})$, $M, C, K, \tilde{M}, \tilde{C}, \tilde{K} \in \mathbb{C}^{m\times n}$, of matrices are called (LOCALLY) EQUIVALENT if there exist matrices $P \in \mathbb{C}^{m\times m}$ and $Q, A, B \in \mathbb{C}^{n\times n}$, $P, Q$ nonsingular, such that

$$\tilde{M} = PMQ, \quad \tilde{C} = PCQ + 2PMA, \quad \tilde{K} = PKQ + PCA + PMB. \tag{3.10}$$

In general, we have the following definition of *(local) equivalence relation* between two tuples of matrices.

**Definition 3.3** *Two* $(l+1)$*-tuples* $(A_l,\ldots,A_1,A_0)$ *and* $(B_l,\ldots,B_1,B_0)$, $A_i, B_i \in \mathbb{C}^{m\times n}$, $i=0,1,\ldots,l$, $l\in\mathbb{N}_0$, *of matrices are called* (LOCALLY) EQUIVALENT *if there exist matrices* $P\in\mathbb{C}^{m\times m}$, $Q, R_1,\ldots,R_l\in\mathbb{C}^{n\times n}$, $P, Q$ *nonsingular, such that*

$$[B_l,\ldots,B_0]=P[A_l,\ldots,A_0]\begin{bmatrix} Q & \binom{l}{1}R_1 & \cdots & \cdots & \binom{l}{l}R_l \\ & Q & \binom{l-1}{1}R_1 & \cdots & \binom{l-1}{l-1}R_{l-1} \\ & & \ddots & \ddots & \vdots \\ & & & Q & \binom{1}{1}R_1 \\ & & & & Q \end{bmatrix}, \qquad (3.11)$$

*Again, we write* $(A_l,\ldots,A_1,A_0)\sim(B_l,\ldots,B_1,B_0)$ *if the context is clear.*

**Proposition 3.4** *Relation (3.11) introduced in Definition 3.3 is an equivalence relation on the set of* $(l+1)$*-tuples of matrices.*

**Proof**. The proof can be immediately carried out along the lines of the proof of Proposition 3.2. □

Recalling the condensed form and the invariants for matrix triples obtained under (strong) equivalence transformations in Section 2.3, we can introduce a set of invariants for matrix triples under local equivalence transformations, as the following lemma shows.

**Lemma 3.5** *Under the same assumption and the same notation as in Lemma 2.9, the quantities defined in (2.30) are invariant under the local equivalence relation (3.10) and* $(M,C,K)$ *is locally equivalent to the form* (2.28).

**Proof**. Since the strong equivalence relation (2.17) is the special case of the local equivalence relation (3.11) by setting $R_i=0$, $i=1,\ldots,l$, by Lemma 2.9, it is immediate that $(M,C,K)$ is locally equivalent to the form (2.28).

In view of the proof of Lemma 2.9, it remains to show that the quantities defined in (2.30) are invariant under the local equivalence relation (3.10). Here, again, we just take $s^{(MCK)}$ as an example. Indeed, let $(M,C,K)$ and $(\tilde{M},\tilde{C},\tilde{K})$ be locally equivalent, namely, identity (3.10) holds. Let the columns of $\tilde{Z}_1$ form a basis for $\mathcal{N}(\tilde{M}^T)$, and let the columns of $\tilde{Z}_3$ form a basis for $\mathcal{N}(\tilde{M}^T)\cap\mathcal{N}(\tilde{C}^T)$. Then, from (3.10) it follows that the columns of $Z_1:=P^T\tilde{Z}_1$ form a basis for $\mathcal{N}(M^T)$. Since, for any $z\in\tilde{Z}_3$,

$$Q^TM^TP^Tz=0,\quad Q^TC^TP^Tz+2A^TM^TP^Tz=0,$$

if and only if

$$M^TP^Tz=0,\quad C^TP^Tz=0,$$

it follows that the columns of $Z_3 := P^T\tilde{Z}_3$ form a basis for $\mathcal{N}(M^T)\cap\mathcal{N}(C^T)$. Thus, the invariance of $s^{(MCK)}$ follows from

$$\begin{aligned}
\tilde{s}^{(MCK)} &= \dim\Big(\mathcal{R}(\tilde{M}^T)\cap\mathcal{R}(\tilde{C}^T\tilde{Z}_1)\cap\mathcal{R}(\tilde{K}^T\tilde{Z}_3)\Big)\\
&= \dim\Big(\mathcal{R}(Q^TM^TP^T)\cap\mathcal{R}(Q^TC^TP^T\tilde{Z}_1+2A^TM^TP^T\tilde{Z}_1)\\
&\qquad \cap\,\mathcal{R}(Q^TK^TP^T\tilde{Z}_3+A^TC^TP^T\tilde{Z}_3+B^TM^TP^T\tilde{Z}_3)\Big)\\
&= \dim\Big(\mathcal{R}(Q^TM^TP^T)\cap\mathcal{R}(Q^TC^TP^T\tilde{Z}_1)\cap\mathcal{R}(Q^TK^TP^T\tilde{Z}_3)\Big)\\
&= \dim\Big(\mathcal{R}(M^TP^T)\cap\mathcal{R}(C^TP^T\tilde{Z}_1)\cap\mathcal{R}(K^TP^T\tilde{Z}_3)\Big)\\
&= \dim\big(\mathcal{R}(M^T)\cap\mathcal{R}(C^TZ_1)\cap\mathcal{R}(K^TZ_3)\big)\\
&= s^{(MCK)}.
\end{aligned}$$

Similarly, the invariance of the other quantities in (2.30) can be proved. □

Now, from the matrix triple $(M, C, K)$ passing to the triple $(M(t), C(t), K(t))$ of matrix-valued functions, we can calculate, based on Lemma 3.5, the characteristic quantities in (2.30) for $(M(t), C(t), K(t))$ at each fixed value $t \in [t_0, t_1]$. Then, we obtain nonnegative-integer valued functions

$$r, a, s^{(MCK)}, s^{(CK)}, d^{(1)}, s^{(MC)}, s^{(MK)}, d^{(2)}, v, u : [t_0, t_1] \to \mathbb{N}_0.$$

For the triple $(M(t), C(t), K(t))$ of matrix-valued functions, in order to derive a condensed form which is similar in form to the condensed form (2.28) for the matrix triple $(M, C, K)$, we introduce the following assumption of *regularity conditions* for the triple $(M(t), C(t), K(t))$ on $[t_0, t_1]$:

$$\begin{array}{l} r(t)\equiv r,\ a(t)\equiv a,\ s^{(MCK)}(t)\equiv s^{(MCK)},\ s^{(CK)}(t)\equiv s^{(CK)},\\ d^{(1)}(t)\equiv d^{(1)},\ s^{(MC)}(t)\equiv s^{(MC)},\ s^{(MK)}(t)\equiv s^{(MK)}. \end{array} \tag{3.12}$$

By (2.30) and (3.12), it immediately follows that $d^{(2)}(t), v(t), u(t)$ are also constant on $[t_0, t_1]$.

We can see that the regularity conditions (3.12) imply that the sizes of the blocks in the condensed form (2.28) do not depend on $t \in [t_0, t_1]$. Then, the assumption (3.12) allows for the application of the following property of a matrix-valued function with a constant rank, which may be regarded as a generalization of the property of a matrix shown in Lemma 2.6.

**Lemma 3.6** *([34], p. 71) Let $A(t) \in \mathcal{C}^l([t_0, t_1], \mathbb{C}^{m\times n})$, $l \in \mathbb{N}_0 \cup \{\infty\}$, and* rank $A(t) \equiv r$*, $r \in \mathbb{N}_0$, for all $t \in [t_0, t_1]$. Then there exist pointwise unitary (and therefore nonsingular) matrix-valued functions $U(t) \in \mathcal{C}^l([t_0, t_1], \mathbb{C}^{m\times m})$ and $V(t) \in \mathcal{C}^l([t_0, t_1], \mathbb{C}^{n\times n})$, such that*

$$U^H(t)A(t)V(t) = \begin{bmatrix} \Sigma(t) & 0 \\ 0 & 0 \end{bmatrix}, \tag{3.13}$$

*where* $\Sigma(t) \in \mathcal{C}^l([t_0, t_1], \mathbb{C}^{r \times r})$ *is nonsingular for all* $t \in [t_0, t_1]$.

Using Lemma 3.6 we can then obtain the following global condensed form for triples of matrix-valued functions via global equivalence transformations (3.6). For convenience of expression, in the following condensed form and its proof, we drop the subscripts of the blocks and omit the argument $t$ unless they are needed for clarification.

**Lemma 3.7** *Let* $M(t), C(t), K(t) \in \mathcal{C}([t_0, t_1], \mathbb{C}^{m \times n})$ *be sufficiently smooth, and suppose that the regularity conditions (3.12) hold for the local characteristic values of* $(M(t), C(t), K(t))$. *Then,* $(M(t), C(t), K(t))$ *is globally equivalent to a triple* $(\tilde{M}(t), \tilde{C}(t), \tilde{K}(t))$ *of matrix-valued functions of the following condensed form*

$$\left(\begin{bmatrix} I_{s^{(MCK)}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I_{s^{(MC)}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{s^{(MK)}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{d^{(2)}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & 0 & 0 & I_{s^{(CK)}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{d^{(1)}} & 0 & 0 \\ I_{s^{(MCK)}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I_{s^{(MC)}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \right.$$

$$\left. \begin{bmatrix} 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & K & 0 & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & 0 & 0 & I_a & 0 \\ 0 & 0 & 0 & 0 & I_{s^{(CK)}} & 0 & 0 & 0 \\ 0 & 0 & I_{s^{(MK)}} & 0 & 0 & 0 & 0 & 0 \\ I_{s^{(MCK)}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}\right) \begin{array}{l} s^{(MCK)} \\ s^{(MC)} \\ s^{(MK)} \\ d^{(2)} \\ s^{(CK)} \\ d^{(1)} \\ s^{(MCK)} \\ s^{(MC)} \\ a \\ s^{(CK)} \\ s^{(MK)} \\ s^{(MCK)} \\ u \end{array} . \tag{3.14}$$

*All blocks except the identity matrices in (3.14) are again matrix-valued functions on* $[t_0, t_1]$.

Note that $C_{5,4}(t) \equiv 0$ in (3.14) whereas $C_{5,4}$ in (2.28) may be a nonzero matrix, which is the only difference in form between condensed forms (3.14) and (2.28). This difference is due to the equivalence relation (3.5) via time varying transformations. $C_{5,4}(t) \equiv 0$ is obtained by solving an initial value problem for ordinary differential equations; see the details of the proof at the end of page 48.

**Proof**. The proof of Lemma 3.7 is given in Appendix (on page 48) to this chapter. $\square$

## 3.2 The Solution Behaviour of Higher-Order Systems of DAEs

In this section, we shall briefly discuss the solution behaviour (solvability, uniqueness of solutions and consistency of initial conditions) of higher-order systems of DAEs with variable coefficients and of the initial value problems associated with them.

First, using the results obtained in Section 3.1, we discuss the solution behaviour of the second-order system (3.3) and of its associated initial value problem (3.3)-(3.4).

For the triple $(M(t), C(t), K(t))$ of matrix-valued functions associated with (3.3), we assume that $M(t), C(t), K(t)$ are sufficiently smooth and the regularity conditions (3.12) hold. Thus, based on the condensed form (3.14) for $(M(t), C(t), K(t))$ which is obtained in Lemma 3.7, it is clear that, as in the case of constant coefficients (Section 2.4), we can write down the system of differential-algebraic equations corresponding to (3.14), and apply those differentiation-and-elimination steps (on pages 24-25) to it. Then, we can again compute the characteristic values $r, a, s^{(MCK)}, s^{(CK)}, d^{(1)}, s^{(MC)}, s^{(MK)}, d^{(2)}, v, u$ and the condensed form and then proceed inductively to the final stage. Here, the only difference of the case of variable coefficients from the constant case is that, in order to carry out the procedure to the final stage, we must assume that at every stage of the inductive procedure, the regularity conditions (3.12) hold. If this is the case, then it is immediate that we can obtain, finally, a theorem which is parallel to Theorem 2.12. From the final theorem we can directly read off the solution behaviour of (3.3) and of (3.3)-(3.4), and obtain a consequence which is parallel to Corollary 2.13. Clearly, there is no difference in form between the final theorem and Theorem 2.12 if in the former case we omit the argument $t$ in the variable coefficients, nor is there between the consequence and Corollary 2.13. Therefore, here, for the sake of space of writing we do not state them again.

In addition, it should be pointed out that, at this writing, since we do not know whether two different but globally equivalent triples of matrix-valued functions, after the differentiation-and-elimination steps are applied to them respectively, will lead to new triples with *same* characteristic values $r, a, s^{(MCK)}, s^{(CK)}, d^{(1)}, s^{(MC)}, s^{(MK)}, d^{(2)}$, $v$, and $u$, we can not guarantee that these values obtained in every step of the above inductive procedure are globally characteristic for the triple $(M(t), C(t), K(t))$.

Analogously, in the general case of higher-order systems of DAEs with variable coefficients, we can obtain a final theorem which is similar in form to Theorem 2.19, and its consequence similar to Corollary 2.20, which can show the solution behaviour of (3.1) and of (3.1)-(3.2). For the same reason, we omit them here.

**Appendix: Proof of Lemma 3.7**. By the global equivalent relation (3.6) and Lemma 3.6, we obtain the following sequence of globally equivalent triples of matrix-valued functions.

$$
\begin{aligned}
&(M, C, K) \\
&\sim \left( \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} C & C \\ C & C \end{bmatrix}, \begin{bmatrix} K & K \\ K & K \end{bmatrix} \right) \\
&\sim \left( \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} C & C \\ C & U_1^H C V_1 \end{bmatrix} + 2 \begin{bmatrix} I & 0 \\ 0 & U_1^H \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \dot{V}_1 \end{bmatrix}, \begin{bmatrix} K & K \\ K & K \end{bmatrix} \right) \\
&\sim \left( \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} C & C & C \\ C & I & 0 \\ C & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K \\ K & K & K \\ K & K & K \end{bmatrix} \right) \\
&\sim \left( \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} C & 0 & C \\ C & I & 0 \\ C & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K \\ K & K & K \\ K & K & K \end{bmatrix} \right) \\
&\sim \left( \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} C & 0 & C \\ 0 & I & 0 \\ C & 0 & 0 \end{bmatrix} + 2 \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ -\dot{C} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K \\ K & K & K \\ K & K & K \end{bmatrix} \right) \\
&\sim \left( \begin{bmatrix} V_2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} C & 0 & C \\ 0 & I & 0 \\ U_2^H C V_2 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K \\ K & K & K \\ K & K & K \end{bmatrix} \right) \\
&\sim \left( \begin{bmatrix} V_{11} & V_{12} & 0 & 0 \\ V_{21} & V_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} C & C & 0 & C \\ C & C & 0 & C \\ 0 & 0 & I & 0 \\ I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & K \\ K & K & K & K \\ K & K & K & K \\ K & K & K & K \\ K & K & K & K \end{bmatrix} \right) \\
&\sim \left( \begin{bmatrix} V_{11} & V_{12} & 0 & 0 \\ V_{21} & V_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & C \\ 0 & C & 0 & C \\ 0 & 0 & I & 0 \\ I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & K \\ K & K & K & K \\ K & K & K & K \\ K & K & K & K \\ K & K & K & K \end{bmatrix} \right) \\
&\sim \left( \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & C \\ 0 & C & 0 & C \\ 0 & 0 & I & 0 \\ I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & K \\ K & K & K & K \\ K & K & K & K \\ K & K & K & K \\ K & K & K & K \end{bmatrix} \right) \\
&\sim \left( \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & Q_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & C \\ 0 & CQ_1 + 2\dot{Q}_1 & 0 & C \\ 0 & 0 & I & 0 \\ I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & K \\ K & K & K & K \\ K & K & K & K \\ K & K & K & K \\ K & K & K & K \end{bmatrix} \right)
\end{aligned}
$$

(where pointwise nonsingular matrix-valued function $Q_1(t)$ is chosen as the solution of the initial value problem $\dot{Q}_1(t) = -\frac{1}{2}C_{2,2}(t)Q_1(t)$, $t \in [t_0, t_1]$, $Q_1(t_0)$=I)

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & C \\ 0 & 0 & 0 & C \\ 0 & 0 & I & 0 \\ I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & K \\ K & K & K & K \\ K & K & K & K \\ K & K & K & K \\ K & K & K & K \end{bmatrix} \right)$$

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & C \\ 0 & 0 & 0 & C \\ 0 & 0 & I & 0 \\ I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + 2 \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \dot{V}_3 \end{bmatrix}, \right.$$

$$\left. \begin{bmatrix} K & K & K & K \\ K & K & K & K \\ K & K & K & K \\ K & K & K & K \\ K & K & K & U_3^H K V_3 \end{bmatrix} + \begin{bmatrix} 0 & C & 0 & C \\ 0 & 0 & 0 & C \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \dot{V}_3 \end{bmatrix} + \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \ddot{V}_3 \end{bmatrix} \right)$$

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & C & C \\ 0 & 0 & 0 & C & C \\ 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & K & K \\ K & K & K & K & K \\ K & K & K & K & K \\ K & K & K & K & K \\ K & K & K & I & 0 \\ K & K & K & 0 & 0 \end{bmatrix} \right)$$

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & C & C \\ 0 & 0 & 0 & C & C \\ 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & K & I & 0 \\ K & K & K & 0 & 0 \end{bmatrix} \right)$$

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & C & C \\ 0 & 0 & 0 & C & C \\ 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} + 2 \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\dot{K} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \right.$$

$$\begin{bmatrix} K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & 0 & I & 0 \\ K & K & K & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & C & 0 & C & C \\ 0 & 0 & 0 & C & C \\ 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\dot{K} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\left. + \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\ddot{K} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right)$$

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & C & C \\ 0 & 0 & 0 & C & C \\ 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & K & K \\ K & K & K & K & K \\ K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & 0 & I & 0 \\ K & K & K & 0 & 0 \end{bmatrix} \right)$$

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & C & C \\ 0 & 0 & 0 & C & C \\ 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & 0 & I & 0 \\ K & K & K & 0 & 0 \end{bmatrix} \right)$$

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & C & C \\ 0 & 0 & 0 & C & C \\ 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & K & 0 & K \\ K & K & K & 0 & K \\ 0 & 0 & 0 & I & 0 \\ K & K & K & 0 & 0 \end{bmatrix} \right)$$

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & 0 & C & C \\ 0 & 0 & 0 & 0 & C & C \\ 0 & 0 & V_{11} & V_{12} & 0 & 0 \\ 0 & 0 & V_{21} & V_{22} & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & K & 0 & K \\ K & K & K & K & 0 & K \\ K & K & K & K & 0 & K \\ K & K & K & K & 0 & K \\ K & K & K & K & 0 & K \\ 0 & 0 & 0 & 0 & I & 0 \\ K & K & I & 0 & 0 & 0 \\ K & K & 0 & 0 & 0 & 0 \end{bmatrix} \right)$$

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & 0 & C & C \\ 0 & 0 & 0 & 0 & C & C \\ 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & I & 0 \\ K & K & I & 0 & 0 & 0 \\ K & K & 0 & 0 & 0 & 0 \end{bmatrix} \right)$$

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & 0 & C & C \\ 0 & 0 & 0 & 0 & C & C \\ 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & I & 0 \\ K & 0 & I & 0 & 0 & 0 \\ K & K & 0 & 0 & 0 & 0 \end{bmatrix} \right)$$

$$\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & 0 & 0 & C & C \\ 0 & 0 & 0 & 0 & C & C \\ 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ K & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & I & 0 & 0 & 0 \\ K & K & 0 & 0 & 0 & 0 \end{bmatrix} \right)$$

$$
\begin{aligned}
\sim \Biggl( & \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & V_{11} & V_{12} & 0 & 0 & 0 & 0 \\ 0 & V_{21} & V_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & 0 & 0 & 0 & C & C \\ 0 & 0 & 0 & 0 & 0 & C & C \\ 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} + 2 \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \dot{V}_{11} & \dot{V}_{12} & 0 & 0 & 0 & 0 \\ 0 & \dot{V}_{21} & \dot{V}_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \\
& \begin{bmatrix} K & K & K & 0 & K & 0 & K \\ K & K & K & 0 & K & 0 & K \\ K & K & K & 0 & K & 0 & K \\ K & K & K & 0 & K & 0 & K \\ K & K & K & 0 & K & 0 & K \\ K & K & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 \\ K & I & 0 & 0 & 0 & 0 & 0 \\ K & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & 0 & 0 & 0 & C & C \\ 0 & 0 & 0 & 0 & 0 & C & C \\ 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \dot{V}_{11} & \dot{V}_{12} & 0 & 0 & 0 & 0 \\ 0 & \dot{V}_{21} & \dot{V}_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\
& + \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddot{V}_{11} & \ddot{V}_{12} & 0 & 0 & 0 & 0 \\ 0 & \ddot{V}_{21} & \ddot{V}_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \Biggr)
\end{aligned}
$$

$$
\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & C & 0 & 0 & C & C \\ 0 & C & C & 0 & 0 & C & C \\ 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & K & K & 0 & K & 0 & K \\ K & K & K & 0 & K & 0 & K \\ K & K & K & 0 & K & 0 & K \\ K & K & K & 0 & K & 0 & K \\ K & K & K & 0 & K & 0 & K \\ K & K & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 \\ K & I & 0 & 0 & 0 & 0 & 0 \\ K & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right)
$$

$$
\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & C & 0 & 0 & C & C \\ 0 & C & C & 0 & 0 & C & C \\ 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 \\ K & I & 0 & 0 & 0 & 0 & 0 \\ K & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right)
$$

$$
\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 & 0 \\ -K & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} C & C & C & 0 & 0 & C & C \\ C & C & C & 0 & 0 & C & C \\ C & C & C & 0 & 0 & C & C \\ 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} + 2 \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\dot{K} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \right.
$$

$$
\left. \begin{bmatrix} K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 \\ K & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & C & 0 & 0 & 0 & 0 & 0 \\ 0 & C & 0 & 0 & 0 & 0 & 0 \\ 0 & C & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\ddot{K} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right)
$$

$$
\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 & 0 \\ -K & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} C & C & C & 0 & 0 & C & C \\ C & C & C & 0 & 0 & C & C \\ C & C & C & 0 & 0 & C & C \\ 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & C & K & 0 & K & 0 & K \\ K & C & K & 0 & K & 0 & K \\ K & C & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 \\ K & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right)
$$

$$
\sim \left( \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & C & C & 0 & 0 & C & C \\ 0 & C & C & 0 & 0 & C & C \\ 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ K & 0 & K & 0 & K & 0 & K \\ 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 \\ K & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right)
$$

$$
\sim \left( \begin{bmatrix} V_{11} & V_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ V_{21} & V_{22} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & C & C & 0 & 0 & C & C \\ 0 & 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I & 0 & 0 \\ V_{11} & V_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ V_{21} & V_{22} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right.
$$

$$
+2\begin{bmatrix}
\dot{V}_{11} & \dot{V}_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\
\dot{V}_{21} & \dot{V}_{22} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix},
\begin{bmatrix}
K & K & 0 & K & 0 & K & 0 & K \\
K & K & 0 & K & 0 & K & 0 & K \\
K & K & 0 & K & 0 & K & 0 & K \\
K & K & 0 & K & 0 & K & 0 & K \\
K & K & 0 & K & 0 & K & 0 & K \\
K & K & 0 & K & 0 & K & 0 & K \\
K & K & 0 & K & 0 & K & 0 & K \\
K & K & 0 & K & 0 & K & 0 & K \\
0 & 0 & 0 & 0 & 0 & 0 & I & 0 \\
0 & 0 & 0 & 0 & I & 0 & 0 & 0 \\
0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\
I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
$$

$$
+\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\dot{V}_{11} & \dot{V}_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\
\dot{V}_{21} & \dot{V}_{22} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
+\begin{bmatrix}
\ddot{V}_{11} & \ddot{V}_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\
\ddot{V}_{21} & \ddot{V}_{22} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}\Bigg)
$$

$$
\sim\left(\begin{bmatrix}
I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & I & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix},
\begin{bmatrix}
C & C & C & C & 0 & 0 & C & C \\
C & C & C & C & 0 & 0 & C & C \\
0 & 0 & C & C & 0 & 0 & C & C \\
0 & 0 & C & C & 0 & 0 & C & C \\
0 & 0 & 0 & 0 & I & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & I & 0 & 0 \\
I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix},
\begin{bmatrix}
K & K & 0 & K & 0 & K & 0 & K \\
K & K & 0 & K & 0 & K & 0 & K \\
K & K & 0 & K & 0 & K & 0 & K \\
K & K & 0 & K & 0 & K & 0 & K \\
K & K & 0 & K & 0 & K & 0 & K \\
K & K & 0 & K & 0 & K & 0 & K \\
K & K & 0 & K & 0 & K & 0 & K \\
K & K & 0 & K & 0 & K & 0 & K \\
0 & 0 & 0 & 0 & 0 & 0 & I & 0 \\
0 & 0 & 0 & 0 & I & 0 & 0 & 0 \\
0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\
I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}\right)
$$

$$
\sim\left(\begin{bmatrix}
I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & I & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix},
\begin{bmatrix}
0 & 0 & C & C & 0 & 0 & C & C \\
0 & 0 & C & C & 0 & 0 & C & C \\
0 & 0 & C & C & 0 & 0 & C & C \\
0 & 0 & C & C & 0 & 0 & C & C \\
0 & 0 & 0 & 0 & I & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & I & 0 & 0 \\
I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix},
\begin{bmatrix}
0 & K & 0 & K & 0 & K & 0 & K \\
0 & K & 0 & K & 0 & K & 0 & K \\
0 & K & 0 & K & 0 & K & 0 & K \\
0 & K & 0 & K & 0 & K & 0 & K \\
0 & K & 0 & K & 0 & K & 0 & K \\
0 & K & 0 & K & 0 & K & 0 & K \\
0 & K & 0 & K & 0 & K & 0 & K \\
0 & K & 0 & K & 0 & K & 0 & K \\
0 & 0 & 0 & 0 & 0 & 0 & I & 0 \\
0 & 0 & 0 & 0 & I & 0 & 0 & 0 \\
0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\
I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}\right). \quad \square
$$

# Chapter 4

# Regularity and Singularity of Matrix Polynomials

## 4.1 Introduction

Now let us turn to the second part of this thesis, in which we shall study the properties of *regularity* and *singularity* of matrix polynomials.

A polynomial with matrix coefficients is called a *matrix polynomial*, or a *polynomial matrix* if we regard it as a matrix whose elements are polynomials. It is well known that matrix polynomials play an important role in the analytical theory of elementary divisors, i.e., the theory by which a square matrix can be reduced to some normal forms (esp. the Smith canonical form and Jordan canonical form) of which important applications have been made to the analysis of differential and difference equations; see, for example, Gantmacher [14] (Chapter VI), Lancaster and Tismenetsky [36] (Chapter 7).

The motivation for our study of *regularity* and *singularity* of matrix polynomials comes mainly from two sources. One is the study of differential-algebraic equations, which is due to the close connection, as we have presented in Chapter 2, between regularity and singularity of a matrix polynomial and the properties of the solutions of the system of DAEs which is associated with the matrix polynomial; the other is the study of the polynomial eigenvalue problems:

$$A(\lambda)x = 0,\ x \neq 0; \quad y^H A(\lambda) = 0,\ y \neq 0;$$

where $A(\lambda) = \sum_{i=0}^{l} \lambda^i A_i$ is an $n \times n$ *matrix polynomial* of *degree* $l$, $A_i \in \mathbb{C}^{n\times n}$, $i = 0, 1, \ldots, l$, $A_l \neq 0$, the nonzero vector $x \in \mathbb{C}^n$ (respectively, $y \in \mathbb{C}^n$) is the right (respectively, left) eigenvector associated with the eigenvalue $\lambda$.

For *regular* matrix polynomials, a spectral theory has been well established (see Lancaster [35], Gohberg, Lancaster, and Rodman [17], and Lancaster and Tismenetsky

[36], Chapter 14). This theory allows us to solve, at least theoretically, many polynomial eigenvalue problems which arise not only as the underlying algebraic problems of the analysis and numerical solution of higher order systems of ordinary differential equations and difference equations, but also in linear algebra problems (for example, the constrained least squares problems and signal processing problems; cf. Meerbergen and Tisseur [44]). However, for *non-regular*, or in another term, *singular* matrix polynomials, especially those of *degrees* greater than or equal to 2, the general theoretical analysis has been largely ignored; see, for example, the concluding remark by P. Van Dooren and P. Dewilde [59] (pp. 575-578).

Traditionally, for polynomial eigenvalue problems, especially those of *degrees* greater than or equal to 2, most research results including spectral analysis, canonical forms, linearization, Jordan pairs, etc., and numerical methods such as numerical algorithms, model reduction, and perturbation analysis (conditioning, backward error, pseudospectra), etc., are mainly based on the *regularity assumption* that *the matrix polynomial* $A(\lambda)$ *is regular*, namely, *it is square and its determinant* $\det(A(\lambda))$ *is not identically equal to zero*. For more details, see, for example, [17] and [44]. There are two major reasons for the regularity assumption. The first is that the regular case frequently occurs in applications. Take, for example, the quadratic eigenvalue problem associated with a gyroscopic system (cf. [44] and the references therein):

$$Q(\lambda) = \lambda^2 M + \lambda C + K,$$

where $M, C, K \in \mathbb{C}^{n\times n}$, $M = M^H$ positive definite, $C = -C^H$, and $K = K^H$. Since the leading coefficient matrix $M$ is nonsingular, the determinant of the quadratic matrix polynomial $Q(\lambda)$ is a polynomial in $\lambda$ of degree $2n$, and therefore $Q(\lambda)$ is *regular*. Such *regular* polynomial eigenvalue problems with a nonsingular leading coefficient matrix frequently arise from the analysis of structural mechanical and acoustic systems, electrical circuit simulation, fluid mechanics, and modelling microelectronics mechanical systems; see [44] and the references therein. The second reason for the regularity assumption is that the study of regular matrix polynomials clearly shows the main features of spectral theory. Take, for instance, the monograph of Lancaster [35], as well as that of Gohberg, Lancaster, and Rodman [17], which has regular matrix polynomials as its whole subject.

However, there are applications from which *singular* polynomial eigenvalue problems of *degrees* greater than or equal to 2, not to mention *singular* generalized eigenvalue problems, arise, as the following examples show.

**Example 4.1 (Signal processing)** *[44]* Consider the symmetric quadratic eigenvalue problem

$$A(\lambda)v = (\lambda^2 A_2 + \lambda A_1 + A_0)v = 0,$$

where $A_0, A_1, A_2 \in \mathbb{R}^{(p+1)\times(p+1)}$, $p \in \mathbb{Z}$, and

$$A_0 = A_0^T \geq 0, \quad A_1 = A_1^T, \quad A_2 = \begin{bmatrix} 0 & 0 \\ 0 & I_p \end{bmatrix}.$$

Since the leading coefficient matrix $A_2$ is singular and the last coefficient matrix $A_0$ may also be singular, the determinant of the quadratic matrix polynomial $A(\lambda)$ may be identically equal to zero. Therefore, $A(\lambda)$ may be a *singular* matrix polynomial. $\diamondsuit$

**Example 4.2 (Vibration of rail tracks)** *[41]* Consider the complex quadratic eigenvalue problem

$$\frac{1}{\kappa}(M_1^T + \kappa M_0 + \kappa^2 M_1)y = 0,$$

where $\kappa \neq 0$, $M_1, M_0 \in \mathbb{C}^{n\times n}$, and $M_1$ is singular. Since the leading coefficient matrix $M_1$ and the last $M_1^T$ of the corresponding matrix polynomial $A(\lambda) := \lambda^2 M_1 + \lambda M_0 + M_1^T$ are singular, $A(\lambda)$ may be *singular*. $\diamondsuit$

In addition, although the study of *singular* matrix pencils, which can be regarded as matrix polynomials of degree 1, has a long history (see, for example, Gantmacher [15], Chapter XII), some related theoretical and numerical aspects have not yet been completely clarified or solved, such as geometrical characterization of *singular* matrix pencils (we shall return to this topic in Subsection 4.2.4), detecting the regularity or singularity, and the nearness to singularity problem for *regular* matrix pencils (see Byers, He, and Mehrmann [4]).

Thus, from a theoretical and/or numerical point of view, the following tasks naturally arise:

1 To obtain characterizations of the *regularity* and *singularity* of matrix polynomials.

2 To detect whether or not a given matrix polynomial is *regular*.

3 To find a solution of or a useful characterization for the nearness to singularity problem for a *regular* matrix polynomial.

The investigations of the above tasks will be carried out in this chapter. In Section 4.2 we present sufficient and necessary conditions for the *singularity* and *regularity* of matrix polynomials, which lay a theoretical foundation for the investigations conducted in the subsequent sections 4.3 and 4.4. In addition, we will present a simple sufficient and necessary geometrical characterization of the *column-singularity* of rectangular matrix pencils, as well as a canonical form, under equivalence transformations (2.17), for $2 \times 2$ *singular* quadratic matrix polynomials. In Section 4.3 we will present a natural approach to detect the regularity or singularity of a given square matrix polynomial via the rank information of its coefficient matrices. At last, Section 4.4 deals with the nearness to singularity problem for square and regular matrix polynomials. We will give a definition, some general properties, and theoretical characterizations of the nearest distance to singularity, and derive two types of lower bounds on the nearest distance.

## 4.2 Sufficient and Necessary Conditions for Regular and Singular Matrix Polynomials

In this section, after giving definitions and examples and stating our main idea in Subsection 4.2.1, we will present in Subsection 4.2.2 the main contribution of our investigation – Theorem 4.32 – which describes sufficient and necessary conditions for the *singularity* of matrix polynomials. In Subsection 4.2.3, Theorem 4.32 will lead to corollaries which give sufficient and necessary conditions for the *singularity* and *regularity* of matrix polynomials. Finally, Subsection 4.2.4 deals with *column-singular* matrix pencils and $2 \times 2$ *singular* quadratic matrix polynomials.

### 4.2.1 Definitions and Main Idea

To set notation, we begin with the definition of *matrix polynomials*.

**Definition 4.3** *A* MATRIX POLYNOMIAL $A(\lambda)$ *over* $\mathbb{C}$ (or $\mathbb{R}$) *is a polynomial in* $\lambda$ *with matrix coefficients:*

$$A(\lambda) = \sum_{i=0}^{l} \lambda^i A_i = \lambda^l A_l + \lambda^{l-1} A_{l-1} + \cdots + \lambda A_1 + A_0, \tag{4.1}$$

*where* $\lambda \in \mathbb{C}$ *and the matrices* $A_i$, $i = 1, \ldots, l$, *are from* $\mathbb{C}^{m\times n}$ (or $\mathbb{R}^{m\times n}$).

*If* $m = n$, *then the matrix polynomial* $A(\lambda)$ *is called* SQUARE, *and the number* $n$ *is called the* ORDER *of the matrix polynomial.*

*The number* $l$ *is called the* DEGREE *of the matrix polynomial if* $A_l \neq 0$.

*If* $m = 1$ (*respectively,* $n = 1$), *then the matrix polynomial* $A(\lambda)$ *is also called a* ROW- (*respectively,* COLUMN-) VECTOR POLYNOMIAL.

**Remark 4.4** We may represent the matrix polynomial $A(\lambda)$ in the form of a POLYNOMIAL MATRIX , i.e., in the form of an $m \times n$ matrix whose elements are polynomials in $\lambda$:

$$A(\lambda) = [a_{i,j}(\lambda)]_{i,j=1}^{m,n} = \left[ a_{i,j}^{(l)}\lambda^l + a_{i,j}^{(l-1)}\lambda^{l-1} + \cdots + a_{i,j}^{(0)} \right]_{i,j=1}^{m,n},$$

where $l$ is the degree of the matrix polynomial. If $m = 1$ (respectively, $n = 1$), then the polynomial matrix $A(\lambda)$ is also called a POLYNOMIAL ROW- (respectively, COLUMN-) VECTOR. $\diamondsuit$

In order to introduce the concepts of *regularity* and *singularity* of matrix polynomials, we need the concepts of the *minor* and *rank* of a matrix polynomial which are natural generalizations of those of the minor and rank of a matrix, as follows.

**Definition 4.5** *Let $A(\lambda)$ be an $m \times n$ rectangular matrix polynomial. A* MINOR OF ORDER *$p$ ($1 \leq p \leq \min(m, n)$) of $A(\lambda)$ is defined to be the determinant of a $p \times p$ sub-matrix polynomial of $A(\lambda)$ obtained from $A(\lambda)$ by striking out $m-p$ rows and $n-p$ columns. If the retained rows and columns are given by subscripts*

$$1 \leq i_1 < i_2 < \cdots < i_p \leq m, \quad 1 \leq j_1 < j_2 < \cdots < j_p \leq n,$$

*respectively, then the corresponding $p$-th-order minor is denoted by*

$$A(\lambda)\begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix} := \det\left[a_{i_k, j_k}(\lambda)\right]_{k=1}^{p}.$$

**Definition 4.6** *([14], p.139; [36], p.259) An integer $r$ is said to be the* RANK *of a matrix polynomial if it is the order of its largest minor that is not identically equal to zero.*

**Remark 4.7** By Definition 4.6, the rank $r$ of a matrix polynomial $A(\lambda)$ can be represented as:

$$r = \operatorname{rank}(A(\lambda)) = \max_{\nu \in \mathbb{C}} \operatorname{rank}(A(\nu)).$$

Obviously, $r \leq \min\{m, n\}$. $\diamondsuit$

**Example 4.8** We consider the following matrix polynomials:

$$(a) \quad \text{Let } A_1(\lambda) := \lambda \begin{bmatrix} 1 & 2 \\ 3 & 0 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \lambda+1 & 2\lambda \\ 3\lambda & 0 \\ \lambda & \lambda \end{bmatrix}.$$

Then, $\operatorname{rank}(A_1(\lambda)) = 2$.

$$(b) \quad \text{Let } A_2(\lambda) := \lambda^2 \begin{bmatrix} 2 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 2 \\ -2 & 1 \\ -1 & 1 \end{bmatrix} + \begin{bmatrix} -2 & 2 \\ 1 & -1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2(\lambda+1)(\lambda-1) & 2(\lambda+1) \\ (\lambda-1)^2 & \lambda-1 \\ \lambda(\lambda-1) & \lambda \end{bmatrix}.$$

Then, $\operatorname{rank}(A_2(\lambda)) = 1$.

$(c)$ Let $A_3(\lambda) := \lambda \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 2\lambda+1 \\ 0 & \lambda+2 \end{bmatrix}.$

Then, $\operatorname{rank}(A_3(\lambda)) = 1.$

$(d)$ Let $A_4(\lambda) := \lambda^2 \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2\lambda^2 & 2\lambda \\ \lambda & 1 \end{bmatrix}.$

Then, $\operatorname{rank}(A_4(\lambda)) = 1.$

$\Diamond$

The idea involved in our investigation is in essence quite simple and natural, and can be regarded as a direct generalization of the corresponding concept for matrices. Let us first recall the definition and a basic property of *regularity* and *singularity* of matrices.

**Definition 4.9** *A matrix $A \in \mathbb{C}^{m\times n}$ is said to be* COLUMN-REGULAR, *or to have* FULL COLUMN RANK, *if* $\operatorname{rank}(A) = n$; *otherwise, it is said to be* COLUMN-SINGULAR, *or* COLUMN-RANK DEFICIENT. *$A$ is said to be* ROW-REGULAR, *or to have* FULL ROW RANK, *if* $\operatorname{rank}(A) = m$; *otherwise, it is said to be* ROW-SINGULAR, *or* ROW-RANK DEFICIENT.

**Proposition 4.10** *A matrix $A \in \mathbb{C}^{m\times n}$ is column-regular (respectively, column-singular) if and only if its conjugate transpose $A^H \in \mathbb{C}^{n\times m}$ (or transpose $A^T \in \mathbb{C}^{n\times m}$) is row-regular (respectively, row-singular).*

Similarly, we give the following definition and propositions of *regularity* and *singularity* of matrix polynomials.

**Definition 4.11** *Let $A(\lambda) = \sum_{i=0}^{l} \lambda^i A_i$ be an $m \times n$ matrix polynomial of degree $l$, where $m, n \in \mathbb{N}$, $l \in \mathbb{N}_0$, $A_i \in \mathbb{C}^{m\times n}$, $i = 0, \ldots, l$. Let $r$ be the rank of $A(\lambda)$.*

*If $r < n$, then $A(\lambda)$ is said to be* COLUMN-SINGULAR; *otherwise, if $r = n$ then the matrix polynomial is said to be* COLUMN-REGULAR.

*If $r < m$, then $A(\lambda)$ is said to be* ROW-SINGULAR; *otherwise, if $r = m$, then $A(\lambda)$ is said to be* ROW-REGULAR.

*If $A(\lambda)$ is both column- and row-regular, i.e., $r = m = n$, then it is said to be* REGULAR; *if $r < m = n$, then it is said to be* SINGULAR.

**Remark 4.12** In order to be consistent with the concepts of the regularity and singularity of matrix pencils (cf. Section 2.4 and Gantmacher [15], p. 25, Def. 2), we always call a non-square matrix polynomial SINGULAR, though by Definition 4.11 it may be column-regular or row-regular. $\Diamond$

By Definition 4.11, it is clear that a square matrix polynomial is column-regular if and only if it is row-regular. Proceeding from Definition 4.11, the following proposition describes, in terms of determinants, the regularity and singularity of a square matrix polynomial.

**Proposition 4.13** *Let* $A(\lambda) = \sum_{i=0}^{l} \lambda^i A_i$ *be an* $n \times n$ *matrix polynomial of degree* $l$, *where* $n \in \mathbb{N}$, $l \in \mathbb{N}_0$, $A_i \in \mathbb{C}^{n\times n}$, $i = 0, \ldots, l$. *Then* $A(\lambda)$ *is singular if and only if*

$$\forall \lambda \in \mathbb{C}: \ \det(A(\lambda)) = \det(\lambda^l A_l + \lambda^{l-1} A_{l-1} + \cdots + A_0) = 0.$$

**Proof**. By Definition 4.11, the proof follows from the fact that

$$\begin{aligned} & A(\lambda) \ \text{ is singular} \\ \Longleftrightarrow\ & \operatorname{rank}(A(\lambda)) = \max_{\nu\in\mathbb{C}} \operatorname{rank}(A(\nu)) < n \\ \Longleftrightarrow\ & \forall \lambda \in \mathbb{C}: \ \det(A(\lambda)) = \det(\lambda^l A_l + \lambda^{l-1} A_{l-1} + \cdots + A_0) = 0. \end{aligned}$$

$\square$

In Example 4.8, we can see by Definition 4.11 that the matrix polynomial $A_1(\lambda)$ is column-regular and at the same time row-singular, $A_2(\lambda)$ is not only column-singular but row-singular, and $A_i(\lambda)$, $i = 3, 4$, are singular.

**Remark 4.14** From the point of view of polynomial eigenvalue problems, most of the related literature agree on the definition of singularity of square matrix polynomials described in Proposition 4.13, which is in essence consistent with Definition 4.11; see, for example, Gantmacher [15] (Chapter XII), Gohberg, Lancaster, and Rodman[17], Lancaster and Tismenetsky [36] (Chapter 7 and 14), Van Dooren and Dewilde [59], and Meerbergen and Tisseur [44], etc. However, form the point view of matrix polynomials themselves, there may be different definitions of regularity and singularity of matrix polynomials with respect to different objectives of study. Take, for example, the one in Gantmacher [14] (Chapter IV), namely, a matrix polynomial is called *regular* if it is square and its leading coefficient matrix is nonsingular. This definition, which is a special case of Definition 4.11 and used in [14] in studying the *right and left division* of matrix polynomials, also appears in the earlier monograph of Lancaster [35] which concentrates on the study of polynomial eigenvalue problems that arise in dynamic vibrating systems. In this chapter, since the motivation for our study proceeds mainly from polynomial eigenvalue problems which arise in a wide variety of applications, we naturally prefer Definition 4.11 to the one in [14] and [35]. $\Diamond$

From Definition 4.6 it immediately follows that the rank of a matrix polynomial is equal to that of its conjugate transpose (or of its transpose). Hence, the following statement is plain.

**Proposition 4.15** *A matrix polynomial $A(\lambda)$ of size $m \times n$ is column-regular (respectively, column-singular) if and only if its conjugate transpose $A^H(\lambda)$ (or transpose $A^T(\lambda)$) is row-regular (respectively, row-singular) .*

Thus, throughout this chapter we shall mainly concentrate our discussion on the column-regularity and column-singularity of rectangular matrix polynomials, since the results of the column cases can be parallelly extended by Proposition 4.15 to those of the row cases.

Now let us recall a fundamental result of the relation between a linear subspace and its matrix representation whose columns (or rows) span the whole linear subspace.

**Proposition 4.16** *([36], pp.93-4) A matrix $A \in \mathbb{C}^{m\times n}$ is column-singular if and only if its column vectors are linearly dependent in $\mathbb{C}^m$, or in other words, there exists a nonzero vector $x \in \mathbb{C}^n$ such that $Ax = 0$.*

Along the same lines, our main idea is to prove that a matrix polynomial $A(\lambda)$ is column-singular if and only if there exists a vector polynomial $x(\lambda)$, which is not identically equal to zero, such that $A(\lambda)x(\lambda) = 0$. As far as singular matrix pencils are concerned, in Chapter XII of [15], in order to prove the Kronecker canonical form, it is asserted that the equation $(A_1\lambda + A_0)x(\lambda) = 0$ has a nonzero vector polynomial $x(\lambda)$ as its solution for any given singular matrix pencil $A_1\lambda + A_0$. However, this result is not proved; see [15] p.29. In the next subsection, we shall generalize this statement to the cases of higher-degree singular matrix polynomials and prove it. More important, we shall give an attainable upper bound on the least possible degree of such nonzero vector polynomials $x(\lambda)$. The upper bound is expressed in terms of the rank, or the minimum of the row and column dimensions, of the related singular matrix polynomial, and conversely, it can be regarded as a criterion for determining or detecting in finite steps whether a given matrix polynomial is column-regular or column-singular.

### 4.2.2 Sufficient and Necessary Conditions for Singularity I

For any given column-singular matrix polynomial $A(\lambda)$, our main objective in this section is to construct a nonzero vector polynomial $x(\lambda)$ such that $A(\lambda)x(\lambda) = 0$ and, at the same time to obtain an upper bound on the least possible degree of such $x(\lambda)$. In order to achieve this objective, we will conduct our investigation in three stages. First, we will reduce, via *left-equivalence transformations*, a given matrix polynomial $A(\lambda)$ to a matrix polynomial $B(\lambda)$ in an upper-triangular form. Then, we will estimate upper bounds of the degrees of some minors of the upper-triangular matrix polynomial $B(\lambda)$. Finally, we will construct by Cramer's rule a nonzero vector polynomial $x(\lambda)$ such that $B(\lambda)x(\lambda) = 0$, which is equivalent to $A(\lambda)x(\lambda) = 0$, and we will, based on the estimates presented in the second stage, give an attainable upper bound to the least possible degree of such $x(\lambda)$.

Since the main results required in the first stage have been obtained in the literature (cf., for example, [14] (Chapter VI), [36] (Chapter 7)), here, for the sake of completeness and coherence, we just present a review of some definitions and results which will be used in the later stages.

We start by defining *elementary row* and *column operations* for a matrix polynomial and the corresponding *elementary matrix polynomials* that are to be applied to reduce the matrix polynomial to an upper- (or lower-) triangular form by means of *equivalence transformations.*

**Definition 4.17 (Elementary operations)** *([14], pp.130-1; [36], p.253) The following operations are referred to as* ELEMENTARY ROW *and* COLUMN OPERATIONS *on a matrix polynomial:*

(*1*) *multiply any row* (*column*) *by a nonzero* $c \in \mathbb{C}$*;*

(*2*) *interchange any two rows* (*columns*)*;*

(*3*) *add to any row* (*column*) *any other row* (*column*) *multiplied by an arbitrary polynomial* $b(\lambda)$ *over* $\mathbb{C}$*,*

*where row elementary operations are also called* LEFT *elementary operations, and column elementary operations are also called* RIGHT *elementary operations.*

**Definition 4.18 (Elementary matrix polynomials)** *([14], pp.131-2; [36], pp.254-5) The following square matrix polynomials* $E^{(1)}$*,* $E^{(2)}$*, and* $E^{(3)}(\lambda)$ *are called* ELEMENTARY MATRIX POLYNOMIALS OF TYPES 1, 2, AND 3*, respectively:*

$$E^{(1)} = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & c & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}, \quad E^{(2)} = \begin{bmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 0 & & \cdots & 1 & & \\ & & & 1 & & & & \\ & & \vdots & & \ddots & \vdots & & \\ & & & & 1 & & & \\ & & 1 & & \cdots & 0 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{bmatrix},$$

$$E^{(3)}(\lambda) = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \cdots & b(\lambda) & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \quad or \quad E^{(3)}(\lambda) = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & \vdots & \ddots & & & \\ & & b(\lambda) & \cdots & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix},$$

*where $c \in \mathbb{C}$, $c \neq 0$, and $b(\lambda)$ is an arbitrary polynomial over $\mathbb{C}$.*

**Remark 4.19** The elementary matrix polynomials of types 1 and 2 are also called ELEMENTARY MATRICES OF TYPES 1 and 2; as for the elementary matrix polynomials of types 3, if $b(\lambda)$ is a nonzero number, then $E^{(3)}$ degenerates into an ELEMENTARY MATRIX OF TYPE 3. ♢

**Remark 4.20** From Definitions 4.17 and 4.18 it immediately follows that performing an elementary row (respectively, column) operation on an $m \times n$ matrix polynomial is equivalent to premultiplication (respectively, postmultiplication) of the matrix polynomial by an $m \times m$ (respectively, $n \times n$) elementary matrix polynomial of the corresponding type. ♢

Some simple but important properties of elementary matrix polynomials should be mentioned, such as: *the determinant of every elementary matrix polynomial is a nonzero constant*, and, *the inverse of every elementary matrix polynomial is also an elementary matrix polynomial*. Here, a matrix polynomial $B(\lambda)$ is called the *inverse* of a matrix polynomial $A(\lambda)$, and vice versa, if $B(\lambda)A(\lambda) = A(\lambda)B(\lambda) = I$.

**Definition 4.21** *([36], p.247.) An $n \times n$ square matrix polynomial $A(\lambda)$ with nonzero constant determinant is referred to as* UNIMODULAR *matrix polynomial.*

Clearly, from Definition 4.21 it follows that elementary matrix polynomials are unimodular.

Making use of the concepts of left and right elementary operations, we can give the following definition of *equivalence transformations.*

**Definition 4.22** *([14], p.132.) Two $m \times n$ matrix polynomials $A(\lambda)$ and $B(\lambda)$ are said to be* 1) LEFT-EQUIVALENT, 2) RIGHT-EQUIVALENT, 3) EQUIVALENT, *or to be connected by* 1) *a* LEFT-EQUIVALENCE TRANSFORMATION, 2*) a* RIGHT-EQUIVALENCE TRANSFORMATION, 3) *an* EQUIVALENCE TRANSFORMATION, *if one of them can be obtained from the other by means of a finite sequence of* 1) *left elementary* 2) *right elementary,* 3) *left or right elementary operations, respectively.*

It follows from the equivalence of elementary operations and operations with elementary matrix polynomials that, $A(\lambda)$ and $B(\lambda)$ are 1) left-equivalent, 2) right-equivalent, 3) equivalent, if and only if there are elementary matrix polynomials $E_1(\lambda)$, $E_2(\lambda)$, ..., $E_k(\lambda)$, $E_{k+1}(\lambda)$, ..., $E_s(\lambda)$ such that 1) $B(\lambda) = P(\lambda)A(\lambda)$, 2) $B(\lambda) = A(\lambda)Q(\lambda)$, 3) $B(\lambda) = P(\lambda)A(\lambda)Q(\lambda)$, respectively, where $P(\lambda) = E_k(\lambda)\cdots E_1(\lambda)$, and $Q(\lambda) = E_{k+1}(\lambda)\cdots E_s(\lambda)$.

**Proposition 4.23** *([36], pp.259-60.) The rank of a matrix polynomial is invariant under* a) *left-equivalence,* b) *right-equivalence, or* c) *equivalence transformations, respectively.*

**Proof**. The result follows directly from Definition 4.6 of the rank of matrix polynomials and the fact that the determinant of every elementary matrix polynomial is a nonzero constant. □

**Proposition 4.24** *The column-regularity (or column-singularity, or row-regularity, or row-singularity) of a matrix polynomial is invariant under* a) *left-equivalence,* b) *right-equivalence, or* c) *equivalence transformations, respectively.*

**Proof**. The result follows directly from Definition 4.11 and Proposition 4.23. □

Through left elementary operations only, we may reduce a rectangular matrix polynomial $A(\lambda)$ to an upper-triangular form $B(\lambda)$ as described in the following theorem. In the second stage of our investigation we will estimate the degrees of some minors of $B(\lambda)$. These minors will be used in the third stage to construct a nonzero vector polynomial $x(\lambda)$ which satisfies the equation $B(\lambda)x(\lambda) = 0$, where $B(\lambda) = P(\lambda)A(\lambda)$ and $A(\lambda)$ is column-singular. Since $P(\lambda)$ is invertible, the equation $B(\lambda)x(\lambda) = 0$ is equivalent to the equation $A(\lambda)x(\lambda) = 0$. Here, the reason why we apply only left-equivalence transformation to matrix polynomials is that in this way we can, in the third stage, estimate the degree of $x(\lambda)$ which remains unaltered, using only left-equivalence transformations of $A(\lambda)$, from $B(\lambda)x(\lambda) = 0$ to $A(\lambda)x(\lambda) = 0$.

**Theorem 4.25** *([14], p.135; [36], p.259) An arbitrary $m \times n$ rectangular matrix polynomial $A(\lambda) = [a_{i,j}(\lambda)]_{i,j=1}^{m,n}$ of degree $l$ can be transformed via left elementary operations into an $m \times n$ upper-triangular matrix polynomial $B(\lambda)$ that is described in the following form (4.2), where the polynomials $b_{1,j}(\lambda)$, $b_{2,j}(\lambda)$, $\ldots$, $b_{j-1,j}(\lambda)$ are of degree less than that of $b_{j,j}(\lambda)$, provided $b_{j,j}(\lambda)$ is not identically equal to zero, and all are identically equal to zero if $b_{j,j} = const. \neq 0$ $(j = 2, 3, \ldots, \min(m, n))$.*

$$\underset{(m \leq n)}{\begin{bmatrix} b_{1,1}(\lambda) & b_{1,2}(\lambda) & \cdots & \cdots & b_{1,m}(\lambda) & \cdots & b_{1,n}(\lambda) \\ 0 & b_{2,2}(\lambda) & \cdots & \cdots & b_{2,m}(\lambda) & \cdots & b_{2,n}(\lambda) \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & b_{m,m}(\lambda) & \cdots & b_{m,n}(\lambda) \end{bmatrix}}, \quad \underset{(m \geq n)}{\begin{bmatrix} b_{1,1}(\lambda) & b_{1,2}(\lambda) & \cdots & b_{1,n}(\lambda) \\ 0 & b_{2,2}(\lambda) & \cdots & b_{2,n}(\lambda) \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & b_{n,n}(\lambda) \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}}. \tag{4.2}$$

**Proof**. See, for example, Gantmacher [14], pp. 134-135. □

As an application of Theorem 4.25, the following corollary sheds light on the nature of unimodular matrix polynomials.

**Corollary 4.26** *([14], p.136; [36], p.256) An $n \times n$ square matrix polynomial $A(\lambda)$ is unimodular if and only if it can be decomposed into a product of elementary matrices.*

**Proof**. See, [14], p. 136. □

**Remark 4.27** *([14], p.133)* By Corollary 4.26 we can restate the equivalence of matrix polynomials in terms of unimodular matrix polynomials, as follows.

*Two rectangular matrix polynomials $A(\lambda)$ and $B(\lambda)$ are* 1) LEFT-EQUIVALENT, 2) RIGHT-EQUIVALENT, 3) EQUIVALENT, *if and only if* 1) $B(\lambda) = P(\lambda)A(\lambda)$, 2) $B(\lambda) = A(\lambda)Q(\lambda)$, 3) $B(\lambda) = P(\lambda)A(\lambda)Q(\lambda)$, *respectively, where $P(\lambda)$ and $Q(\lambda)$ are unimodular matrix polynomials. Moreover, all these equivalences between matrix polynomials are equivalence relations.* ♢

In order to estimate the degrees of some minors of matrix polynomial $B(\lambda)$ in (4.2), we need the following lemmas.

**Lemma 4.28** *Let $A(\lambda)$ be an $m \times n$ rectangular matrix polynomial of degree $l$, and let $A(\lambda)$ be left-equivalent to an $m \times n$ matrix polynomial $B(\lambda)$. If*

$$B(\lambda) = \begin{bmatrix} B_{1,1}(\lambda) & B_{1,2}(\lambda) \\ 0 & B_{2,2}(\lambda) \end{bmatrix}, \tag{4.3}$$

*where the polynomial matrix $B_{1,1}(\lambda)$ is of dimension $p \times p$, $p \in \mathbb{N}$, $1 \leq p \leq \min\{m, n\}$, and $\det(B_{1,1}(\lambda))$ is not identically equal to zero, then*

$$\deg\left(\det(B_{1,1}(\lambda))\right) \leq pl. \tag{4.4}$$

**Proof**. Since the matrix polynomial $A(\lambda)$ is left-equivalent to $B(\lambda)$, by Remark 4.27, there exists a unimodular matrix polynomial $P(\lambda)$, such that

$$A(\lambda) = P(\lambda)B(\lambda). \tag{4.5}$$

Let $A(\lambda) = [A_1(\lambda), A_2(\lambda)]$, where the matrix polynomial $A_1(\lambda)$ is of dimension $m \times p$. Then, by (4.5) and (4.3), we have

$$A_1(\lambda) = P(\lambda) \begin{bmatrix} B_{1,1}(\lambda) \\ 0 \end{bmatrix}, \tag{4.6}$$

namely, $A_1(\lambda)$ is left-equivalent to $[B_{1,1}(\lambda), 0]^T$. Since $\det(B_{1,1}(\lambda))$ is not identically equal to zero, by Definition 4.6, we have

$$\mathrm{rank}([B_{1,1}(\lambda), 0]^T) = \mathrm{rank}(B_1(\lambda)) = p. \tag{4.7}$$

Because $A_1(\lambda)$ is left-equivalent to $[B_{1,1}(\lambda), 0]^T$, by Proposition 4.23 and (4.7), we have

$$\mathrm{rank}(A_1(\lambda)) = p. \tag{4.8}$$

By Definition 4.6, there exists a permutation matrix $E$ of dimension $m \times m$, such that the leading principal submatrix $\tilde{A}_{1,1}(\lambda)$ of $EA_1(\lambda)$ has full rank $p$, where $\tilde{A}_{1,1}(\lambda)$ is of dimension $p \times p$. Hence, by (4.6), we get

$$EA_1(\lambda) = \begin{bmatrix} \tilde{A}_{1,1}(\lambda) \\ \tilde{A}_{2,1}(\lambda) \end{bmatrix} = EP(\lambda) \begin{bmatrix} B_{1,1}(\lambda) \\ 0 \end{bmatrix}, \tag{4.9}$$

where $\tilde{A}_{2,1}(\lambda)$ is of dimension $(m-p) \times p$. We rewrite (4.9) in the following form:

$$\begin{bmatrix} \tilde{A}_{1,1}(\lambda) \\ \tilde{A}_{2,1}(\lambda) \end{bmatrix} = \begin{bmatrix} \tilde{P}_{1,1}(\lambda) & \tilde{P}_{1,2}(\lambda) \\ \tilde{P}_{2,1}(\lambda) & \tilde{P}_{2,2}(\lambda) \end{bmatrix} \begin{bmatrix} B_{1,1}(\lambda) \\ 0 \end{bmatrix}, \tag{4.10}$$

where $EP(\lambda) = \begin{bmatrix} \tilde{P}_{1,1}(\lambda) & \tilde{P}_{1,2}(\lambda) \\ \tilde{P}_{2,1}(\lambda) & \tilde{P}_{2,2}(\lambda) \end{bmatrix}$, and the matrix polynomials $\tilde{P}_{1,1}(\lambda)$ is of dimension $p \times p$; then we have

$$\tilde{A}_{1,1}(\lambda) = \tilde{P}_{1,1}(\lambda) B_{1,1}(\lambda), \tag{4.11}$$

and therefore, it follows that

$$\det(\tilde{A}_{1,1}(\lambda)) = \det(\tilde{P}_{1,1}(\lambda)) \det(B_{1,1}(\lambda)). \tag{4.12}$$

Note that $\tilde{A}_{1,1}(\lambda)$ has full rank, or in other words, $\det(\tilde{A}_{1,1}(\lambda))$ is a polynomial in $\lambda$ which is not identically equal to zero. By (4.12), it follows that $\det(\tilde{P}_{1,1}(\lambda))$ is not identically equal to zero, and therefore also

$$0 \leq \deg\left(\det(B_{1,1}(\lambda))\right) \leq \deg\left(\det(\tilde{A}_{1,1}(\lambda))\right). \tag{4.13}$$

Since $EA_1(\lambda)$ is obtained from $A_1(\lambda)$ through interchanges of rows, every entry of the submatrix polynomial $\tilde{A}_{1,1}(\lambda)$ of $EA_1(\lambda)$ is either a polynomial in $\lambda$ with its degree less than or equal to $l$, or zero. Thus, we have

$$\deg\left(\det(\tilde{A}_{1,1}(\lambda))\right) \leq pl. \tag{4.14}$$

And finally, from (4.13) and (4.14), it follows that $\deg\left(\det(B_{1,1}(\lambda))\right) \leq pl$. $\square$

**Lemma 4.29** *Let $B(\lambda)$ be left-equivalent to an $m \times n$ matrix polynomial $A(\lambda)$ of degree of $l$, and let $E = [e_{j_1}, e_{j_2}, \ldots, e_{j_n}]$ be a permutation matrix. If*

$$B(\lambda)E = \begin{bmatrix} \hat{B}_{1,1}(\lambda) & \hat{B}_{1,2}(\lambda) \\ 0 & \hat{B}_{2,2}(\lambda) \end{bmatrix}, \tag{4.15}$$

*where the polynomial matrix $\hat{B}_{1,1}(\lambda)$ is of dimension $p \times p$, $p \in \mathbb{N}$, $1 \leq p \leq \min\{m, n\}$, with its determinant $\det(\hat{B}_{1,1}(\lambda))$ being not identically equal to zero, then*

$$\deg(\det(\hat{B}_{1,1}(\lambda))) = \deg\left(\det\left(\begin{bmatrix} e_1^T \\ e_2^T \\ \vdots \\ e_p^T \end{bmatrix} B(\lambda)[e_{j_1}, e_{j_2}, \cdots, e_{j_p}]\right)\right) \leq pl. \tag{4.16}$$

**Proof**. Note that $A(\lambda)E$, which is left-equivalent to $B(\lambda)E$, is still a matrix polynomial of degree $l$. Thus, by Lemma 4.28, we immediately have $\deg\left(\det(\hat{B}_{1,1}(\lambda))\right) \leq pl$. Then the result follows by the fact that

$$\hat{B}_{1,1}(\lambda) = \begin{bmatrix} e_1^T \\ \vdots \\ e_p^T \end{bmatrix} B(\lambda)\left[e_{j_1}, e_{j_2}, \ldots, e_{j_p}\right]. \qquad \square$$

We will now derive upper bounds on the degree of the minors of the matrix polynomial $B(\lambda)$ in (4.2) which are required in the third stage of our investigation to construct a nonzero vector polynomial $x(\lambda)$ such that $B(\lambda)x(\lambda) = 0$, provided that $B(\lambda)$ is column-singular.

**Proposition 4.30** *Under the same assumption and the same notation as in Theorem 4.25, let $k = \min\{m, n\}$. If there exists $p \in \mathbb{N}$, $1 \le p \le k$, such that every diagonal element $b_{j,j}(\lambda)$ in (4.2), $1 \le j \le p$, is not identically equal to zero, then we have*

$$\deg\left(B(\lambda)\begin{pmatrix} 1 & \cdots & p \\ 1 & \cdots & p \end{pmatrix}\right) = \sum_{j=1}^{p} \deg(b_{j,j}(\lambda)) \le pl. \tag{4.17}$$

**Proof**. Note that $B(\lambda)$ in (4.2) is upper-triangular; then the result follows directly from Lemma 4.28. $\square$

**Proposition 4.31** *Under the same assumption and the same notation as in Theorem 4.25, let $k = \min\{m, n\}$. For any $p \in \mathbb{N}$, $1 \le p \le k$, and $j_i \in \mathbb{N}$, $1 \le j_i \le n$, $i = 1, \ldots, p$, which are pairwise distinct from each other, if the determinant* $\det\left(\begin{bmatrix} e_1^T \\ e_2^T \\ \vdots \\ e_p^T \end{bmatrix} B(\lambda)[e_{j_1}, e_{j_2}, \cdots, e_{j_p}]\right)$ *is not identically equal to zero, then*

$$\deg\left(\det\left(\begin{bmatrix} e_1^T \\ e_2^T \\ \vdots \\ e_p^T \end{bmatrix} B(\lambda)[e_{j_1}, e_{j_2}, \cdots, e_{j_p}]\right)\right) \le pl. \tag{4.18}$$

**Proof**. Let $\{j_1, j_2, \ldots, j_p, j_{p+1}, \ldots, j_n\} = \{1, \ldots, n\}$, and let $E = [e_{j_1}, e_{j_2}, \ldots, e_{j_n}]$ be a permutation matrix. Since $B(\lambda)$ in (4.2) is upper-triangular, we have

$$B(\lambda)E = \begin{bmatrix} \hat{B}_{1,1}(\lambda) & \hat{B}_{1,2}(\lambda) \\ 0 & \hat{B}_{2,2}(\lambda) \end{bmatrix}, \tag{4.19}$$

where $\hat{B}_{1,1}(\lambda) = \begin{bmatrix} e_1^T \\ e_2^T \\ \vdots \\ e_p^T \end{bmatrix} B(\lambda)[e_{j_1}, e_{j_2}, \ldots, e_{j_p}]$. If $\det(\hat{B}_{1,1}(\lambda))$ is not identically equal to zero, then from Lemma 4.29 the result follows. $\square$

Now we have paved the way for the final stage. The main result of this section is as follows.

**Theorem 4.32** *Let $A(\lambda) = \sum_{i=0}^{l} \lambda^i A_i$ be an $m \times n$ matrix polynomial of degree $l$, where $m, n \in \mathbb{N}$, $l \in \mathbb{N}_0$, $A_i \in \mathbb{C}^{m\times n}$, $i = 0, \ldots, l$. Let $k = \min(m, n)$, and let $r$ be the rank of $A(\lambda)$. If $A(\lambda)$ is column-singular, i.e., $r < n$, then there exists a nonzero vector polynomial $x(\lambda) = \sum_{i=0}^{d} \lambda^i x_i$, where $x_i \in \mathbb{C}^n$, $i = 1, \ldots, d$, $x_d \neq 0$, and $0 \leq d \leq rl \leq (k-1)l$, such that*

$$\forall \lambda \in \mathbb{C} : \ (\lambda^l A_l + \lambda^{l-1} A_{l-1} + \cdots + A_0)x(\lambda) = 0. \tag{4.20}$$

*Conversely, if there exists an $n$-dimensional nonzero vector polynomial $x(\lambda)$ which is a solution of (4.20), then $A(\lambda)$ is column-singular.*

**Proof**.

1. From Theorem 4.25 it follows that $A(\lambda)$ is left-equivalent to an upper-triangular matrix polynomial $B(\lambda)$ in the form (4.2). Hence, by Remark 4.27, there is a unimodular matrix polynomial $P(\lambda)$ such that $A(\lambda) = P(\lambda)B(\lambda)$. If $A(\lambda)$ is column-singular with rank $r < n$, by Propositions 4.23 and 4.24, $B(\lambda)$ is also column-singular with rank $r$.

The next step is to prove that there exists an $n$-dimensional nonzero vector polynomial $x(\lambda) = \sum_{i=0}^{d} \lambda^i x_i$, where $x_d \neq 0$, $0 \leq d \leq rl$, such that for all $\lambda \in \mathbb{C}$, $B(\lambda)x(\lambda) = 0$.

i) If $m \geq n$, then from the upper-triangular form and column-singularity of $B(\lambda)$, it follows that there exists at least one diagonal element $b_{k,k}(\lambda)$ of $B(\lambda)$ such that $b_{k,k}(\lambda)$ is identically equal to zero. Let $k_0 \in \mathbb{N}$, $1 \leq k_0 \leq (r + 1)$, be the smallest such index.

   (a) If $k_0 = 1$, then $b_{1,1}(\lambda) \equiv 0$. Let $x(\lambda) = x_0 = e_1$. Then we have $B(\lambda)x(\lambda) = 0$.

   (b) If $k_0 \geq 2$, we will, in virtue of Cramer's rule, construct $x(\lambda) = [x_1(\lambda), \ldots, x_n(\lambda)]^T$ in such a way that

$$\begin{cases} x_j(\lambda) = -\det\left( \begin{bmatrix} e_1^T \\ e_2^T \\ \vdots \\ e_{k_0-1}^T \end{bmatrix} B(\lambda)[e_1, \ldots, e_{j-1}, e_{k_0}, e_{j+1}, \ldots, e_{k_0-1}] \right), \\ \qquad\qquad j = 1, \ldots, k_0 - 1; \\ x_{k_0}(\lambda) = b_{1,1}(\lambda) \cdots b_{k_0-1,k_0-1}(\lambda); \\ x_j(\lambda) = 0, \qquad j = k_0 + 1, \ldots, n. \end{cases} \tag{4.21}$$

Since $b_{i,i}(\lambda) \not\equiv 0$, $i = 1, \ldots, k_0 - 1$, we have $x_{k_0}(\lambda) \not\equiv 0$, and therefore $x(\lambda)$ is not identically equal to zero. And, from Cramer's rule it follows that

$$
\begin{aligned}
&B(\lambda)x(\lambda) \\
&= \begin{bmatrix}
b_{1,1}(\lambda) & \cdots & b_{1,k_0-1}(\lambda) & b_{1,k_0}(\lambda) & \cdots & b_{1,n}(\lambda) \\
 & \ddots & \vdots & \vdots & \cdots & \vdots \\
 & & b_{k_0-1,k_0-1}(\lambda) & b_{k_0-1,k_0}(\lambda) & \cdots & b_{k_0-1,n}(\lambda) \\
 & & & 0 & \cdots & b_{k_0,n}(\lambda) \\
 & & & & \ddots & \vdots \\
 & & & & & b_{n,n}(\lambda) \\
0 & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & \cdots & \cdots & \cdots & \cdots & 0
\end{bmatrix}
\begin{bmatrix}
x_1(\lambda) \\ \vdots \\ x_{k_0-1}(\lambda) \\ x_{k_0}(\lambda) \\ 0 \\ \vdots \\ 0
\end{bmatrix} = 0.
\end{aligned}
$$

Now we estimate the degree of $x(\lambda)$. For any $j$, $1 \le j \le k_0 - 1$, if $x_j(\lambda)$ in (4.21) is not identically equal to 0, then by Proposition 4.31 we have

$$
\deg(x_j(\lambda)) = \deg\left(\det\left(\begin{bmatrix} e_1^T \\ e_2^T \\ \vdots \\ e_{k_0-1}^T \end{bmatrix} B(\lambda)[e_1, \ldots, e_{j-1}, e_{k_0}, e_{j+1}, \ldots, e_{k_0-1}]\right)\right) \le (k_0-1)l. \tag{4.22}
$$

Regarding the degree of $x_{k_0}(\lambda)$ in (4.21), by Proposition 4.30 we have

$$
\deg(x_{k_0}(\lambda)) = \sum_{i=1}^{k_0-1} \deg(b_{i,i}(\lambda)) \le (k_0 - 1)l. \tag{4.23}
$$

If we rewrite the polynomial vector $x(\lambda)$ in the form of a vector polynomial $\sum_{i=0}^{d} \lambda^i x_i$, where $x_d \neq 0$, then, from (4.22) and (4.23) it follows that

$$
\deg(x(\lambda)) = d \le (k_0 - 1)l \le rl \le (n-1)l. \tag{4.24}
$$

ii) If $m < n$, then we have the following cases:

(a) If there exists a diagonal element $b_{k,k}(\lambda)$ of $B(\lambda)$ such that $b_{k,k}(\lambda)$ is identically equal to zero, then the proof is analogous to the proof in the case of $m \ge n$.

(b) If for every $k \in \mathbb{N}$, $1 \le k \le m$, $b_{k,k}(\lambda)$ is not identically equal to zero, then let $k_0 = m+1$. We can construct, by (4.21), an $n$-dimensional nonzero vector polynomial $x(\lambda)$, such that $B(\lambda)x(\lambda) = 0$. Analogously, we also have

$$
\deg(x(\lambda)) = d \le (k_0 - 1)l = rl = ml. \tag{4.25}
$$

Finally, the proof of necessity ends with the fact that

$$A(\lambda)x(\lambda) = P(\lambda)B(\lambda)x(\lambda) = 0.$$

2. We prove the second part of the theorem by contradiction. Suppose that $A(\lambda)$ is column-regular, namely, $r = n$. Then, by Definition 4.6, there exists an $n$-order minor $A(\lambda)\begin{pmatrix} i_1 & i_2 & \cdots & i_n \\ 1 & 2 & \cdots & n \end{pmatrix}$, which is not identically equal to zero. Therefore, since the minor is a polynomial in $\lambda$, there are infinitely many values for $\lambda$, such that the minor at those values is not zero. However, note that there are at most finitely many values for $\lambda$, such that $x(\lambda)$ at those values is zero; therefore, by (4.20), there are at most finitely many values for $\lambda$, such that $A(\lambda)$ at those values has a full column rank. In other words, there are at most finitely many values for $\lambda$, such that any of the minors of order $n$ of $A(\lambda)$ at those values is not zero, which is a contradiction. □

**Remark 4.33** With respect to row-singularity, by Proposition 4.15 and Theorem 4.32, we have:

*$A(\lambda)$ is row-singular, i.e., $r < m$, if and only if there exists a nonzero vector polynomial $y(\lambda) = \sum_{i=0}^{d} \lambda^i y_i$, where $y_i \in \mathbb{C}^m$, $i = 1, \ldots, d$, $y_d \neq 0$, and $0 \le d \le rl \le (k-1)l$, such that*

$$\forall \lambda \in \mathbb{C}: \; y^H(\lambda)(\lambda^l A_l + \lambda^{l-1} A_{l-1} + \cdots + A_0) = 0. \tag{4.26}$$

◊

### 4.2.3 Sufficient and Necessary Conditions for Singularity and Regularity II

From Theorem 4.32 we deduce the following corollary which presents a geometric description of the column-singularity of matrix polynomials.

**Corollary 4.34** *Under the same assumptions and the same notation as in Theorem 4.32, $A(\lambda)$ is column-singular if and only if there exists a tuple $(x_0, x_1, \dots, x_d)$ of vectors, $x_i \in \mathbb{C}^n$, $i = 1, \dots, d$, $x_0 \neq 0$, $x_d \neq 0$, and $0 \leq d \leq rl \leq (k-1)l$, where $k = \min(m, n)$ and $r$ is the rank of $A(\lambda)$, such that*

$$\left\{\begin{array}{lllllllll}
A_0x_0 & & & & & & & = & 0, \\
A_1x_0 & + & A_0x_1 & & & & & = & 0, \\
 & & \cdots\cdots & & & & & & \\
A_{l-1}x_0 & + & A_{l-2}x_1 & + & \cdots & + & A_0x_{l-1} & = & 0, \\
A_lx_0 & + & A_{l-1}x_1 & + & \cdots & + & A_0x_l & = & 0, \\
A_lx_1 & + & A_{l-1}x_2 & + & \cdots & + & A_0x_{l+1} & = & 0, \\
 & & \cdots\cdots & & \cdots\cdots & & & & \\
A_lx_{d-l} & + & A_{l-1}x_{d-l+1} & + & \cdots & + & A_0x_d & = & 0, \\
A_lx_{d-l+1} & + & A_{l-1}x_{d-l+2} & + & \cdots & + & A_1x_d & = & 0, \\
 & & & & \cdots\cdots & & & & \\
 & & & & A_lx_{d-1} & + & A_{l-1}x_d & = & 0, \\
 & & & & & & A_lx_d & = & 0,
\end{array}\right. \tag{4.27}$$

*where $x_i = 0$, $i \in \mathbb{Z}$, for any $i < 0$ or $i > d$.*

**Proof**.

"$\Rightarrow$" From Theorem 4.32, it follows that there exists an $n$-dimensional nonzero vector polynomial

$$x(\lambda) = x_d\lambda^d + x_{d-1}\lambda^{d-1} + \cdots + x_0, \tag{4.28}$$

where $x_d \neq 0$, $0 \leq d \leq (k-1)l$, such that

$$\forall\lambda \in \mathbb{C}: \ (\lambda^l A_l + \lambda^{l-1}A_{l-1} + \cdots + A_0)x(\lambda) = 0. \tag{4.29}$$

Let $i \in \mathbb{N}_0$, $0 \leq i \leq d$, be such an index that $x_i \neq 0$ while $x_{i-1} = x_{i-2} = \cdots = x_0 = 0$. We define an $n$-dimensional nonzero vector polynomial

$$x(\lambda) := \frac{x(\lambda)}{\lambda^i} = x_d\lambda^d + x_{d-1}\lambda^{d-1} + \cdots + x_0, \tag{4.30}$$

where $d := d - i$, $x_j := x_{i+j}$, $j = 0, 1, \dots, d - i$. Thus, (4.29) still holds. If we substitute (4.30) in (4.29) and equate to zero the coefficients of every power of $\lambda$, then we immediately obtain the system of equations (4.27).

"$\Leftarrow$" From the given tuple of vectors, we construct a nonzero vector $x(\lambda)$ as in the form of (4.28). Then (4.29) follows from the given system of equations (4.27). Hence, by Theorem 4.32, the matrix polynomial $A(\lambda)$ is column-singular. $\square$

**Remark 4.35** Parallelly, in the case of row-singularity, we have:

*$A(\lambda)$ is row-singular if and only if there exists a tuple $(y_0, y_1, \ldots, y_{\hat{d}})$ of vectors, where $y_i \in \mathbb{C}^m$, $i = 1, \ldots, \hat{d}$, $y_0 \neq 0$, $y_{\hat{d}} \neq 0$, and $0 \leq \hat{d} \leq rl \leq (k-1)l$, such that*

$$\left\{\begin{array}{llllllllll}
y_0^H A_0 & & & & & & & = & 0, \\
y_0^H A_1 & + & y_1^H A_0 & & & & & = & 0, \\
 & & \cdots\cdots & & & & & & \\
y_0^H A_{l-1} & + & y_1^H A_{l-2} & + & \cdots & + & y_{l-1}^H A_0 & = & 0, \\
y_0^H A_l & + & y_1^H A_{l-1} & + & \cdots & + & y_l^H A_0 & = & 0, \\
y_1^H A_l & + & y_2^H A_{l-1} & + & \cdots & + & y_{l+1}^H A_0 & = & 0, \\
 & & \cdots\cdots & & \cdots\cdots & & & & \\
y_{\hat{d}-l}^H A_l & + & y_{\hat{d}-l+1}^H A_{l-1} & + & \cdots & + & y_{\hat{d}}^H A_0 & = & 0, \\
y_{\hat{d}-l+1}^H A_l & + & y_{\hat{d}-l+2}^H A_{l-1} & + & \cdots & + & y_{\hat{d}}^H A_1 & = & 0, \\
 & & & & \cdots\cdots & & & & \\
 & & & & y_{\hat{d}-1}^H A_l & + & y_{\hat{d}}^H A_{l-1} & = & 0, \\
 & & & & & & y_{\hat{d}}^H A_l & = & 0,
\end{array}\right. \tag{4.31}$$

*where $y_i = 0$, $i \in \mathbb{Z}$, for any $i < 0$ or $i > \hat{d}$.* ♢

**Remark 4.36** In the next subsection we shall, in virtue of Corollary 4.34, investigate more deeply the geometrical characteristic of column-singular matrix pencils and $2 \times 2$ quadratic matrix polynomials. ♢

By the rank information of the leading or the last coefficient matrix of a matrix polynomial, sometimes, as the following corollary shows, we can directly judge whether the matrix polynomial is column-regular (or row-regular).

**Corollary 4.37** *Under the same assumptions and the same notation as in Theorem 4.32, $A(\lambda)$ is column-regular if its leading coefficient matrix $A_l$, or its last coefficient matrix $A_0$, has full column rank $n$.*

**Proof**. If the leading coefficient matrix $A_l$ of the matrix polynomial $A(\lambda)$ has full column rank $n$, then there does not exist a nonzero vector $x$ of dimension $n$, such that $A_l x = 0$. Therefore, there does not exist such a tuple $(x_0, x_1, \ldots, x_d)$ of $n$-dimensional vectors, where $x_0 \neq 0$, $x_d \neq 0$, and $0 \leq d \leq rl \leq (k-1)l$, such that (4.27) can be established. Then, from Corollary 4.34 it follows that $A(\lambda)$ is column-regular.

In the case that the last coefficient matrix $A_0$ has full column rank, the proof is analogous to that in the above case of leading coefficient matrix. □

**Remark 4.38** Parallelly, we have:

*$A(\lambda)$ is row-regular if its leading coefficient matrix $A_l$, or its last coefficient matrix $A_0$, has full row rank $m$.* ♢

From Theorem 4.32 and Corollary 4.34 we obtain the following corollary which presents an algebraic description of the column-regularity of matrix polynomials.

**Corollary 4.39** *Under the same assumptions and the same notation as in Theorem 4.32, if $A(\lambda)$ is column-regular, then for all $s \in \mathbb{N}$, the $(s+l)$-by-$s$ block matrix with $m$-by-$n$ blocks*

$$W_s(A_l, A_{l-1}, \ldots, A_0) := \begin{bmatrix} A_0 & & & & & & \\ A_1 & A_0 & & & & & \\ \vdots & \vdots & \ddots & & & & \\ A_{l-1} & A_{l-2} & \cdots & A_0 & & & \\ A_l & A_{l-1} & \cdots & A_1 & A_0 & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & A_l & A_{l-1} & \cdots & A_1 & A_0 \\ & & & A_l & \cdots & A_2 & A_1 \\ & & & & \ddots & \vdots & \vdots \\ & & & & & A_l & A_{l-1} \\ & & & & & & A_l \end{bmatrix} \tag{4.32}$$

*has full column rank $ns$; conversely, if for all $s \in \mathbb{N}$, $s \leq (rl+1) \leq ((k-1)l+1)$, where $k = \min(m,n)$ and $r$ is the rank of $A(\lambda)$, $W_s(A_l, A_{l-1}, \ldots, A_0)$ has full column rank $ns$, then $A(\lambda)$ is column-regular.*

**Proof**.

1. If $A(\lambda)$ is column-regular, we prove the first part of the corollary by contradiction. Suppose that there exists $s \in \mathbb{N}$ such that the column rank of the $(s+l)$-by-$s$ block matrix $W_k(A_l, A_{l-1}, \ldots, A_0)$ in (4.32) is deficient, then there exists $x \in \mathbb{C}^{ns}$, $x \neq 0$, such that

$$W_s(A_l, A_{l-1}, \ldots, A_0)x = 0. \tag{4.33}$$

Let $x = [x_0^T, x_1^T, \ldots, x_{s-1}^T]^T$, $x_i \in \mathbb{C}^n$, $i = 0, 1, \ldots, (s-1)$, and let $x(\lambda) = \sum_{i=0}^{s-1} \lambda^i x_i$. Since $x \neq 0$, $x(\lambda) \neq 0$. From (4.33) it follows that (4.27) holds, where $d := s$. By (4.27) we obtain $A(\lambda)x(\lambda) = 0$. Then from Theorem 4.32 we get the conclusion that the matrix polynomial $A(\lambda)$ must be column-singular, which is a contradiction.

2. If for all $s \in \mathbb{N}$, $s \leq (rl+1) \leq ((k-1)l+1)$, $W_s(A_l, A_{l-1}, \ldots, A_0)$ has full column rank $ns$, then we prove the second part of the corollary also by contradiction. Suppose that the matrix polynomial $A(\lambda)$ is column-singular, then from Corollary 4.34 it follows that there is a tuple $(x_0, x_1, \ldots, x_d)$ of vectors, where $x_i \in \mathbb{C}^n$, $i = 1, \ldots, d$, $x_0 \neq 0$, $x_d \neq 0$, and $0 \leq d \leq rl \leq (k-1)l$, such that the system of equations (4.27) is satisfied. Let $s := d+1$, and let $x =: [x_0^T, x_1^T, \ldots, x_{s-1}^T]^T$. Then we have

$1 \leq s \leq (rl+1) \leq ((k-1)l+1)$, and $x \neq 0$. Thus, we can rewrite the system of equations (4.27) in the following form:

$$W_s(A_l, A_{l-1}, \ldots, A_0)x = 0,$$

which is a contradiction. $\square$

**Remark 4.40** Parallelly, we have:

*Under the same assumptions and the same notation as in Theorem 4.32, if $A(\lambda)$ is row-regular, then for all $s \in \mathbb{N}$, the $s$-by-$(s+l)$ block matrix with $m$-by-$n$ blocks*

$$\widehat{W}_s(A_l, A_{l-1}, \ldots, A_0) = \begin{bmatrix} A_0 & A_1 & \cdots & A_l & & & & & \\ & A_0 & \cdots & A_{l-1} & A_l & & & & \\ & & \ddots & \ddots & \ddots & \ddots & & & \\ & & & A_0 & \cdots & A_{l-1} & A_l & & \\ & & & & \ddots & \ddots & \ddots & \ddots & \\ & & & & & A_0 & \cdots & A_{l-1} & A_l \end{bmatrix} \tag{4.34}$$

*has full row rank $ms$; conversely, if for all $s \in \mathbb{N}$, $s \leq (rl+1) \leq ((k-1)l+1)$, $\widehat{W}_s(A_l, A_{l-1}, \ldots, A_0)$ has full row rank $ms$, then $A(\lambda)$ is row-regular.* $\diamondsuit$

**Remark 4.41** By the second part of Corollary 4.39 we know that, given any matrix polynomial $A(\lambda)$, we can judge within a finite number of steps of computation whether or not $A(\lambda)$ is column-regular. In Section 4.3, we shall present another way by which we can numerically detect whether a square matrix polynomial is regular or singular. $\diamondsuit$

The following example shows that the upper bound $(rl+1)$ or $((k-1)l+1)$ on $s$ in Corollary 4.39 is attainable.

**Example 4.42** We consider the quadratic matrix polynomial $A(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0$, where

$$A_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_1 = 0, \quad A_0 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Here, $m = n = 3$, $k = 3$, $r = 2$, $l = 2$, and $rl = ((k-1)l+1) = 5$. Since $\det(A(\lambda)) \equiv 0$, $A(\lambda)$ is singular. If we investigate the column ranks of $W_s(A_2, A_1, A_0)$, $s = 1, \ldots, 5$, we find that

$$\underset{(s=1)}{\operatorname{rank}\left(\begin{bmatrix} A_0 \\ A_1 \\ A_2 \end{bmatrix}\right) = 3}, \quad \underset{(s=2)}{\operatorname{rank}\left(\begin{bmatrix} A_0 & \\ A_1 & A_0 \\ A_2 & A_1 \\ & A_2 \end{bmatrix}\right) = 6}, \quad \underset{(s=3)}{\operatorname{rank}\left(\begin{bmatrix} A_0 & & \\ A_1 & A_0 & \\ A_2 & A_1 & A_0 \\ & A_2 & A_1 \\ & & A_2 \end{bmatrix}\right) = 9},$$

$$\operatorname{rank}\left(\begin{bmatrix} A_0 & & & \\ A_1 & A_0 & & \\ A_2 & A_1 & A_0 & \\ & A_2 & A_1 & A_0 \\ & & A_2 & A_1 \\ & & & A_2 \end{bmatrix}\right) = 12, \quad \operatorname{rank}\left(\begin{bmatrix} A_0 & & & & \\ A_1 & A_0 & & & \\ A_2 & A_1 & A_0 & & \\ & A_2 & A_1 & A_0 & \\ & & A_2 & A_1 & A_0 \\ & & & A_2 & A_1 \\ & & & & A_2 \end{bmatrix}\right) = 14 < 15,$$
$$(s = 4) \qquad\qquad\qquad\qquad (s = 5)$$

which shows that, not until $s = rl + 1 = ((k-1)l + 1) = 5$, the rank of the block matrix $W_s(A_2, A_1, A_0)$ is deficient. $\Diamond$

**Remark 4.43** If a matrix polynomial $A(\lambda)$ is column-singular, then by Theorem 4.32 we know that there exists a nonzero vector polynomial $x(\lambda)$ such that

$$A(\lambda)x(\lambda) \equiv 0. \tag{4.35}$$

Let $d_{\min}$ be the least possible degree of such solutions $x(\lambda)$ of (4.35). By Corollaries 4.34 and 4.39, we know that $d_{\min} = s_{\min} - 1$, where $s_{\min}$ is the least integer $s$ such that the column rank of $W_s(A_l, A_{l-1}, \ldots, A_0)$ in (4.32) is deficient. From Example 4.42 we know that $rl$ or $(k-1)l$ is an attainable upper bound on $d_{\min}$. $\Diamond$

### 4.2.4 Singular Matrix Pencils and Quadratic Matrix Polynomials

In Subsection 4.2.3 we have seen that Corollary 4.34 describes the relations between the column spaces of the coefficient matrices of a given column-singular matrix polynomial. From Corollary 4.34 we now proceed to investigate more deeply and obtain a geometrical characterization of singular matrix pencils, as the following theorem presents.

**Theorem 4.44** *A matrix pencil $A_0 + \lambda A_1$, $A_0, A_1 \in \mathbb{C}^{m\times n}$ is column-singular if and only if there exists a subspace $\mathcal{X}$ of $\mathbb{C}^n$ such that*

$$\dim(A_0\mathcal{X} \dotplus A_1\mathcal{X}) < \dim(\mathcal{X}), \tag{4.36}$$

*where $A\mathcal{X} := \{Ax : x \in \mathcal{X}\}$ if $A \in \mathbb{C}^{m\times n}$ and $\mathcal{X} \subseteq \mathbb{C}^n$, and $\mathcal{X} \dotplus \mathcal{Y} := \{x + y : x \in \mathcal{X}, y \in \mathcal{Y}\}$ if $\mathcal{X}, \mathcal{Y} \subseteq \mathbb{C}^m$.*

**Proof**.

"$\Rightarrow$" Since the matrix pencil $A_0+\lambda A_1$ is column-singular, by Corollary 4.34 we have

$$\left\{\begin{array}{lcccl} A_0x_0 & & & = & 0, \\ A_1x_0 & + & A_0x_1 & = & 0, \\ A_1x_1 & + & A_0x_2 & = & 0, \\ & \cdots & \cdots & & \\ A_1x_{d-1} & + & A_0x_d & = & 0, \\ & & A_1x_d & = & 0, \end{array}\right. \tag{4.37}$$

where $x_i \in \mathbb{C}^n$, $i = 0, 1, \ldots, d$, $d \in \mathbb{N}_0$, $0 \le d \le \min\{m, n\} - 1$, $x_0 \neq 0$, and $x_d \neq 0$. Let $\mathcal{X} = \text{span}\{x_0, x_1, \ldots, x_d\}$, $\mathcal{X}_0 = \text{span}\{x_1, \ldots, x_d\}$, and let $\mathcal{X}_d = \text{span}\{x_0, x_1, \ldots, x_{d-1}\}$. From (4.37) it follows that

$$A_0\mathcal{X} = A_0\mathcal{X}_0 = A_1\mathcal{X}_d = A_1\mathcal{X}. \tag{4.38}$$

Since $x_0 \neq 0$ and $A_0x_0 = 0$, we have $A_0\mathcal{X} \subsetneqq \mathcal{X}$. Thus,

$$\dim(A_0\mathcal{X}) < \dim(\mathcal{X}). \tag{4.39}$$

Then, (4.38) and (4.39) imply (4.36).

"$\Leftarrow$" We conduct the proof of sufficiency in the following two cases.

i) In the case $m < n$, clearly, by Definition 4.11, $A_0 + \lambda A_1$ is column-singular.

ii) In the case $m \ge n$, let $Z := [Z_1, Z_2]$ be a nonsingular matrix with $Z_1 \in \mathbb{C}^{n\times l}$, $\mathcal{R}(Z_1) = \mathcal{X}$, and let $l = \dim(\mathcal{X})$. From the given inequality (4.36), we see that both $A_0\mathcal{X}$ and $A_1\mathcal{X}$ lie in a $k$-dimensional subspace $\mathcal{Y}$ of $\mathbb{C}^m$, where $k < l$. If we let $Q := [Q_1, Q_2]$ be a unitary matrix with $Q_1 \in \mathbb{C}^{n\times k}$, $\mathcal{R}(Q_1) = \mathcal{Y}$, then we have

$$\begin{aligned} Q^H A_i Z &= \begin{array}{c} \\ k \\ m-k \end{array} \begin{array}{c} \begin{array}{cc} \quad l \quad & \quad n-l \quad \end{array} \\ \begin{bmatrix} Q_1^H A_i Z_1 & Q_1^H A_i Z_2 \\ 0 & Q_2^H A_i Z_2 \end{bmatrix} \end{array} \\ &= \begin{array}{c} \\ k \\ l-k \\ m-l \end{array} \begin{array}{c} \begin{array}{ccc} k & l-k & n-l \end{array} \\ \begin{bmatrix} * & * & * \\ 0 & 0 & * \\ 0 & 0 & * \end{bmatrix} \end{array}, \quad i = 0, 1. \end{aligned} \tag{4.40}$$

From (4.40) it follows that any minor of order $n$ of the matrix polynomial $Q^H(A_0+\lambda A_1)Z$ is 0. Hence, the rank of $Q^H(A_0 + \lambda A_1)Z$ is strictly less than $n$. By Definition 4.11, $Q^H(A_0+\lambda A_1)Z$ is column-singular. Since $Q$ and $Z$ are nonsingular matrices, it follows that $A_0 + \lambda A_1$ is column-singular. $\square$

**Remark 4.45** By Theorem 4.44, we can see from the point of view of generalized eigenvalue problems that, for a *regular* matrix pencil $A_0 + \lambda A_1$, $A_i \in \mathbb{C}^{n\times n}$, $i = 0, 1$, a subspace $\mathcal{X}$ is a DEFLATING SUBSPACE if and only if

$$\dim(A_0\mathcal{X} \dotplus A_1\mathcal{X}) = \dim(\mathcal{X}).$$

For more details about deflating subspaces for a matrix pencil, such as properties of them and perturbation analysis of them, see, for example, [54], [55] (Chapter VI). $\Diamond$

**Remark 4.46** Later, in Subsection 4.4.2, we will see that in the special case of square matrix pencils, the geometrical characterization (4.36) coincides with the algebraic characterization of the nearest distance to singularity for matrix pencils. The algebraic characterization will be given in Theorem 4.82. $\Diamond$

For a singular matrix polynomial other than a matrix pencil, the relations between the column spaces of its coefficient matrices become, as Corollary 4.34 indicates, very complicated. The reason for the complexity is the nonlinearity which comes from the higher degree (greater than 1) of the matrix polynomial. Nevertheless, based on Corollary 4.34, we can explore the simplest case of singular quadratic matrix polynomials, and obtain the following theorem.

**Theorem 4.47** *Let matrices $A_2, A_1, A_0 \in \mathbb{C}^{2\times 2}$. Then, the quadratic matrix polynomial $\lambda^2 A_2 + \lambda A_1 + A_0$ is singular if and only if one of the following three cases happens:*

1. $\mathcal{N}(A_2) \cap \mathcal{N}(A_1) \cap \mathcal{N}(A_0) \neq \emptyset$.
2. $\mathcal{N}(A_2^H) \cap \mathcal{N}(A_1^H) \cap \mathcal{N}(A_0^H) \neq \emptyset$.
3. *There exist nonsingular matrices $X, Y \in \mathbb{C}^{2\times 2}$, such that:*

$$Y^{-1}A_2X = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \; Y^{-1}A_1X = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \; Y^{-1}A_0X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

**Proof.**
"$\Rightarrow$" Since $\lambda^2 A_2 + \lambda A_1 + A_0$ is singular, by Corollary 4.34, there exists a tuple of vectors $(x_0, x_1, \ldots, x_d)$, where $x_0 \neq 0$, $x_d \neq 0$, $0 \leq d \leq 2$, such that

$$\left\{\begin{array}{lclclcl} A_0x_0 & & & & & = & 0, \\ A_1x_0 & + & A_0x_1 & & & = & 0, \\ A_2x_0 & + & A_1x_1 & + & A_0x_2 & = & 0, \\ A_2x_1 & + & A_1x_2 & + & A_0x_3 & = & 0, \\ & & \cdots\cdots & & & & \\ A_2x_{d-2} & + & A_1x_{d-1} & + & A_0x_d & = & 0, \\ & & A_2x_{d-1} & + & A_1x_d & = & 0, \\ & & & & A_2x_d & = & 0. \end{array}\right. \tag{4.41}$$

According to different values of $d$ in (4.41), we conduct the proof in the following three cases.

i) If $d = 0$, then by (4.41) there exists $x_0 \neq 0$ such that $A_2 x_0 = A_1 x_0 = A_0 x_0 = 0$, in other words, $\mathcal{N}(A_2) \cap \mathcal{N}(A_1) \cap \mathcal{N}(A_0) \neq \emptyset$, which is Case 1.

ii) If $d = 1$, then by (4.41) there exist $x_0 \neq 0$ and $x_1 \neq 0$, such that

$$\begin{cases} A_0 x_0 & & & = & 0, \\ A_1 x_0 & + & A_0 x_1 & = & 0, \\ A_2 x_0 & + & A_1 x_1 & = & 0, \\ & & A_2 x_1 & = & 0. \end{cases} \tag{4.42}$$

(a) If $x_0$ and $x_1$ are linearly dependent, then $A_2 x_0 = A_1 x_0 = A_0 x_0 = 0$, which is Case 1.

(b) If $x_0$ and $x_1$ are linearly independent, and $A_1 x_0$ and $A_1 x_1$ are linearly dependent, then there exists $y \in \mathbb{C}^2$, $y \neq 0$, such that $y^H A_1 x_0 = y^H A_1 x_1 = 0$. From (4.42) it follows that $y^H A_2 x_0 = y^H A_2 x_1 = 0$ and $y^H A_0 x_0 = y^H A_0 x_1 = 0$. Since $x_0$ and $x_1$ are linearly independent, we obtain $y^H A_2 = y^H A_1 = y^H A_0 = 0$, namely, $\mathcal{N}(A_2^H) \cap \mathcal{N}(A_1^H) \cap \mathcal{N}(A_0^H) \neq \emptyset$, which is Case 2.

(c) If $x_0$ and $x_1$ are linearly independent, and $A_1 x_0$ and $A_1 x_1$ are linearly independent, then we have, by setting $X := [-x_0, x_1]$ and $Y := [-A_1 x_0, A_1 x_1]$,

$$A_2 X = Y \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \; A_1 X = Y \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \; A_0 X = Y \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

which is Case 3, because $X$ and $Y$ are nonsingular.

iii) If $d = 2$, then by (4.41) there exist $x_0 \neq 0$, $x_1$, and $x_2 \neq 0$, such that,

$$\begin{cases} A_0 x_0 & & & & & = & 0, \\ A_1 x_0 & + & A_0 x_1 & & & = & 0, \\ A_2 x_0 & + & A_1 x_1 & + & A_0 x_2 & = & 0, \\ & & A_2 x_1 & + & A_1 x_2 & = & 0, \\ & & & & A_2 x_2 & = & 0. \end{cases} \tag{4.43}$$

(a) If $\dim(\operatorname{span}\{x_0, x_1, x_2\}) = 1$, then from (4.43) it is derived that $A_2 x_0 = A_1 x_0 = A_0 x_0 = 0$, which is Case 1.

(b) If $\dim(\operatorname{span}\{x_0, x_1, x_2\}) = 2$, and $\dim(\operatorname{span}\{A_1 x_0, A_1 x_1, A_1 x_2\}) \leq 1$, then there exists $y \in \mathbb{C}^2$, $y \neq 0$, such that $y^H A_1 x_0 = y^H A_1 x_1 = y^H A_1 x_2 = 0$. From (4.43) it follows that $y^H A_2 x_0 = y^H A_2 x_1 = y^H A_2 x_2 = 0$ and $y^H A_0 x_0 = y^H A_0 x_1 = y^H A_0 x_2 = 0$. Since $\dim(\operatorname{span}\{x_0, x_1, x_2\}) = 2$, we obtain $y^H A_2 = y^H A_1 = y^H A_0 = 0$, namely, $\mathcal{N}(A_2^H) \cap \mathcal{N}(A_1^H) \cap \mathcal{N}(A_0^H) \neq \emptyset$, which is Case 2.

(c) If $\dim(\mathrm{span}\{x_0, x_1, x_2\}) = 2$, and $\dim(\mathrm{span}\{A_1x_0, A_1x_1, A_1x_2\}) = 2$, then it is clear that $A_1$ is nonsingular. We shall prove via contradiction that $x_0$ and $x_2$ are linearly independent. Suppose that $x_2 = ax_0$, $a \neq 0$. Then $A_1x_2 = aA_1x_0$, and $A_0x_0 = A_0x_2 = A_2x_0 = A_2x_2 = 0$; therefore, by (4.43), $A_1x_1 = 0$. Thus, it follows that $\dim(\mathrm{span}\{A_1x_0, A_1x_1, A_1x_2\}) < 2$, which is a contradiction.

Let $x_1 := ax_0 + bx_2$. We prove via contradiction that $a \neq 0$ and $b \neq 0$. Supposing $a = 0$, we get $x_1 = bx_2$. From (4.43) it follows that $A_2x_1 = bA_2x_2$ and therefore $A_1x_2 = 0$, which is a contradiction to the fact that $A_1$ in nonsingular and $x_2 \neq 0$. Similarly, supposing $b = 0$, we also get a contradiction.

Thus, substituting $ax_0 + bx_2$ for $x_1$ in (4.43), we obtain

$$A_0x_0 = 0,\ A_0x_2 = -\frac{1}{b}A_1x_0, \tag{4.44}$$

$$\left(a - \frac{1}{b}\right) A_1x_0 + \left(b - \frac{1}{a}\right) A_1x_2 = 0, \tag{4.45}$$

$$A_2(-ax_0) = A_1x_2,\ A_2x_2 = 0. \tag{4.46}$$

Since $A_1x_0$ and $A_1x_2$ are linearly independent, from (4.45) it is derived that $ab = 1$. Therefore, setting $X := [-ax_0, x_2]$ and $Y := [-aA_1x_0, A_1x_2]$, we have, by (4.44) -(4.46),

$$A_2X = Y\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},\ A_1X = Y\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ A_0X = Y\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

which is Case 3, because $X$ and $Y$ are nonsingular.

"$\Leftarrow$" The statement is straightforward. □

## 4.3 Detecting Regularity/Singularity of Square Matrix Polynomials by Rank Information

### 4.3.1 Introduction

As we have pointed out in Section 4.1, one of the major motivations for studying the regularity and singularity of matrix polynomials comes from the analysis of problems in which *square* matrix polynomials are involved. Numerically speaking, it is reasonable for a mathematical software developed for solving polynomial eigenvalue problems to be able to determine the regularity or singularity of any square matrix polynomial involved, and the *distance to singularity* if it is regular. However, this is not the case with classical commercial software packages. For instance, let us use MATLAB (Version 6.0.0.88 Release 12) to analyze the following example.

**Example 4.48** We consider the matrix pencil:

$$A - \lambda E = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}.$$

Evidently, for any $\lambda \in \mathbb{C}$, we have $\det(A - \lambda E) \equiv 0$; therefore, the matrix pencil is singular, and the corresponding generalized eigenvalue problem $Ax = \lambda Ex$ has infinitely many eigenpairs $(\lambda, x)$, where

$$\lambda \in \mathbb{C}, \ x = \begin{bmatrix} \lambda - 1 \\ -2\lambda + 1 \end{bmatrix}.$$

Nevertheless, if we use the following Matlab function:

$$[V, D] = \mathrm{eig}(A, E),$$

which produces a diagonal matrix $D$ of generalized eigenvalues and a full matrix $V$ whose columns are the corresponding eigenvectors so that $AV = EVD$ (cf. the "MATLAB Function Reference" [39]), then we get the following output $[V, D]$ which, unfortunately, does not indicate any information about the singularity of $A - \lambda E$:

$$V = \begin{bmatrix} -1.0000 & -0.3385 \\ 0 & 1.0000 \end{bmatrix}, \quad D = \begin{bmatrix} 0.5000 & 0 \\ 0 & 2.0482 \end{bmatrix}.$$

$\Diamond$

It is the purpose of this section to derive methods to detect the regularity and singularity of *square* matrix polynomials. We will in the following subsection present a natural approach to detect the regularity or singularity of a given square matrix polynomial, provided that the rank information of its coefficient matrices is known. And later, in Section 4.4 we will define and discuss the *nearness to singularity problem* for square matrix polynomials.

### 4.3.2 Testing for Regularity and Singularity

In Section 4.2 we have presented general sufficient and necessary conditions for regularity and singularity of *rectangular* matrix polynomials. Among those conditions is Corollary 4.37, by which we can directly determine the *column-regularity* of a matrix polynomial as long as either its leading or its last coefficient matrix has full column rank. The question of interest here is how we can test whether or not a given matrix polynomial is column-regular if the column ranks of both its leading and last coefficient matrices are deficient and the rank information of its coefficient matrices is assumed to be known beforehand. The following proposition gives a sufficient condition for a matrix polynomial to be *column-singular*, provided that the sum of the ranks of its coefficient matrices is sufficiently small.

**Proposition 4.49** *Let $A(\lambda) = \sum_{i=0}^{l} \lambda^i A_i$ be an $m \times n$ matrix polynomial of degree $l$, where $m, n \in \mathbb{N}$, $l \in \mathbb{N}_0$, $A_i \in \mathbb{C}^{m\times n}$, $i = 0, \ldots, l$. Then $A(\lambda)$ is column-singular if $\sum_{i=0}^{l} \operatorname{rank}(A_i) \leq (n-1)$.*

**Proof**. Noting that

$$\operatorname{rank}\left(\begin{bmatrix} A_0 \\ A_1 \\ \vdots \\ A_l \end{bmatrix}\right) \leq \sum_{i=0}^{l} \operatorname{rank}(A_i) \leq (n-1),$$

we know that the column rank of the matrix $W_1(A_l, A_{l-1}, \ldots, A_0) := \begin{bmatrix} A_0 \\ A_1 \\ \vdots \\ A_l \end{bmatrix}$ is deficient. Therefore, by Corollary 4.39, $A(\lambda)$ is column-singular. □

**Remark 4.50** Similarly, in the case of row-singularity, we have:

*Under the same assumption and the same notation as in Proposition 4.49, $A(\lambda)$ is row-singular if $\sum_{i=0}^{l} \operatorname{rank}(A_i) \leq (m-1)$.* ◊

Corollary 4.37 and Proposition 4.49 show that, under specified rank conditions for the coefficient matrices of a matrix polynomial in Corollary 4.37 and Proposition 4.49, we can immediately judge whether it is column-singular. A natural question then arises, namely, under other rank conditions than those in Corollary 4.37 and Proposition 4.49, how do we make use of the beforehand rank information to test for column-regularity or column-singularity?

For *square* matrix polynomials we can answer the above question. First, let us recall Proposition 4.13 in Subsection 4.2.1, which can be regarded, from the point of view of polynomial eigenvalue problems, as an equivalent definition for the singularity of square matrix polynomials. By Proposition 4.13, the determinant of any *regular* square matrix polynomial $A(\lambda)$, or regarded from the point of view of eigenvalue problems as the characteristic polynomial of $A(\lambda)$, is a nonzero polynomial in $\lambda$, and therefore has only a *finite number* of roots. Our objective in this subsection is to analyze the determinant $\det(A(\lambda))$ with the use of the rank information of the coefficient matrices of $A(\lambda)$, and to derive an attainable upper bound on the *finite number* of nonzero roots of $\det(A(\lambda))$. Then, the upper bound will lead to a stopping criterion of an algorithm which we will present at the end of this subsection to detect the regularity or singularity of any given square matrix polynomial.

In order to conduct the analysis, we need the following lemma which shows, roughly speaking, that the determinant of the sum of matrices can be represented as a certain sum of products of the minors of the matrices.

**Lemma 4.51** *Let matrices* $A_k = \left[a_{i,j}^{(k)}\right]_{i,j=1}^{n} \in \mathbb{C}^{n\times n}$, $k = 0, 1, \ldots, l$, $l \in \mathbb{N}$. *Then,*

$$\det\left(\sum_{k=0}^{l} A_k\right) = \sum_{\substack{0\le p\le n\\ i,j}} \left\{(-1)^s A_l\begin{pmatrix} i_1 & \cdots & i_{p_l}\\ j_1 & \cdots & j_{p_l}\end{pmatrix}\cdot A_{l-1}\begin{pmatrix} i_{p_l+1} & \cdots & i_{p_{l-1}}\\ j_{p_l+1} & \cdots & j_{p_{l-1}}\end{pmatrix}\cdot\right.$$
$$\left.\cdots\cdot A_1\begin{pmatrix} i_{p_2+1} & \cdots & i_{p_1}\\ j_{p_2+1} & \cdots & j_{p_1}\end{pmatrix}\cdot A_0\begin{pmatrix} i_{p_1+1} & \cdots & i_{n}\\ j_{p_1+1} & \cdots & j_{n}\end{pmatrix}\right\}, \quad (4.47)$$

*where* $p = p_l + \cdots + p_1, p_q \in \mathbb{N}_0,\ q = 1, \ldots, l,\ \ 0 \le p_l \le p_{l-1} \le \cdots \le p_1 \le n;$

$i := (i_1, \ldots, i_{p_l}, i_{p_l+1}, \ldots, i_{p_{l-1}}, \ldots, i_{p_1+1}, \ldots, i_n)$ *is a permutation of* $\{1, 2, \ldots, n\}$,

$i_1 < i_2 < \cdots < i_{p_l}, i_{p_l+1} < i_{p_l+2} < \cdots < i_{p_{l-1}},\ \ldots,\ i_{p_1+1} < i_{p_1+2} < \cdots < i_n;$

$j := (j_1, \ldots, j_{p_l}, j_{p_l+1}, \ldots, j_{p_{l-1}}, \ldots, j_{p_1+1}, \ldots, j_n)$ *is a permutation of* $\{1, 2, \ldots, n\}$,

$j_1 < j_2 < \cdots < j_{p_l}, j_{p_l+1} < j_{p_l+2} < \cdots < j_{p_{l-1}},\ \ldots,\ j_{p_1+1} < j_{p_1+2} < \cdots < j_n;$

$$s = \left(\sum_{q=1}^{p_l} i_q + \sum_{q=p_l+1}^{p_{l-1}} i_q + \cdots + \sum_{q=p_2+1}^{p_1} i_q\right) + \left(\sum_{q=1}^{p_l} j_q + \sum_{q=p_l+1}^{p_{l-1}} j_q + \cdots + \sum_{q=p_2+1}^{p_1} j_q\right).$$

**Proof**. We conduct the proof by induction on $l$.

1. $l = 1$: Let $B = [b_{i,j}]_{i,j=1}^{n} := A_0 + A_1 = \left[a_{i,j}^{(0)} + a_{i,j}^{(1)}\right]_{i,j=1}^{n}$. Then, by the definition of the determinant of a matrix (cf., for example, [36] p. 26.) we have

$$\begin{aligned}\det(B) &= \sum_k (-1)^{t(k)} b_{1,k_1} b_{2,k_2} \cdots b_{n,k_n}\\ &= \sum_k (-1)^{t(k)} \left(a_{1,k_1}^{(0)} + a_{1,k_1}^{(1)}\right)\left(a_{2,k_2}^{(0)} + a_{2,k_2}^{(1)}\right)\cdots\left(a_{n,k_n}^{(0)} + a_{n,k_n}^{(1)}\right),\end{aligned} \quad (4.48)$$

where $k := (k_1, \ldots, k_n)$ is any of the permutations of $1, 2, \ldots, n$, and $t(k)$ is the number of inversions in the permutation $k$. Throughout the proof we denote by $t(i)$, where $i$ is a permutation of $1, 2, \ldots, n$, the number of inversions in the permutation $i$, and $i$ is called *odd* or *even* according to whether the number $t(i)$ is odd or even.

Decomposing each term in the sum in (4.48) and rearranging the terms obtained after the decomposition, we have

$$\det(B) = \sum_{\substack{0 \le p \le n \\ k, i}} (-1)^{t(k)} a^{(1)}_{i_1,k_{i_1}} a^{(1)}_{i_2,k_{i_2}} \cdots a^{(1)}_{i_p,k_{i_p}} a^{(0)}_{i_{p+1},k_{i_{p+1}}} a^{(0)}_{i_{p+2},k_{i_{p+2}}} \cdots a^{(0)}_{i_n,k_{i_n}}, \tag{4.49}$$

where $p \in \mathbb{N}_0$ and the permutation $i := (i_1, \ldots, i_n)$, $i_1 < i_2 < \cdots < i_p$, $i_{p+1} < i_{p+2} < \cdots < i_n$. Let $\tilde{j} := (\tilde{j}_1, \ldots, \tilde{j}_n)$, where $\tilde{j}_m = k_{i_m}$, $m = 1, \ldots, n$. Note that, to get the permutation $\tilde{j}$ from the permutation $k$, we can interchange two neighbouring elements in $k$ by altogether $\left(\sum_{q=1}^{p} i_q - \frac{p(p+1)}{2}\right)$ times. Since every time of interchanging two neighbouring elements in a permutation changes it from odd to even or vice versa (cf. [36] p. 26.), we have

$$(-1)^{t(k)} = (-1)^{\left(t(\tilde{j}) + \sum_{q=1}^{p} i_q - \frac{p(p+1)}{2}\right)}. \tag{4.50}$$

Thus, by (4.49) and (4.50) we have

$$\det(B) = \sum_{\substack{0 \le p \le n \\ \tilde{k}, i}} (-1)^{\left(t(\tilde{j}) + \sum\limits_{q=1}^{p} i_q - \frac{p(p+1)}{2}\right)} a^{(1)}_{i_1,\tilde{j}_1} a^{(1)}_{i_2,\tilde{j}_2} \cdots a^{(1)}_{i_p,\tilde{j}_p} a^{(0)}_{i_{p+1},\tilde{j}_{p+1}} \cdots a^{(0)}_{i_n,\tilde{j}_n}. \tag{4.51}$$

Let

$$\begin{aligned}
&j_{q_m} := \tilde{j}_m, \; m = 1, 2, \ldots, n; \\
&j_{1:p} := (j_{q_1}, j_{q_2}, \ldots, j_{q_p}) = (\tilde{j}_1, \tilde{j}_2 \ldots, \tilde{j}_p) \text{ be a permutation of } \{j_1, j_2, \ldots, j_p\}, \\
&\qquad\qquad\qquad\qquad \text{where } j_1 < j_2 < \cdots < j_p; \\
&j_{p+1:n} := (j_{q_{p+1}}, j_{q_{p+2}}, \ldots, j_{q_n}) = (\tilde{j}_{p+1}, \tilde{j}_{p+2} \ldots, \tilde{j}_n) \text{ be a permutation of} \\
&\qquad\qquad\qquad\qquad \{j_{p+1}, j_{p+2}, \ldots, j_n\}, \text{ where } j_{p+1} < j_{p+2} < \cdots < j_n; \\
&j := (j_1, j_2, \ldots, j_n).
\end{aligned}$$

Note that, to get the permutation $j$ from the permutation $(1, 2, \ldots, n)$, we can interchange two neighbouring elements in $(1, 2, \ldots, n)$ by altogether $\left(\sum_{q=1}^{p} j_q - \frac{p(p+1)}{2}\right)$ times; therefore, we have

$$(-1)^{t(j)} = (-1)^{\left(\sum_{q=1}^{p} j_q - \frac{p(p+1)}{2}\right)} \tag{4.52}$$

Note further that the number of inversions in the permutation $\tilde{j}$ is larger than that in the permutation $j$ by $(t(j_{1:p}) + t(j_{p+1:n}))$, namely,

$$t(\tilde{j}) = t(j_{1:p}) + t(j_{p+1:n}) + t(j). \tag{4.53}$$

Thus, by (4.51)-(4.53), we have

$$\det(B) = \sum_{\substack{0 \le p \le n \\ j,i}} (-1)^t a^{(1)}_{i_1,j_{q_1}} a^{(1)}_{i_2,j_{q_2}} \cdots a^{(1)}_{i_p,j_{q_p}} a^{(0)}_{i_{p+1},j_{q_{p+1}}} a^{(0)}_{i_{p+2},j_{q_{p+2}}} \cdots a^{(0)}_{i_n,j_{q_n}}, \tag{4.54}$$

where $t = t(j_{1:p}) + t(j_{p+1:n}) + \sum_{q=1}^{p} i_q - \frac{p(p+1)}{2} + \sum_{q=1}^{p} j_q - \frac{p(p+1)}{2}$. Note that $p(p+1)$ is even, and

$$\sum_q (-1)^{t(j_{1:p})} a^{(1)}_{i_1,j_{q_1}} a^{(1)}_{i_2,j_{q_2}} \cdots a^{(1)}_{i_p,j_{q_p}} = A_1 \begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix}, \tag{4.55}$$

$$\sum_q (-1)^{t(j_{p+1:n})} a^{(0)}_{i_{p+1},j_{q_{p+1}}} a^{(0)}_{i_{p+2},j_{q_{p+2}}} \cdots a^{(0)}_{i_n,j_{q_n}} = A_0 \begin{pmatrix} i_{p+1} & i_{p+2} & \cdots & i_n \\ j_{p+1} & j_{p+2} & \cdots & j_n \end{pmatrix}. \tag{4.56}$$

Rearranging the terms of the sum in (4.54) and substituting (4.55) and (4.56) in (4.54), we obtain that

$$\det(A_0 + A_1) = \sum_{\substack{0 \le p \le n \\ j,i}} (-1)^s A_1 \begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix} A_0 \begin{pmatrix} i_{p+1} & i_{p+2} & \cdots & i_n \\ j_{p+1} & j_{p+2} & \cdots & j_n \end{pmatrix}, \tag{4.57}$$

where $s = \sum_{q=1}^{p} i_q + \sum_{q=1}^{p} j_q$. Thus, we have finished the first step of the induction proof.

2. $l - 1 \Rightarrow l$: Suppose that we have already proved (4.47) for all $m$, where $1 \le m \le (l-1)$, $l \ge 2$; we shall prove that (4.47) also holds for $l$. Let $B = \sum_{k=0}^{l-1} A_k$. Since (4.57) has been proven for the determinant of the sum of two matrices, we have

$$\det(B + A_l) = \sum_{\substack{0 \le p_l \le n \\ \hat{j},\hat{i}}} (-1)^{\hat{s}} A_l \begin{pmatrix} i_1 & i_2 & \cdots & i_{p_l} \\ j_1 & j_2 & \cdots & j_{p_l} \end{pmatrix} B \begin{pmatrix} i_{p_l+1} & i_{p_l+2} & \cdots & i_n \\ j_{p_l+1} & j_{p_l+2} & \cdots & j_n \end{pmatrix}, \tag{4.58}$$

where $\hat{s} = \sum_{q=1}^{p_l} i_q + \sum_{q=1}^{p_l} j_q$, the permutation $\hat{i} = (i_1, \ldots, i_n)$, $i_1 < i_2 < \cdots < i_{p_l}$, $i_{p_l+1} < i_{p_l+2} < \cdots < i_n$, and the permutation $\hat{j} = (j_1, \ldots, j_n)$, $j_1 < j_2 < \cdots < j_{p_l}$, $j_{p_l+1} < j_{p_l+2} < \cdots < j_n$. Since we have supposed that (4.47) holds for $l-1$, for each term of $B\begin{pmatrix} i_{p_l+1} & i_{p_l+2} & \cdots & i_n \\ j_{p_l+1} & j_{p_l+2} & \cdots & j_n \end{pmatrix}$ in the sum in (4.58) we have

$$B\begin{pmatrix} i_{p_l+1} & i_{p_l+2} & \cdots & i_n \\ j_{p_l+1} & j_{p_l+2} & \cdots & j_n \end{pmatrix} = \sum_{\substack{0 \le \bar{p} \le n - p_l \\ \bar{i}, \bar{j}}} \left\{(-1)^{\bar{s}} A_{l-1}\begin{pmatrix} i_{p_l+1} & \cdots & i_{p_{l-1}} \\ j_{p_l+1} & \cdots & j_{p_{l-1}} \end{pmatrix} \cdot \right.$$

$$\left. A_{l-2}\begin{pmatrix} i_{p_{l-1}+1} & \cdots & i_{p_{l-2}} \\ j_{p_{l-1}+1} & \cdots & j_{p_{l-2}} \end{pmatrix} \cdots A_0 \begin{pmatrix} i_{p_1+1} & \cdots & i_n \\ j_{p_1+1} & \cdots & j_n \end{pmatrix}\right\}, \quad (4.59)$$

where $\bar{p} = p_{l-1} + \cdots + p_1, p_q \in \mathbb{N}_0,\ q = 1, \ldots, l-1,\ \ 0 \le p_{l-1} \le \cdots \le p_1 \le n;$

$\bar{i} := (i_{p_l+1}, \ldots, i_{p_{l-1}}, \ldots, i_{p_1+1}, \ldots, i_n)$ is a permutation of $\{i_{p_l+1}, i_{p_l+2}, \ldots, i_n\}$,

$$i_{p_l+1} < i_{p_l+2} < \cdots < i_{p_{l-1}},\ \ldots,\ i_{p_1+1} < i_{p_1+2} < \cdots < i_n;$$

$\bar{j} := (j_{p_l+1}, \ldots, j_{p_{l-1}}, \ldots, j_{p_1+1}, \ldots, j_n)$ is a permutation of $\{j_{p_l+1}, j_{p_l+2}, \ldots, j_n\}$;

$$j_{p_l+1} < j_{p_l+2} < \cdots < j_{p_{l-1}},\ \ldots,\ j_{p_1+1} < j_{p_1+2} < \cdots < j_n;$$

$$\bar{s} = \left(\sum_{q=p_l+1}^{p_{l-1}} i_q + \cdots + \sum_{q=p_2+1}^{p_1} i_q\right) + \left(\sum_{q=p_l+1}^{p_{l-1}} j_q + \cdots + \sum_{q=p_2+1}^{p_1} j_q\right).$$

Let

$$p := p_l + \bar{p}; \quad (4.60)$$

$$s := \hat{s} + \bar{s}; \quad (4.61)$$

$$i := (i_1, \ldots, i_{p_l}, i_{p_l+1}, \ldots, i_{p_{l-1}}, \ldots, i_{p_1+1}, \ldots, i_n); \quad (4.62)$$

$$j := (j_1, \ldots, j_{p_l}, j_{p_l+1}, \ldots, j_{p_{l-1}}, \ldots, j_{p_1+1}, \ldots, j_n). \quad (4.63)$$

Then, substituting (4.59) in (4.58) and rewriting the sum in (4.58) in terms of $p, s, i, j$ defined by (4.60)-(4.63), we finally obtain that (4.47) also holds for $l$. $\square$

Now in virtue of Lemma 4.51 we proceed to analyze the characteristic polynomials of square matrix polynomials and to derive attainable upper bounds on the number of nonzero roots of the characteristic polynomials if they are not identically equal to zero. The upper bounds are presented in terms of the ranks of the coefficient matrices of matrix polynomials.

**Theorem 4.52** *Let $A(\lambda) = \lambda A_1 + A_0$, $A_1, A_0 \in \mathbb{C}^{n \times n}$, be a regular matrix pencil, and let $\lambda_1, \ldots, \lambda_m$, $m \in \mathbb{N}_0$, be all the* nonzero *roots of* $\det(A(\lambda))$ (*in which repetitions may be included, and $m = 0$ means that all its roots are zero*). *Then,*

$$0 \le m \le (\operatorname{rank}(A_1) + \operatorname{rank}(A_0) - n). \quad (4.64)$$

**Proof**. From (4.47) in Lemma 4.51 it follows that the characteristic polynomial of $\lambda A_1 + A_0$ can be computed via the following minor expansions:

$$\begin{aligned} &\det(\lambda A_1 + A_0) \\ = &\sum_{\substack{0 \le p \le n \\ j,i}} (-1)^s (\lambda A_1) \begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix} A_0 \begin{pmatrix} i_{p+1} & i_{p+2} & \cdots & i_n \\ j_{p+1} & j_{p+2} & \cdots & j_n \end{pmatrix}, \\ = &\sum_{\substack{0 \le p \le n \\ j,i}} (-1)^s A_1 \begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix} A_0 \begin{pmatrix} i_{p+1} & i_{p+2} & \cdots & i_n \\ j_{p+1} & j_{p+2} & \cdots & j_n \end{pmatrix} \lambda^p, \end{aligned} \tag{4.65}$$

where the permutations $i = (i_1, \ldots, i_n)$, $j = (j_1, \ldots, j_n)$, and $s = \sum_{q=1}^{p} i_q + \sum_{q=1}^{p} j_q$. Now based on (4.65) we calculate the highest and lowest possible orders of $\lambda$ in $\det(A(\lambda))$. Since the rank of a matrix is equal to the order of its largest nonzero minor, we have

$$A_1 \begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix} = 0, \qquad \text{if } p > \operatorname{rank}(A_1); \tag{4.66}$$

$$A_0 \begin{pmatrix} i_{p+1} & i_{p+2} & \cdots & i_n \\ j_{p+1} & j_{p+2} & \cdots & j_n \end{pmatrix} = 0, \quad \text{if } p < (n - \operatorname{rank}(A_0)). \tag{4.67}$$

Hence, from (4.65) and (4.66) it follows that the highest possible order of $\lambda$ in $\det(A(\lambda))$ is $\operatorname{rank}(A_1)$, and from (4.65) and (4.67) it follows that the lowest possible order of $\lambda$ in $\det(A(\lambda))$ is $(n - \operatorname{rank}(A_0))$. Since $A(\lambda)$ is regular, its characteristic polynomial $\det(A(\lambda))$ is not identically equal to zero. Therefore, $\det(A(\lambda))$ can have at most $\operatorname{rank}(A_1) - (n - \operatorname{rank}(A_0)) = \operatorname{rank}(A_1) + \operatorname{rank}(A_0) - n$ nonzero roots, or in other words, $m \le (\operatorname{rank}(A_1) + \operatorname{rank}(A_0) - n)$. □

**Remark 4.53** From the above proof of Theorem 4.52, it immediately follows that for a regular matrix pencil $\lambda A_1 + A_0$, where $A_1, A_0 \in \mathbb{C}^{n \times n}$, the corresponding generalized eigenvalue problem $A_0 x = -\lambda A_1 x$ has eigenvalue $\infty$ with algebraic multiplicity *greater than or equal to* $(n - \operatorname{rank}(A_1))$, and eigenvalue 0 with algebraic multiplicity *greater than or equal to* $(n - \operatorname{rank}(A_0))$. ◊

As a straightforward consequence of Theorem 4.52, we have the following corollary.

**Corollary 4.54** *Let $A_1, A_0 \in \mathbb{C}^{n \times n}$. Then the matrix pencil $\lambda A_1 + A_0$ is singular if there exist at least $\tilde{m}$* distinct *nonzero numbers $\tilde{\lambda}_i$, $i = 1, \ldots, \tilde{m}$, such that all the matrices $\tilde{\lambda}_i A_1 + A_0$, $i = 1, \ldots, \tilde{m}$, are singular, where $\tilde{m} := \operatorname{rank}(A_1) + \operatorname{rank}(A_0) - n + 1$.*

**Proof**. The result follows directly from Theorem 4.52 via contradiction. □

**Remark 4.55** If $\tilde{m}$ in Corollary 4.54 is not greater than 0, i.e.,

$$\mathrm{rank}(A_1) + \mathrm{rank}(A_0) - n + 1 \leq 0,$$

then, by Proposition 4.49 or Theorem 4.52, the matrix pencil $\lambda A_1 + A_0$ is singular. $\Diamond$

With respect to a given regular and square quadratic matrix polynomial, we analyze via Lemma 4.51 its characteristic polynomial in a similar way in which we analyze regular and square matrix pencils, and the following result is obtained.

**Theorem 4.56** *Let $A(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0$, $A_i \in \mathbb{C}^{n\times n}$, $i = 0, 1, 2$, be a regular quadratic matrix polynomial, and let $\lambda_1, \ldots, \lambda_m$, $m \in \mathbb{N}_0$, be all the* nonzero *roots of* $\det(A(\lambda))$ (*in which repetitions may be included, and $m = 0$ means all its roots are zero*). *Then,*

$$\begin{aligned} 0 \leq m \leq (\mathrm{rank}(A_2) + \mathrm{rank}(A_0) \quad &+ \quad \min\{\mathrm{rank}(A_2) + \mathrm{rank}(A_1) - n, 0\} \\ &+ \quad \min\{\mathrm{rank}(A_1) + \mathrm{rank}(A_0) - n, 0\}) \,. \end{aligned} \tag{4.68}$$

**Proof**. By (4.47) in Lemma 4.51, the characteristic polynomial of $\lambda^2 A_2 + \lambda A_1 + A_0$ can be computed via the following minor expansions:

$$\begin{aligned} &\det\left(A(\lambda)\right) \\ = &\sum_{\substack{0 \leq p+q \leq n \\ i,j}} \left\{ (-1)^s (\lambda^2 A_2) \begin{pmatrix} i_1 & \cdots & i_p \\ j_1 & \cdots & j_p \end{pmatrix} \right. \\ &\qquad \left. \cdot (\lambda A_1) \begin{pmatrix} i_{p+1} & \cdots & i_{p+q} \\ j_{p+1} & \cdots & j_{p+q} \end{pmatrix} \cdot A_0 \begin{pmatrix} i_{p+q+1} & \cdots & i_n \\ j_{p+q+1} & \cdots & j_n \end{pmatrix} \right\} \\ = &\sum_{\substack{0 \leq p+q \leq n \\ i,j}} \left\{ (-1)^s A_2 \begin{pmatrix} i_1 & \cdots & i_p \\ j_1 & \cdots & j_p \end{pmatrix} \right. \\ &\qquad \left. \cdot A_1 \begin{pmatrix} i_{p+1} & \cdots & i_{p+q} \\ j_{p+1} & \cdots & j_{p+q} \end{pmatrix} \cdot A_0 \begin{pmatrix} i_{p+q+1} & \cdots & i_n \\ j_{p+q+1} & \cdots & j_n \end{pmatrix} \cdot \lambda^{2p+q} \right\}, \end{aligned} \tag{4.69}$$

where $p, q \in \mathbb{N}_0$, the permutations $i = (i_1, \ldots, i_p, i_{p+1}, \ldots, i_{p+q}, i_{p+q+1}, \ldots, i_n)$, $j = (j_1, \ldots, j_p, j_{p+1}, \ldots, j_{p+q}, j_{p+q+1}, \ldots, j_n)$, and $s = \sum_{t=1}^{p+q} i_t + \sum_{t=1}^{p+q} j_t$. Let $O_h$, $O_l$ denote the highest and lowest possible orders of $\lambda$ in $\det(A(\lambda))$, respectively. Now according to the minor expansions (4.69) we calculate $O_h$ in the following two cases.

(a) In the case $(\mathrm{rank}(A_2) + \mathrm{rank}(A_1)) \leq n$, let $p_1 = \mathrm{rank}(A_2)$ and let $q_1 = \mathrm{rank}(A_1)$. It follows that $p_1 + q_1 \leq n$; therefore, it is possible that there exist permutations $i$ and $j$ so that

$$C^{(i,j)} := A_2 \begin{pmatrix} i_1 & \cdots & i_{p_1} \\ j_1 & \cdots & j_{p_1} \end{pmatrix} \cdot A_1 \begin{pmatrix} i_{p_1+1} & \cdots & i_{p_1+q_1} \\ j_{p_1+1} & \cdots & j_{p_1+q_1} \end{pmatrix} \cdot A_0 \begin{pmatrix} i_{p+q+1} & \cdots & i_n \\ j_{p+q+1} & \cdots & j_n \end{pmatrix} \neq 0.$$

Moreover, it is also possible that the sum of the terms $\sum_{i,j} C^{(i,j)}\lambda^{2p_1+q_1} \neq 0$. Hence, by (4.69), it is possible that $\det(A(\lambda))$ has the term $\lambda^{2p_1+q_1}$ with nonzero coefficient $\sum_{i,j} C^{(i,j)}$. Since the rank of a matrix is equal to the order of its largest nonzero minor, and $p_1 = \text{rank}(A_2)$ and $q_1 = \text{rank}(A_1)$, it is obvious that the highest possible degree of $\det(A(\lambda))$ is $2p_1 + q_1$, namely,

$$O_h = 2p_1 + q_1 = 2\text{rank}(A_2) + \text{rank}(A_1). \tag{4.70}$$

(b) In the case $(\text{rank}(A_2)+\text{rank}(A_1)) > n$, let $p_2 = \text{rank}(A_2)$ and let $q_2 = n - \text{rank}(A_2)$. It follows that $p_2 + q_2 = n$; therefore, it is possible that there exist permutations $i$ and $j$ so that

$$C^{(i,j)} := A_2 \begin{pmatrix} i_1 & \cdots & i_{p_2} \\ j_1 & \cdots & j_{p_2} \end{pmatrix} \cdot A_1 \begin{pmatrix} i_{p_2+1} & \cdots & i_n \\ j_{p_2+1} & \cdots & j_n \end{pmatrix} \neq 0.$$

Moreover, it is also possible that the sum of the terms $\sum_{i,j} C^{(i,j)}\lambda^{2p_2+q_2} \neq 0$. Hence, by (4.69), it is possible that $\det(A(\lambda))$ has the term $\lambda^{2p_2+q_2}$ with nonzero coefficient $\sum_{i,j} C^{(i,j)}$. Since the rank of a matrix is equal to the order of its largest nonzero minor, and $p_2 = \text{rank}(A_2)$ and $q_2 = n - \text{rank}(A_2)$, it is obvious that the highest possible degree of $\det(A(\lambda))$ is $2p_2 + q_2$, namely,

$$O_h = 2p_2 + q_2 = 2\text{rank}(A_2) + n - \text{rank}(A_1) = \text{rank}(A_2) + n. \tag{4.71}$$

For conciseness of expression, we rewrite (4.70) and (4.71) into a single formula:

$$O_h = \text{rank}(A_2) + \min\{\text{rank}(A_2) + \text{rank}(A_1), n\}. \tag{4.72}$$

Analogously, according to (4.69) we calculate $O_l$ in the following two cases.

(a) In the case $(\text{rank}(A_1) + \text{rank}(A_0)) \leq n$, we have

$$O_l = 2(n - \text{rank}(A_1) - \text{rank}(A_0)) + \text{rank}(A_1) = 2n - \text{rank}(A_1) - 2\text{rank}(A_0). \tag{4.73}$$

(b) In the case $(\text{rank}(A_1) + \text{rank}(A_0)) > n$, we have

$$O_l = n - \text{rank}(A_0). \tag{4.74}$$

We also rewrite (4.73) and (4.74) into the following formula:

$$O_l = n - \text{rank}(A_0) + \max\{n - \text{rank}(A_1) - \text{rank}(A_0), 0\}. \tag{4.75}$$

Since $A(\lambda)$ is regular, its characteristic polynomial $\det(A(\lambda))$ is not identically equal to zero. Therefore, $\det(A(\lambda))$ can have at most $O_h - O_l$ nonzero roots, namely, in terms of $m$ given in the condition, we have $m \leq (O_h - O_l)$. Finally, the proof ends with the following calculation of $O_h - O_l$:

$$\begin{aligned}
O_h - O_l &= \operatorname{rank}(A_2) + \min\{\operatorname{rank}(A_2) + \operatorname{rank}(A_1), n\} \\
&\quad -(n - \operatorname{rank}(A_0)) - \max\{n - \operatorname{rank}(A_1) - \operatorname{rank}(A_0), 0\} \\
&= \operatorname{rank}(A_2) + \min\{\operatorname{rank}(A_2) + \operatorname{rank}(A_1) - n, 0\} \\
&\quad +\operatorname{rank}(A_0) - \max\{n - \operatorname{rank}(A_1) - \operatorname{rank}(A_0), 0\} \\
&= \operatorname{rank}(A_2) + \operatorname{rank}(A_0) + \min\{\operatorname{rank}(A_2) + \operatorname{rank}(A_1) - n, 0\} \\
&\qquad\qquad + \min\{\operatorname{rank}(A_1) + \operatorname{rank}(A_0) - n, 0\}.
\end{aligned}$$

□

**Remark 4.57** From the proof of Theorem 4.56, it immediately follows that, for a regular quadratic matrix polynomial $\lambda^2 A_2 + \lambda A_1 + A_0$, where $A_2, A_1, A_0 \in \mathbb{C}^{n\times n}$, the corresponding quadratic eigenvalue problem $(\lambda^2 A_2 + \lambda A_1 + A_0)x = 0$ has eigenvalue $\infty$ with algebraic multiplicity

$$m_\infty \geq (n - \operatorname{rank}(A_2) - \min\{\operatorname{rank}(A_2) + \operatorname{rank}(A_1) - n, 0\}) \tag{4.76}$$

and eigenvalue 0 with algebraic multiplicity

$$m_0 \geq (n - \operatorname{rank}(A_0) - \min\{\operatorname{rank}(A_1) + \operatorname{rank}(A_0) - n, 0\}). \tag{4.77}$$

◊

From Theorem 4.56 it is straightforward to obtain the following corollaries.

**Corollary 4.58** *Let $A_2, A_1, A_0 \in \mathbb{C}^{n\times n}$. Then the quadratic matrix polynomial $\lambda^2 A_2 + \lambda A_1 + A_0$ is singular if there exist at least $\tilde{m}$* distinct *nonzero numbers $\tilde{\lambda}_i$, $i = 1, \ldots, \tilde{m}$, such that all the matrices $\tilde{\lambda}_i^2 A_2 + \tilde{\lambda}_i A_1 + A_0$, $i = 1, \ldots, \tilde{m}$, are singular, where*

$$\begin{aligned}
\tilde{m} := \operatorname{rank}(A_2) + \operatorname{rank}(A_0) &+ \min\{\operatorname{rank}(A_2) + \operatorname{rank}(A_1) - n, 0\} \\
&+ \min\{\operatorname{rank}(A_1) + \operatorname{rank}(A_0) - n, 0\} + 1.
\end{aligned} \tag{4.78}$$

**Corollary 4.59** *Under the same assumption and the same notation as in Corollary 4.58, if $\tilde{m} \leq 0$, then the quadratic matrix polynomial $\lambda^2 A_2 + \lambda A_1 + A_0$ is singular.*

**Remark 4.60** It should be noted that the sufficient condition $\tilde{m} \le 0$ for the singularity in Corollary 4.59 is equivalent to the sufficient condition for the singularity in Proposition 4.49, since

$$\begin{aligned}
\operatorname{rank}(A_2)+\operatorname{rank}(A_0) &+ \min\{\operatorname{rank}(A_2)+\operatorname{rank}(A_1)-n,0\} \\
&+ \min\{\operatorname{rank}(A_1)+\operatorname{rank}(A_0)-n,0\}+1 \le 0 \\
\Longleftrightarrow \operatorname{rank}(A_2)+\operatorname{rank}(A_0) &+ \operatorname{rank}(A_2)+\operatorname{rank}(A_1)-n \\
&+ \operatorname{rank}(A_1)+\operatorname{rank}(A_0)-n+1 \le 0 \\
\Longleftrightarrow \operatorname{rank}(A_2)+\operatorname{rank}(A_1) &+ \operatorname{rank}(A_0)-n+\frac{1}{2} \le 0 \\
\Longleftrightarrow \operatorname{rank}(A_2)+\operatorname{rank}(A_1) &+ \operatorname{rank}(A_0)-n+1 \le 0. \qquad \diamond
\end{aligned}$$

To illustrate how to make use of Corollary 4.37, Proposition 4.49, and Corollary 4.58 to detect the regularity or singularity of square quadratic matrix polynomials, we give the following example.

**Example 4.61** Consider a quadratic matrix polynomial $A(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0$, where

$$A_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \ A_1 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \text{ and } A_0 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{4.79}$$

We carry out the test for regularity or singularity of $A(\lambda)$ in the following three steps.

1. Check the rank information of the leading and last coefficient matrices of $A(\lambda)$: since $\operatorname{rank}(A_2) = \operatorname{rank}(A_0) = 1 < 3$, by Corollary 4.37, $A(\lambda)$ may be singular.

2. Check the sum of the ranks of all coefficient matrices of $A(\lambda)$: since $\operatorname{rank}(A_2) + \operatorname{rank}(A_1) + \operatorname{rank}(A_0) = 3 > 2 = 3 - 1$, by Proposition 4.49, the matrix polynomial $A(\lambda)$ may be regular.

3. Compute $\tilde{m}$ defined by (4.78) in Corollary 4.58 and test for regularity or singularity: since $\tilde{m} = 1 + 1 - 1 - 1 + 1 = 1$, by Corollary 4.58, we need only to test, for any *one* nonzero number $\tilde{\lambda}_1$, whether $A(\tilde{\lambda}_1)$ is singular. If we randomly let $\tilde{\lambda}_1 = 1$, clearly, $A(1) = A_2 + A_1 + A_0$ is nonsingular; therefore, it is concluded that the quadratic matrix polynomial $A(\lambda)$ is regular.

Moreover, it is immediate from (4.79) that $\det(A(\lambda)) = -\lambda^3$, and therefore $\det(A(\lambda))$ has no nonzero root, which shows that the upper bound $\tilde{m} - 1 = 0$ obtained in (4.68) in Theorem 4.56 is *attainable*. $\diamond$

Similar to Theorem 4.52 and Theorem 4.56, we have the following general theorem which sets an attainable upper bound on the numbers of nonzero roots of the characteristic polynomials of square and regular matrix polynomials of degree $l$.

**Theorem 4.62** *Let $A(\lambda) = \sum_{i=0}^{l} \lambda^i A_i$, $A_i \in \mathbb{C}^{n\times n}$, $i = 0, 1, \ldots, l$, $l \in \mathbb{N}$, be a regular matrix polynomial of degree $l$, and let $\lambda_1, \ldots, \lambda_m$, $m \in \mathbb{N}_0$, be all the* nonzero *roots of* $\det(A(\lambda))$ (*in which repetitions may be included, and $m = 0$ means all its roots are zero*). *Then,*

$$\begin{aligned} 0 \le m \quad \le \quad & (\operatorname{rank}(A_l) + \operatorname{rank}(A_0) + (l-2)n \\ & + \sum_{j=1}^{l-1} \min\left\{ \sum_{k=j}^{l} \operatorname{rank}(A_k) - n, 0 \right\} + \sum_{j=1}^{l-1} \min\left\{ \sum_{k=j}^{l} \operatorname{rank}(A_{l-k}) - n, 0 \right\} \Bigg). \end{aligned} \tag{4.80}$$

**Proof**. Since the proof is similar to that of Theorem 4.56, we only present a brief sketch of the facts which lead to the result, as follows.

(a) By Lemma 4.51, the minor expansions of the characteristic polynomial of $A(\lambda)$ is:

$$\begin{aligned} \det(A(\lambda)) \quad = \quad & \sum_{\substack{0 \le p \le n \\ i,j}} \Bigg\{ (-1)^s \, (\lambda^l A_l) \begin{pmatrix} i_1 & \cdots & i_{p_l} \\ j_1 & \cdots & j_{p_l} \end{pmatrix} \cdot (\lambda^{l-1} A_{l-1}) \begin{pmatrix} i_{p_l+1} & \cdots & i_{p_{l-1}} \\ j_{p_l+1} & \cdots & j_{p_{l-1}} \end{pmatrix} \cdot \\ & \cdots \cdot (\lambda A_1) \begin{pmatrix} i_{p_2+1} & \cdots & i_{p_1} \\ j_{p_2+1} & \cdots & j_{p_1} \end{pmatrix} \cdot A_0 \begin{pmatrix} i_{p_1+1} & \cdots & i_n \\ j_{p_1+1} & \cdots & j_n \end{pmatrix} \Bigg\} \\ = \quad & \sum_{\substack{0 \le p \le n \\ i,j}} \Bigg\{ (-1)^s A_l \begin{pmatrix} i_1 & \cdots & i_{p_l} \\ j_1 & \cdots & j_{p_l} \end{pmatrix} \cdot A_{l-1} \begin{pmatrix} i_{p_l+1} & \cdots & i_{p_{l-1}} \\ j_{p_l+1} & \cdots & j_{p_{l-1}} \end{pmatrix} \cdot \\ & \cdots \cdot A_1 \begin{pmatrix} i_{p_2+1} & \cdots & i_{p_1} \\ j_{p_2+1} & \cdots & j_{p_1} \end{pmatrix} \cdot A_0 \begin{pmatrix} i_{p_1+1} & \cdots & i_n \\ j_{p_1+1} & \cdots & j_n \end{pmatrix} \lambda^d \Bigg\}, \end{aligned} \tag{4.81}$$

where $p = p_l + \cdots + p_1, p_q \in \mathbb{N}_0,\ q = 1, \ldots, l,\ \ 0 \le p_l \le p_{l-1} \le \cdots \le p_1 \le n$;
$i := (i_1, \ldots, i_{p_l}, i_{p_l+1}, \ldots, i_{p_{l-1}}, \ldots, i_{p_1+1}, \ldots, i_n)$ is a permutation of $\{1, 2, \ldots, n\}$,
$i_1 < i_2 < \cdots < i_{p_l}, i_{p_l+1} < i_{p_l+2} < \cdots < i_{p_{l-1}},\ \ldots,\ i_{p_1+1} < i_{p_1+2} < \cdots < i_n$;
$j := (j_1, \ldots, j_{p_l}, j_{p_l+1}, \ldots, j_{p_{l-1}}, \ldots, j_{p_1+1}, \ldots, j_n)$ is a permutation of $\{1, 2, \ldots, n\}$,
$j_1 < j_2 < \cdots < j_{p_l}, j_{p_l+1} < j_{p_l+2} < \cdots < j_{p_{l-1}},\ \ldots,\ j_{p_1+1} < j_{p_1+2} < \cdots < j_n$;

$$s = \left( \sum_{q=1}^{p_l} i_q + \sum_{q=p_l+1}^{p_{l-1}} i_q + \cdots + \sum_{q=p_2+1}^{p_1} i_q \right) + \left( \sum_{q=1}^{p_l} j_q + \sum_{q=p_l+1}^{p_{l-1}} j_q + \cdots + \sum_{q=p_2+1}^{p_1} j_q \right);$$

and $d = l p_l + (l-1) p_{l-1} + \cdots + p_1$.

(b) Based on (4.81), we calculate the highest possible order of $\lambda$ in $\det(A(\lambda))$:

$$\begin{aligned} O_h &= \operatorname{rank}(A_l) + \sum_{j=1}^{l-1} \min\left\{\sum_{k=j}^{l} \operatorname{rank}(A_k), n\right\} \\ &= \operatorname{rank}(A_l) + (l-1)n + \sum_{j=1}^{l-1} \min\left\{\sum_{k=j}^{l} \operatorname{rank}(A_k) - n, 0\right\}. \end{aligned} \tag{4.82}$$

(c) Analogously, we calculate the lowest possible order of $\lambda$ in $\det(A(\lambda))$:

$$\begin{aligned} O_l &= n - \operatorname{rank}(A_0) + \sum_{j=1}^{l-1} \max\left\{n - \sum_{k=j}^{l} \operatorname{rank}(A_{l-k}), 0\right\} \\ &= n - \operatorname{rank}(A_0) - \sum_{j=1}^{l-1} \min\left\{\sum_{k=j}^{l} \operatorname{rank}(A_{l-k}) - n, 0\right\}. \end{aligned} \tag{4.83}$$

(d) By (4.82) and (4.83), we derive an upper bound on the number $m$:

$$\begin{aligned} m \leq O_h - O_l &= \operatorname{rank}(A_l) + (l-1)n + \sum_{j=1}^{l-1} \min\left\{\sum_{k=j}^{l} \operatorname{rank}(A_k) - n, 0\right\} \\ &\quad - \left(n - \operatorname{rank}(A_0) - \sum_{j=1}^{l-1} \min\left\{\sum_{k=j}^{l} \operatorname{rank}(A_{l-k}) - n, 0\right\}\right) \\ &= (\operatorname{rank}(A_l) + \operatorname{rank}(A_0) + (l-2)n \\ &\quad + \sum_{j=1}^{l-1} \min\left\{\sum_{k=j}^{l} \operatorname{rank}(A_k) - n, 0\right\} + \sum_{j=1}^{l-1} \min\left\{\sum_{k=j}^{l} \operatorname{rank}(A_{l-k}) - n, 0\right\}\Bigg). \end{aligned}$$

□

**Remark 4.63** From (4.82) and (4.83) in the proof of Theorem 4.62, it follows that for a regular matrix polynomial $A(\lambda) = \sum_{i=0}^{l} \lambda^i A_i$ of degree $l$, where $A_i \in \mathbb{C}^{n\times n}$, $i = 0, 1, \ldots, l$, $l \in \mathbb{N}$, the corresponding polynomial eigenvalue problem $\left(\sum_{i=0}^{l} \lambda^i A_i\right) x = 0$ has eigenvalue $\infty$ with algebraic multiplicity

$$m_\infty \geq \left(n - \operatorname{rank}(A_l) - \sum_{j=1}^{l-1} \min\left\{\sum_{k=j}^{l} \operatorname{rank}(A_k) - n, 0\right\}\right), \tag{4.84}$$

and eigenvalue 0 with algebraic multiplicity

$$m_0 \geq \left(n - \operatorname{rank}(A_0) - \sum_{j=1}^{l-1} \min\left\{\sum_{k=j}^{l} \operatorname{rank}(A_{l-k}) - n, 0\right\}\right). \tag{4.85}$$

◊

Also, Theorem 4.62 directly implies the following corollaries.

**Corollary 4.64** *Let $A_i \in \mathbb{C}^{n\times n}$, where $i = 0, 1, \ldots, l$, $l \in \mathbb{N}$. Then the matrix polynomial $\sum_{i=0}^{l} \lambda^i A_i$ of degree $l$ is singular if there exist at least $\tilde{m}$* distinct nonzero *numbers $\tilde{\lambda}_i$, $i = 1, \ldots, \tilde{m}$, such that all the matrices $\sum_{i=0}^{l} \tilde{\lambda}^i A_i$, $i = 1, \ldots, \tilde{m}$, are singular, where*

$$\begin{aligned} \tilde{m} : \;&=\; (\operatorname{rank}(A_l) + \operatorname{rank}(A_0) + (l-2)n \\ &\quad + \sum_{j=1}^{l-1} \min\left\{\sum_{k=j}^{l} \operatorname{rank}(A_k) - n, 0\right\} + \sum_{j=1}^{l-1} \min\left\{\sum_{k=j}^{l} \operatorname{rank}(A_{l-k}) - n, 0\right\}\Bigg) + 1. \end{aligned} \tag{4.86}$$

**Corollary 4.65** *Under the same assumption and the same notation as in Corollary 4.64, if $\tilde{m} \leq 0$, then the matrix polynomial $\sum_{i=0}^{l} \lambda^i A_i$ is singular.*

**Remark 4.66** The sufficient condition $\tilde{m} \leq 0$ for singularity in Corollary 4.65 is equivalent to the sufficient condition for singularity in Proposition 4.49, since

$$\begin{aligned} &(\operatorname{rank}(A_l) + \operatorname{rank}(A_0) + (l-2)n \\ &\quad + \sum_{j=1}^{l-1} \min\left\{\sum_{k=j}^{l} \operatorname{rank}(A_k) - n, 0\right\} + \sum_{j=1}^{l-1} \min\left\{\sum_{k=j}^{l} \operatorname{rank}(A_{l-k}) - n, 0\right\}\Bigg) + 1 \leq 0 \\ \Longleftrightarrow\quad &(\operatorname{rank}(A_l) + \operatorname{rank}(A_0) + (l-2)n \\ &\quad + \sum_{j=1}^{l-1} \sum_{k=j}^{l} (\operatorname{rank}(A_k) - n) + \sum_{j=1}^{l-1} \sum_{k=j}^{l} (\operatorname{rank}(A_{l-k}) - n)\Bigg) + 1 \leq 0 \\ \Longleftrightarrow\quad &\sum_{i=0}^{l} \operatorname{rank}(A_i) - n + \frac{1}{l} \leq 0 \\ \Longleftrightarrow\quad &\sum_{i=0}^{l} \operatorname{rank}(A_i) - n + 1 \leq 0. \qquad\qquad \lozenge \end{aligned}$$

Although it is a widely known fact that it is possible to study $p$th ($p \geq 2$) degree square matrix polynomials, or more precisely, their corresponding polynomial eigenvalue problems, via a *linearization* method (see [17]), it should be noted that, for a regular $p$th ($p \geq 2$) degree polynomial eigenvalue problem, the upper bound of the number of its nonzero eigenvalues derived in (4.80) in Theorem 4.62 may be much sharper than that derived, after a linearization, in (4.64) in Theorem 4.52, as the next remark shows.

**Remark 4.67** Given a square and *regular* matrix polynomial $A(\lambda) = \sum_{i=0}^{l} \lambda^i A_i$ of degree $l \geq 2$, where $A_i \in \mathbb{C}^{n\times n}$, $i = 0, 1, \ldots, l$, we can *linearize* $A(\lambda)$ in the classical way into its *companion polynomial* $C_A(\lambda)$ (cf. [17], p. 186):

$$C_A(\lambda) := \mathcal{A}_1\lambda + \mathcal{A}_0 = \begin{bmatrix} I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \\ 0 & 0 & \cdots & 0 & A_l \end{bmatrix} \lambda + \begin{bmatrix} 0 & -I & 0 & \cdots & 0 \\ 0 & 0 & -I & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & -I \\ A_0 & A_1 & \cdots & \cdots & A_{l-1} \end{bmatrix}. \tag{4.87}$$

Since $C_A(\lambda)$ is a *linearization* of $A(\lambda)$, we have

$$E(\lambda)C_A(\lambda)F(\lambda) = \operatorname{diag}[A(\lambda), I_{n(l-1)}], \tag{4.88}$$

where $E(\lambda)$ and $F(\lambda)$ are $nl \times nl$ unimodular matrix polynomials, or in other words, the determinants of $E(\lambda)$ and $F(\lambda)$ are nonzero constants. From (4.88) it follows that

$$\det(C_A(\lambda)) = c_0 \cdot \det(A(\lambda)),$$

with $c_0 \in \mathbb{C}\backslash\{0\}$. Therefore, $\det(C_A(\lambda))$ has exactly the same roots as $\det(A(\lambda))$. If we apply Theorem 4.52 to the linearization $C_A(\lambda)$ as a matrix pencil to bound the number $m$ of its nonzero roots, we obtain that

$$0 \leq m \leq \operatorname{rank}(\mathcal{A}_1) + \operatorname{rank}(\mathcal{A}_0) - ln = \operatorname{rank}(A_l) + \operatorname{rank}(A_0) + (l-2)n, \tag{4.89}$$

since by (4.87) we have

$$\begin{aligned} \operatorname{rank}(\mathcal{A}_1) &= (l-1)n + \operatorname{rank}(A_l), & (4.90)\\ \operatorname{rank}(\mathcal{A}_0) &= (l-1)n + \operatorname{rank}(A_0), & (4.91) \end{aligned}$$

which do not take account of the rank information of $A_1, \ldots, A_{l-1}$. However, if we directly apply Theorem 4.62 to $A(\lambda)$, we can derive a much sharper upper bound on $m$ shown in (4.80) than that in (4.89). Take, for instance, Example 4.61, where $l = 2$. By the upper bound in (4.89) derived after linearization, we have $m \leq 1 + 1 + 0 = 2$, whereas by Theorem 4.62, we have $m \leq 0$.

The same result, as above, can also be available in the case of the so-called *decomposable linearization* $T_A(\lambda)$, which preserves the full spectral information including spectrum at infinity of the original matrix polynomial $A(\lambda)$. The reason is due to the following relation between the companion polynomial $C_A(\lambda)$ and the decomposable linearization $T_A(\lambda)$:

$$C_A(\lambda)S_1 = S_2T_A(\lambda),$$

where $S_1, S_2 \in \mathbb{C}^{nl\times nl}$ are nonsingular matrices. For more details of linearizations and spectral properties of regular matrix polynomials, we refer to [17] (Chapter 7), and more recently, [41] which deals with structure-preserved linearizations. $\Diamond$

Finally, we conclude this section with the following formal procedure which summarizes the results obtained in Proposition 4.13, Corollary 4.37, and Corollaries 4.64 and 4.65 to test regularity or singularity for any given square matrix polynomial.

**Algorithm 4.68** *Given a matrix polynomial $A(\lambda) = \sum_{i=0}^{l} \lambda^i A_i$ of degree $l$, where $A_i \in \mathbb{C}^{n\times n}$, $i = 0, 1, \ldots, l$, and $l \in \mathbb{N}$, this algorithm determines whether or not the matrix polynomial is regular.*

1. *Check the rank information of the leading and last coefficient matrices of $A(\lambda)$.*

   If $\quad \mathrm{rank}(A_l) = n$ or $\mathrm{rank}(A_0) = n$
   $\qquad A(\lambda)$ is regular (by Corollary 4.37), return
   end

2. *Compute $\tilde{m}$ defined by (4.86) in Corollary 4.64.*

3. *Test for regularity or singularity by using $\tilde{m}$.*

```
If     m̃ ≤ 0
       A(λ) is singular (by Corollary 4.65), return
else
       randomly choose distinct nonzero numbers λ̃_1, ..., λ̃_m̃
        for   i = 1 : m̃
              if    det(A(λ̃_i)) ≠ 0
                    A(λ) is regular (by Proposition 4.13), return
              end
        end
       A(λ) is singular (by Corollary 4.64)
end
```

◊

It should be noted that, if the rank of some coefficient matrix $A_i$, $i \neq 0, l$, is not exactly known, it is also valid to substitute its upper bound, if we know, or even $n$ for $\mathrm{rank}(A_i)$ in (4.86) to compute $\tilde{m}$.

## 4.4 Nearness to Singularity Problem for Matrix Polynomials

### 4.4.1 Introduction

In this section we investigate the distance to the nearest singular square matrix polynomial, i.e., for a square and regular matrix polynomial we are interested in by how much its coefficient matrices must be perturbed for the regularity to be lost. This question of interest mainly arises, as we have pointed out in the last section, from polynomial eigenvalue problems. For a *nearly* singular but regular matrix polynomial, the eigenvalues and eigenvectors of its corresponding polynomial eigenvalue problem may become very sensitive to perturbations of the coefficient matrices even if the matrix polynomial after perturbation is still regular, as the following example illustrates.

**Example 4.69** We consider quadratic matrix polynomials

$$\begin{aligned} A(\lambda) &= \lambda^2 A_2 + \lambda A_1 + A_0 := \lambda^2 \begin{bmatrix} 0 & \epsilon \\ 1 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \text{ and} \\ \hat{A}(\lambda) &= \lambda^2 \hat{A}_2 + \lambda \hat{A}_1 + \hat{A}_0 := \lambda^2 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \end{aligned}$$

where $|\epsilon| \ll 1$. From Theorem 4.47 it follows that $\hat{A}(\lambda)$ is singular, and $A(\lambda)$ is regular if and only if $\epsilon \neq 0$. If $\epsilon \neq 0$, it is immediate that the quadratic eigenvalue problem $A(\lambda)x = (\lambda^2 A_2 + \lambda A_1 + A_0)x = 0$ has eigenvalues

$$\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0 \tag{4.92}$$

with the sole eigenvector $x = [1, 0]^T$. If we perturb the coefficient matrices of $A(\lambda)$ to get the following perturbed quadratic matrix polynomial

$$\tilde{A}(\lambda) = \lambda^2 \tilde{A}_2 + \lambda \tilde{A}_1 + \tilde{A}_0 := \lambda^2 \begin{bmatrix} \alpha & 0 \\ 1 & -\alpha \end{bmatrix} + \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ -\beta & 0 \end{bmatrix},$$

then, by Theorem 4.47, $\tilde{A}(\lambda)$ is regular if and only if $\alpha \neq 0$ *or* $\beta \neq 0$. If $\alpha \neq 0$ and $\beta \neq 0$, a short computation shows that the regular quadratic eigenvalue problem $\tilde{A}(\lambda)\tilde{x} = (\lambda^2 \tilde{A}_2 + \lambda \tilde{A}_1 + \tilde{A}_0)\tilde{x} = 0$ has eigenvalues

$$\tilde{\lambda}_i = \frac{\sqrt[4]{|\beta|}}{\sqrt{|\alpha|}} e^{\sqrt{-1}\theta_i}, \text{ where } \theta_i = \frac{1}{4}\arg(\frac{\beta}{\alpha^2}) + i \cdot \frac{\pi}{2}, \; i = 1, \ldots, 4, \tag{4.93}$$

with corresponding eigenvectors

$$\tilde{x}_i = \left[ 1, -\sqrt{|\beta|} e^{\sqrt{-1}2\theta_i} - \frac{\sqrt[4]{|\beta|}}{\sqrt{|\alpha|}} e^{\sqrt{-1}\theta_i} \right]^T, \; i = 1, \ldots, 4, \tag{4.94}$$

respectively. Clearly, if

$$1 \gg |\epsilon| > 0, \; 1 \gg |\alpha| > 0, \text{ and } 1 \gg |\beta| > 0, \tag{4.95}$$

then the *distance* between $A(\lambda)$ and $\hat{A}(\lambda)$

$$\text{dis } (A(\lambda), \hat{A}(\lambda)) := \sqrt{\sum_{i=0}^{2} \|A_i - \hat{A}_i\|_F^2} = |\epsilon| \tag{4.96}$$

is very small, and therefore $A(\lambda)$ is *nearly* singular. Also, under the condition (4.95), the distance between $A(\lambda)$ and $\tilde{A}(\lambda)$

$$\text{dis }(A(\lambda), \tilde{A}(\lambda)) = \sqrt{\sum_{i=0}^{2} \|A_i - \tilde{A}_i\|_F^2} = \sqrt{|\epsilon|^2 + 2|\alpha|^2 + |\beta|^2} \tag{4.97}$$

is very small. However, if we let, in addition to the condition (4.95), $|\beta| \approx |\alpha|$, then, by (4.93) and (4.94), the perturbed eigenvalues $\tilde{\lambda}_i$, $i = 1, \ldots, 4$, may vary from 0 drastically, and so may the perturbed eigenvectors $\tilde{x}_i$, $i = 1, \ldots, 4$, from $x$. $\Diamond$

The distance to the nearest non-regular square matrix *pencils* has been investigated by Byers, He, and Mehrmann in [4]. In Subsection 4.4.2, we will give a definition and some properties of the distance to the nearest singular square matrix *polynomials*, and, based on the results obtained in Subsection 4.2.3, we will present a general theoretical characterization for the distance. Both the definition and the characterization can be regarded as natural generalizations of those given in [4]. We will also show that the nearness problem is in fact a perturbation-structured and -constrained rank-deficiency problem which appears to be an open problem. Moreover, at the end of Subsection 4.4.2, a characterization of the nearness for matrix pencils will be given, which implies a coincidence with the geometrical characterization (4.36) for singular matrix polynomials obtained in Subsection 4.2.4. Subsection 4.4.3 deals with two special cases of matrix polynomials. For each of the two, an explicit formula for the nearest distance is determined. In particular, an example, in which the nearest distance in the spectral norm can be strictly less than that in the Frobenius norm, is given. Finally, in Subsection 4.4.4, in terms of the smallest singular value of a matrix, we will derive two types of lower bounds on the nearest distance for general regular matrix polynomials, which are the generalizations of the results for matrix pencils obtained in [4].

### 4.4.2 Properties and Theoretical Characterizations of the Nearness to Singularity

First of all, we give a definition of the nearness to singularity for matrix polynomials.

**Definition 4.70** *Given a square matrix polynomial $A(\lambda) = \sum_{i=0}^{l} \lambda^i A_i$ of degree $l$, where $A_i \in \mathbb{C}^{n\times n}$, $i = 0, 1, \ldots, l$, and $l \in \mathbb{N}_0$, the distance to the nearest singular matrix polynomial is defined by*

$$\delta_p(A(\lambda)) := \min \left\{ \|[\Delta A_l, \ldots, \Delta A_0]\|_p \;\middle|\; A(\lambda) + \Delta A(\lambda) \text{ is singular}, \Delta A(\lambda) := \sum_{i=0}^{l} \lambda^i \Delta A_i \right\}, \tag{4.98}$$

*where $\|(\cdot)\|_p$ denotes the matrix 2-norm (spectral norm) if $p = 2$, or, the Frobenius matrix norm if $p = F$.*

Let $\Delta := \{[\Delta A_l, \ldots, \Delta A_0] \mid A(\lambda) + \Delta A(\lambda) \text{ is singular}\}$ be a subset of $\mathbb{C}^{n\times(l+1)n}$. Obviously, $[-A_l, \ldots, -A_0] \in \Delta$; therefore, $\Delta$ is not empty. Since, by Proposition 4.13, the singularity of square matrix polynomials can be derived in terms of a determinant which is a continuous function in the entries of a matrix, the property of singularity of matrix polynomials makes $\Delta$ a closed subset of $\mathbb{C}^{n\times(l+1)n}$. Hence, the distance function $\delta_p(A(\lambda))$ in (4.98) is well-defined.

We call a matrix polynomial $\Delta A(\lambda) = \sum_{i=0}^{l} \lambda^i \Delta A_i$, $\Delta A_i \in \mathbb{C}^{n\times n}$, $i = 0, \ldots, l$, a *minimum p-norm de-regularizing perturbation of* $A(\lambda)$, if $A(\lambda) + \Delta A(\lambda)$ is singular and $\delta_p(A(\lambda)) = \|[\Delta A_l, \ldots, \Delta A_0]\|_p$, $p = 2, F$. From the above analysis, we know that for any given $A(\lambda)$, there exists at least one minimum $p$-norm de-regularizing perturbation of $A(\lambda)$.

### Some Properties of the Nearness

The distances in the spectral norm and in the Frobenius norm are *equivalent* to each other, as the following Proposition 4.71 shows.

**Proposition 4.71**

$$\delta_F(A(\lambda)) \geq \delta_2(A(\lambda)) \geq \frac{1}{\sqrt{n}} \cdot \delta_F(A(\lambda)).$$

**Proof**. Assume that matrix polynomial $\Delta A(\lambda) = \sum_{i=0}^{l} \lambda^i \Delta A_i$ is a minimum $F$-norm de-regularizing perturbation of $A(\lambda)$. Since $A(\lambda) + \Delta A(\lambda)$ is singular, by Definition 4.70, we have

$$\delta_2(A(\lambda)) \leq \|[\Delta A_l, \ldots, \Delta A_0]\|_2 \leq \|[\Delta A_l, \ldots, \Delta A_0]\|_F = \delta_F(A(\lambda)).$$

Similarly, we assume that matrix polynomial $\Delta A(\lambda) = \sum_{i=0}^{l} \lambda^i \Delta A_i$ is a minimum 2-norm de-regularizing perturbation of $A(\lambda)$. Since $A(\lambda) + \Delta A(\lambda)$ is singular, by Definition 4.70, we have

$$\delta_F(A(\lambda)) \leq \|[\Delta A_l, \ldots, \Delta A_0]\|_F \leq \sqrt{n}\, \|[\Delta A_l, \ldots, \Delta A_0]\|_2 = \sqrt{n} \cdot \delta_2(A(\lambda)).$$

$\square$

In the next subsection, we will give an example in which $\delta_2(A(\lambda))$ is (strictly) less than $\delta_F(A(\lambda))$.

Like the spectral and Frobenius matrix norms, $\delta_p(A(\lambda))$ is *unitarily invariant*, as the following proposition shows.

**Proposition 4.72** *For any unitary matrices $U, V \in \mathbb{C}^{n\times n}$ and $p = 2, F$,*

$$\delta_p(U^H A(\lambda) V) = \delta_p(A(\lambda)).$$

**Proof**. The proof follows from the fact that $U^H(A(\lambda)+\Delta A(\lambda))V$ is singular if and only if $A(\lambda)+\Delta A(\lambda)$ is singular, and the fact that

$$\begin{aligned} & \left\| \left[U^H\Delta A_l V, \ldots, U^H\Delta A_0 V\right] \right\|_p \\ = & \left\| U^H \left[\Delta A_l, \ldots, \Delta A_0\right] \begin{bmatrix} V & & \\ & \ddots & \\ & & V \end{bmatrix} \right\|_p \\ = & \left\| \left[\Delta A_l, \ldots, \Delta A_0\right] \right\|_p, \ p = 2, F. \end{aligned}$$

□

The following proposition which describes the relation between the $F$-norm distance of $A(\lambda)$ to the nearest singular matrix polynomial and that of its conjugate transpose.

**Proposition 4.73**

$$\delta_F(A(\lambda)) = \delta_F(A^H(\lambda)). \tag{4.99}$$

**Proof**. The proof follows from the fact that $A(\lambda)+\Delta A(\lambda)$ is singular if and only if $A^H(\lambda)+(\Delta A)^H(\lambda)$ is singular, and the fact that

$$\left\| \left[\Delta A_l, \ldots, \Delta A_0\right] \right\|_F = \left\| \left[(\Delta A_l)^H, \ldots, (\Delta A_0)^H\right] \right\|_F.$$

□

To investigate the relation between $\delta_2(A(\lambda))$ and $\delta_2(A^H(\lambda))$, we need the following lemma.

**Lemma 4.74** *Let $B_i \in \mathbb{C}^{m\times n}$, $i = 1, \ldots, l$. Then*

$$\left\| \left[B_l, \ldots, B_1\right] \right\|_2 \leq \min\left\{\sqrt{l}, \sqrt{n}\right\} \cdot \left\| \begin{bmatrix} B_l \\ \vdots \\ B_1 \end{bmatrix} \right\|_2, \tag{4.100}$$

*and*

$$\left\| \begin{bmatrix} B_l \\ \vdots \\ B_1 \end{bmatrix} \right\|_2 \leq \min\left\{\sqrt{l}, \sqrt{n}\right\} \cdot \left\| \left[B_l, \ldots, B_1\right] \right\|_2. \tag{4.101}$$

**Proof**. The proof of inequality (4.100) follows from the following two inequalities (4.102) and (4.103):

$$\begin{aligned}
\|[B_l,\ldots,B_1]\|_2^2 &= \lambda_{\max}\left([B_l,\ldots,B_1]\cdot\begin{bmatrix} B_l^H \\ \vdots \\ B_1^H \end{bmatrix}\right) \\
&= \lambda_{\max}\left(B_lB_l^H+\cdots+B_1B_1^H\right) \\
&\le \lambda_{\max}(B_lB_l^H)+\cdots+\lambda_{\max}(B_1B_1^H) \\
&= \lambda_{\max}(B_l^HB_l)+\cdots+\lambda_{\max}(B_1^HB_1) \\
&\le l\cdot\lambda_{\max}\left(B_l^HB_l+\cdots+B_1^HB_1\right) \\
&= l\cdot\lambda_{\max}\left(\left[B_l^H,\ldots,B_1^H\right]\cdot\begin{bmatrix} B_l \\ \vdots \\ B_1 \end{bmatrix}\right) \\
&= l\cdot\left\|\begin{bmatrix} B_l \\ \vdots \\ B_1 \end{bmatrix}\right\|_2^2 .
\end{aligned} \tag{4.102}$$

$$\|[B_l,\ldots,B_1]\|_2^2 \le \|[B_l,\ldots,B_1]\|_F^2 = \left\|\begin{bmatrix} B_l \\ \vdots \\ B_1 \end{bmatrix}\right\|_F^2 \le \min\{lm,n\}\cdot\left\|\begin{bmatrix} B_l \\ \vdots \\ B_1 \end{bmatrix}\right\|_2^2 . \tag{4.103}$$

Since, for any $A\in\mathbb{C}^{m\times n}$, $\|A\|_2=\|A^H\|_2$, inequality (4.101) immediately follows from inequality (4.100). □

In inequalities (4.100) and (4.101), equality can be attained, as the following example shows.

**Example 4.75** Let $B_1 := \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and let $B_2 := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Then $l=n=2$, and

$$\begin{aligned}
\|[B_2,B_1]\|_2 = \left\|\begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right\|_2 &= \lambda_{\max}^{\frac{1}{2}}\left(\begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\cdot\begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}\right) \\
&= \lambda_{\max}^{\frac{1}{2}}\left(\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}\right) = \sqrt{2};
\end{aligned}$$

whereas,

$$\left\| \begin{bmatrix} B_2 \\ B_1 \end{bmatrix} \right\|_2 = \left\| \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \right\|_2 = \lambda_{\max}^{\frac{1}{2}} \left( \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \right)$$
$$= \lambda_{\max}^{\frac{1}{2}} \left( \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) = 1.$$

Therefore, equality in inequality (4.100) can be attained, and so can equality in inequality (4.101) if we let $B_1 := \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and let $B_2 := \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$. $\Diamond$

**Proposition 4.76**

$$\delta_2(A^H(\lambda)) \leq \min\left\{\sqrt{l+1}, \sqrt{n}\right\} \cdot \delta_2(A(\lambda)). \tag{4.104}$$

**Proof**. We assume that $\Delta A(\lambda) = \sum_{i=0}^{l} \lambda^i \Delta A_i$ is a minimum 2-norm de-regularizing perturbation of $A(\lambda)$, i.e., $A(\lambda) + \Delta A(\lambda)$ is singular and $\delta_2(A(\lambda)) = \|[\Delta A_l, \ldots, \Delta A_0]\|_2$. Then $A^H(\lambda) + (\Delta A)^H(\lambda)$ is singular. Therefore, by Definition 4.70 and Lemma 4.74, we have

$$\begin{aligned}
\delta_2(A^H(\lambda)) \leq \left\| \left[ (\Delta A_l)^H, \ldots, (\Delta A_0)^H \right] \right\|_2 &= \left\| \begin{bmatrix} \Delta A_l \\ \vdots \\ \Delta A_0 \end{bmatrix} \right\|_2 \\
&\leq \min\left\{\sqrt{l+1}, \sqrt{n}\right\} \cdot \|[\Delta A_l, \ldots, \Delta A_0]\|_2 \\
&= \min\left\{\sqrt{l+1}, \sqrt{n}\right\} \cdot \delta_2(A(\lambda)).
\end{aligned}$$

$\square$

### A General Characterization for the Nearness to Singularity

In order to determine a characterization for $\delta_p(A(\lambda))$, $p = 2, F$, let us recall the necessary and sufficient conditions for regularity of matrix polynomials which have been presented in terms of coefficient matrices in Corollary 4.39 and Remark 4.40. Observing the special forms of $W_s(A_l, A_{l-1}, \ldots, A_0)$ in (4.32) and of $\widehat{W}_s(A_l, A_{l-1}, \ldots, A_0)$ in (4.34), we immediately have the following proposition.

**Proposition 4.77** *Let $A(\lambda) = \sum_{i=0}^{l} \lambda^i A_i$ be a matrix polynomial of degree $l$, where $A_i \in \mathbb{C}^{n\times n}$, $i = 0, 1, \ldots, l$, and $l \in \mathbb{N}_0$. Then the following statements are equivalent:*

1. *$A(\lambda)$ is regular.*

2. *The block matrix $W_{(n-1)l+1}(A_l, A_{l-1}, \ldots, A_0)$ has full column rank $n((n-1)l+1)$, where $W_{(n-1)l+1}(A_l, A_{l-1}, \ldots, A_0)$ is defined as in (4.32) in Corollary 4.39.*

3. *The block matrix $\widehat{W}_{(n-1)l+1}(A_l, A_{l-1}, \ldots, A_0)$ has full row rank $n((n-1)l+1)$, where $\widehat{W}_{(n-1)l+1}(A_l, A_{l-1}, \ldots, A_0)$ is defined as in (4.34) in Remark 4.40.*

**Proof**. 1. $\Rightarrow$ 2.: By Corollary 4.39, the proof is immediate.

2. $\Rightarrow$ 1.: Observe that if the block matrix $W_s(A_l, A_{l-1}, \ldots, A_0)$ in (4.32) has full column rank, then for every $s \in \mathbb{N}$, $s \leq (n-1)l$, the block matrix $W_s(A_l, A_{l-1}, \ldots, A_0)$ has also full column rank. Thus, by Corollary 4.39, $A(\lambda)$ is column-regular and therefore regular.

The proofs of 1. $\Rightarrow$ 3. and 3. $\Rightarrow$ 1. are analogous to those of 1. $\Rightarrow$ 2. and 2. $\Rightarrow$ 1., respectively. $\square$

From the regularity conditions given in Corollary 4.39, Remark 4.40 and Proposition 4.77, we directly derive a characterization for the distance $\delta_p(A(\lambda))$, as follows.

**Proposition 4.78** *Let $A(\lambda) = \sum_{i=0}^{l} \lambda^i A_i$ be a matrix polynomial of degree $l$, where $A_i \in \mathbb{C}^{n\times n}$, $i = 0, 1, \ldots, l$, and $l \in \mathbb{N}_0$. Then,*

$$\begin{aligned}
\delta_p(A(\lambda)) &= \min_s \min_\Delta \{\| [\Delta A_l, \ldots, \Delta A_0] \|_p \mid \\
&\qquad \operatorname{rank}\left(W_s\left(A_l + \Delta A_l, A_{l-1} + \Delta A_{l-1}, \ldots, A_0 + \Delta A_0\right)\right) < ns\} \quad (4.105)\\
&= \min_\Delta \{\| [\Delta A_l, \ldots, \Delta A_0] \|_p \mid \\
&\qquad \operatorname{rank}\left(W_{\tilde{s}}\left(A_l + \Delta A_l, A_{l-1} + \Delta A_{l-1}, \ldots, A_0 + \Delta A_0\right)\right) < n\tilde{s}\} \quad (4.106)\\
&= \min_s \min_\Delta \{\| [\Delta A_l, \ldots, \Delta A_0] \|_p \mid \\
&\qquad \operatorname{rank}\left(\widehat{W}_s\left(A_l + \Delta A_l, A_{l-1} + \Delta A_{l-1}, \ldots, A_0 + \Delta A_0\right)\right) < ns\Big\} \quad (4.107)\\
&= \min_\Delta \{\| [\Delta A_l, \ldots, \Delta A_0] \|_p \mid \\
&\qquad \operatorname{rank}\left(\widehat{W}_{\tilde{s}}\left(A_l + \Delta A_l, A_{l-1} + \Delta A_{l-1}, \ldots, A_0 + \Delta A_0\right)\right) < n\tilde{s}\Big\} \quad (4.108)
\end{aligned}$$

*where $s \in \mathbb{N} : 1 \leq s \leq (n-1)l + 1$, $\tilde{s} := (n-1)l + 1$, the $(s+l)$-by-$s$ block matrix*

$W_s(A_l + \Delta A_l, A_{l-1} + \Delta A_{l-1}, \ldots, A_0 + \Delta A_0)$ *is given by*

$$
\begin{aligned}
& W_s(A_l + \Delta A_l, A_{l-1} + \Delta A_{l-1}, \ldots, A_0 + \Delta A_0) \\
= & \begin{bmatrix}
A_0+\Delta A_0 & & & & & & \\
A_1+\Delta A_1 & A_0+\Delta A_0 & & & & & \\
\vdots & \vdots & \ddots & & & & \\
A_{l-1}+\Delta A_{l-1} & A_{l-2}+\Delta A_{l-2} & \cdots & A_0+\Delta A_0 & & & \\
A_l+\Delta A_l & A_{l-1}+\Delta A_{l-1} & \cdots & A_1+\Delta A_1 & A_0+\Delta A_0 & & \\
 & \ddots & \ddots & \ddots & \ddots & \ddots & \\
 & & A_l+\Delta A_l & A_{l-1}+\Delta A_{l-1} & \cdots & A_1+\Delta A_1 & A_0+\Delta A_0 \\
 & & & A_l+\Delta A_l & \cdots & A_2+\Delta A_2 & A_1+\Delta A_1 \\
 & & & & \ddots & \vdots & \vdots \\
 & & & & & A_l+\Delta A_l & A_{l-1}+\Delta A_{l-1} \\
 & & & & & & A_l+\Delta A_l
\end{bmatrix},
\end{aligned}
\tag{4.109}
$$

*and the s-by-$(s+l)$ block matrix* $\widehat{W}_s(A_l + \Delta A_l, A_{l-1} + \Delta A_{l-1}, \ldots, A_0 + \Delta A_0)$ *is given by*

$$
\begin{aligned}
& \widehat{W}_s(A_l + \Delta A_l, A_{l-1} + \Delta A_{l-1}, \ldots, A_0 + \Delta A_0) \\
= & \begin{bmatrix}
A_0+\Delta A_0 & A_1+\Delta A_1 & \cdots & A_l+\Delta A_l & & & & & \\
 & A_0+\Delta A_0 & \cdots & A_{l-1}+\Delta A_{l-1} & A_l+\Delta A_l & & & & \\
 & & \ddots & \ddots & \ddots & \ddots & & & \\
 & & & & \ddots & \ddots & \ddots & \ddots & \\
 & & & & & A_0+\Delta A_0 & \cdots & A_{l-1}+\Delta A_{l-1} & A_l+\Delta A_l
\end{bmatrix}.
\end{aligned}
\tag{4.110}
$$

**Remark 4.79** Note that if $l = 0$ then $A(\lambda) \equiv A_0$, and therefore the nearness problem degenerates into the rank-deficiency problem for a general matrix, which has well-known classical result in terms of singular values of the matrix (see, for example, [20], or Subsection 4.4.3); whereas, if $l \geq 1$, Proposition 4.78 shows that the nearness problem becomes a perturbation-structured and constrained rank-deficiency problem, for which we still do not know how to determine a general explicit formula, except in some cases in which both the order and degree of $A(\lambda)$ are very small (see Proposition 4.88 in Subsection 4.4.3), or, the coefficient matrices of $A(\lambda)$ have very special forms or properties (e.g. they are scalar multiples of one another). For a very brief survey of rank-deficiency nearness problem for matrices, see also [20]. $\Diamond$

### A Characterization of the Nearness to Singularity for Matrix Pencils

In the case of matrix pencils, another kind of characterization for $\delta_F(A(\lambda))$ is, in the light of the *generalized Schur form* ([53, 45, 55, 19, 4]), obtained in [4]. Before extending the result to the 2-norm case, let us recall the generalized Schur form for matrix pencils.

**Theorem 4.80 (Generalized Schur form)** *[53, 45] For every matrix pencil $\lambda A_1 + A_0$, $A_1, A_0 \in \mathbb{C}^{n\times n}$, there exist unitary matrices $Q, Z \in \mathbb{C}^{n\times n}$ such that*

$$Q^H(\lambda A_1 + A_0)Z = \lambda R + S, \tag{4.111}$$

*where $R, S \in \mathbb{C}^{n\times n}$ are upper-triangular matrices.*

In the following lemma, we restate the Frobenius norm result in [4] and extend it to the 2-norm case.

**Lemma 4.81** *([4], Theorem 8) Let $A(\lambda) = \lambda A_1 + A_0$, where $A_1, A_0 \in \mathbb{C}^{n\times n}$. Then, for $p = 2, F$,*

$$\delta_p(A(\lambda)) = \min_{1\le k\le n} \min_{\substack{Q_{n-k+1}\in\mathbb{C}^{(n-k+1)\times n},\\ Q_{n-k+1}Q_{n-k+1}^H = I}} \min_{\substack{Z_k\in\mathbb{C}^{n\times k},\\ Z_k^H Z_k = I}} \|[Q_{n-k+1}A_1Z_k, Q_{n-k+1}A_0Z_k]\|_p . \tag{4.112}$$

**Proof**. In the case of the Frobenius norm $\delta_F(A(\lambda))$, see the proof given in [4].

In the case of the 2-norm, the part which proves that $\delta_2(A(\lambda))$ is bounded from above by the right-hand-side of (4.112) is the same as that in the case of the Frobenius norm. Next, we will prove that the right-hand-side of (4.112) is, conversely, also bounded from above by $\delta_2(A(\lambda))$. Assume that $\Delta A(\lambda) = \lambda\Delta A_1 + \Delta A_0$ is a minimum 2-norm de-regularizing perturbation of $A(\lambda)$, where $\Delta A_1, \Delta A_0 \in \mathbb{C}^{n\times n}$. Since $A(\lambda) + \Delta A(\lambda)$ is singular, the generalized Schur form of $A(\lambda) + \Delta A(\lambda)$ must have the following zero structure for some index $k$, $1 \le k \le n$:

$$Q\left(\lambda\left(A_1 + \Delta A_1\right) + \left(A_0 + \Delta A_0\right)\right)Z = \begin{array}{c c} & \begin{array}{cc} k & n-k \end{array} \\ \begin{array}{c} k-1 \\ n-k+1 \end{array} & \left[\begin{array}{cc} \lambda R_{11} + S_{11} & \lambda R_{12} + S_{12} \\ 0 & \lambda R_{22} + S_{12} \end{array}\right] \end{array}, \tag{4.113}$$

where $Q, Z \in \mathbb{C}^{n\times n}$ are unitary matrices. We partition $Q$ as $Q = \left[\begin{array}{c} Q_{k-1} \\ Q_{n-k+1} \end{array}\right]\mathbb{C}^{n\times n}$, $Z$ as $Z = [Z_k, Z_{n-k}]$, and $QA_1Z$ and $QA_0Z$ conformally with (4.113) as

$$QA_1Z = \begin{array}{c c} & \begin{array}{cc} k & n-k \end{array} \\ \begin{array}{c} k-1 \\ n-k+1 \end{array} & \left[\begin{array}{cc} A_{11}^{(1)} & A_{12}^{(1)} \\ A_{21}^{(1)} & A_{22}^{(1)} \end{array}\right] \end{array}, \; QA_0Z = \begin{array}{c c} & \begin{array}{cc} k & n-k \end{array} \\ \begin{array}{c} k-1 \\ n-k+1 \end{array} & \left[\begin{array}{cc} A_{11}^{(0)} & A_{12}^{(0)} \\ A_{21}^{(0)} & A_{22}^{(0)} \end{array}\right] \end{array}. \tag{4.114}$$

We also partition $Q\Delta A_1Z$ and $Q\Delta A_0Z$ conformally with (4.113) as

$$Q\Delta A_1Z = \begin{array}{c c} & \begin{array}{cc} k & n-k \end{array} \\ \begin{array}{c} k-1 \\ n-k+1 \end{array} & \left[\begin{array}{cc} \Delta A_{11}^{(1)} & \Delta A_{12}^{(1)} \\ \Delta A_{21}^{(1)} & \Delta A_{22}^{(1)} \end{array}\right] \end{array}, \; Q\Delta A_0Z = \begin{array}{c c} & \begin{array}{cc} k & n-k \end{array} \\ \begin{array}{c} k-1 \\ n-k+1 \end{array} & \left[\begin{array}{cc} \Delta A_{11}^{(0)} & \Delta A_{12}^{(0)} \\ \Delta A_{21}^{(0)} & \Delta A_{22}^{(0)} \end{array}\right] \end{array}. \tag{4.115}$$

Noting the zero structure in the right-hand-side of (4.113) and by (4.114) and (4.115), we have

$$\Delta A_{21}^{(1)} = -A_{21}^{(1)} = -Q_{n-k+1}A_1Z_k, \quad \Delta A_{21}^{(0)} = -A_{21}^{(0)} = -Q_{n-k+1}A_0Z_k. \tag{4.116}$$

Since $\Delta A(\lambda)$ is a minimum 2-norm de-regularizing perturbation of $A(\lambda)$, by (4.115) and (4.116), we have

$$\begin{aligned}
\delta_2(A(\lambda)) = \|[\Delta A_1, \Delta A_0]\|_2 &= \|[Q\Delta A_1 Z, Q\Delta A_0 Z]\|_2 \\
&= \left\| \begin{bmatrix} \Delta A_{11}^{(1)} & \Delta A_{12}^{(1)} & \Delta A_{11}^{(0)} & \Delta A_{12}^{(0)} \\ \Delta A_{21}^{(1)} & \Delta A_{22}^{(1)} & \Delta A_{21}^{(0)} & \Delta A_{22}^{(0)} \end{bmatrix} \right\|_2 \\
&\geq \left\| \left[\Delta A_{21}^{(1)}, \Delta A_{22}^{(1)}, \Delta A_{21}^{(0)}, \Delta A_{22}^{(0)}\right] \right\|_2 \\
&= \left\| \left[\Delta A_{21}^{(1)}, \Delta A_{21}^{(0)}, \Delta A_{22}^{(1)}, \Delta A_{22}^{(0)}\right] \right\|_2 \\
&\geq \left\| \left[\Delta A_{21}^{(1)}, \Delta A_{21}^{(0)}\right] \right\|_2 \\
&= \|[-Q_{n-k+1}A_1Z_k, -Q_{n-k+1}A_0Z_k]\|_2 \\
&= \|[Q_{n-k+1}A_1Z_k, Q_{n-k+1}A_0Z_k]\|_2 .
\end{aligned} \tag{4.117}$$

Hence, the right-hand-side of (4.112) is bounded from above by $\delta_2(A(\lambda))$. Thus, finally, we obtain the formula (4.112). $\square$

We investigate further the characterization presented in Lemma 4.81, and obtain another simplified characterization in terms of the singular values of a family of matrices, as the following theorem shows.

**Theorem 4.82** *Let $A(\lambda) = \lambda A_1 + A_0$, where $A_1, A_0 \in \mathbb{C}^{n\times n}$. Then*

$$\delta_2(A(\lambda)) = \min_{1\leq k\leq n} \min_{\substack{Z\,\in\,\mathbb{C}^{n\times k},\\ Z^HZ=I}} \sigma_k\left([A_1Z, A_0Z]\right) \tag{4.118}$$

*and*

$$\delta_F(A(\lambda)) = \min_{1\leq k\leq n} \min_{\substack{Z\,\in\,\mathbb{C}^{n\times k},\\ Z^HZ=I}} \sqrt{\left(\sigma_k^2 + \cdots + \sigma_n^2\right)\left([A_1Z, A_0Z]\right)}, \tag{4.119}$$

*where $\sigma_k(B)$ denotes the $k$-th singular value of an $m\times n$ matrix $B$ with singular values in the descending order*

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min\{m,n\}} \geq \sigma_{\min\{m,n\}+1} = \cdots = \sigma_{\max\{m,n\}} = 0, \tag{4.120}$$

*and $\left(\sigma_k^2 + \cdots + \sigma_n^2\right)(B)$ denotes the sum of the squares of singular values of an $m\times n$ matrix $B$ from the $k$-th to the $n$-th if the singular values of the matrix $B$ are in the descending order as in (4.120).*

The proof of Theorem 4.82 will make use of the Wielandt's Theorem and the Courant-Fischer Theorem ([55]) for simplifying the characterization in (4.112).

**Theorem 4.83 (Wielandt)** *([55], p. 199-201) Let $A \in \mathbb{C}^{n\times n}$ be a Hermitian matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, and let $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, $1 \leq k \leq n$. Then*

$$\lambda_{i_1} + \lambda_{i_2} + \cdots + \lambda_{i_k} = \max_{\substack{\mathcal{X}_{i_j} \subseteq \mathbb{C}^n \\ \mathcal{X}_{i_1} \subset \mathcal{X}_{i_2} \subset \cdots \subset \mathcal{X}_{i_k} \\ \dim(\mathcal{X}_{i_j}) = i_j}} \min_{\substack{X = [x_{i_1}, x_{i_2}, \cdots, x_{i_k}], x_{i_j} \in \mathcal{X}_{i_j} \\ X^H X = I}} \operatorname{trace}(X^H A X), \tag{4.121}$$

*and*

$$\lambda_{i_1} + \lambda_{i_2} + \cdots + \lambda_{i_k} = \min_{\substack{\mathcal{X}_{i_j} \subseteq \mathbb{C}^n \\ \mathcal{X}_{i_1} \supset \mathcal{X}_{i_2} \supset \cdots \supset \mathcal{X}_{i_k} \\ \dim(\mathcal{X}_{i_j}) = n - i_j + 1}} \max_{\substack{X = [x_{i_1}, x_{i_2}, \cdots, x_{i_k}], x_{i_j} \in \mathcal{X}_{i_j} \\ X^H X = I}} \operatorname{trace}(X^H A X). \tag{4.122}$$

The Courant-Fischer Minimax Theorem, which is a direct consequence of Wielandt's Theorem, is as follows.

**Theorem 4.84 (Courant-Fischer)** *([55], p. 201) Let $A \in \mathbb{C}^{n\times n}$ be a Hermitian matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Then*

$$\lambda_i = \max_{\substack{\mathcal{X} \subseteq \mathbb{C}^n \\ \dim(\mathcal{X}) = i}} \min_{\substack{x \in \mathcal{X} \\ x^H x = 1}} x^H A x$$

*and*

$$\lambda_i = \min_{\substack{\mathcal{X} \subseteq \mathbb{C}^n \\ \dim(\mathcal{X}) = n - i + 1}} \max_{\substack{x \in \mathcal{X} \\ x^H x = 1}} x^H A x.$$

**Proof of Theorem 4.82**. By Lemma 4.81, for $p = 2, F$, we have

$$\delta_p(A(\lambda)) = \min_{1\leq k\leq n} \min_{\substack{Q \in \mathbb{C}^{(n-k+1)\times n}, \\ QQ^H = I}} \min_{\substack{Z \in \mathbb{C}^{n\times k}, \\ Z^H Z = I}} \left\| [QA_1 Z, QA_0 Z] \right\|_p. \tag{4.123}$$

If $p = 2$, by (4.123), we have

$$
\begin{aligned}
\delta_2^2(A(\lambda)) &= \min_{1\le k\le n} \min_{\substack{Q\in\mathbb{C}^{(n-k+1)\times n},\\ QQ^H=I}} \min_{\substack{Z\in\mathbb{C}^{n\times k},\\ Z^HZ=I}} \|[QA_1Z, QA_0Z]\|_2^2 \\
&= \min_{1\le k\le n} \min_{\substack{Q\in\mathbb{C}^{(n-k+1)\times n},\\ QQ^H=I}} \min_{\substack{Z\in\mathbb{C}^{n\times k},\\ Z^HZ=I}} \lambda_{\max}\left([QA_1Z, QA_0Z]\cdot \begin{bmatrix} Z^HA_1^HQ^H \\ Z^HA_0^HQ^H \end{bmatrix}\right) \\
&= \min_{1\le k\le n} \min_{\substack{Q\in\mathbb{C}^{(n-k+1)\times n},\\ QQ^H=I}} \min_{\substack{Z\in\mathbb{C}^{n\times k},\\ Z^HZ=I}} \lambda_{\max}\left(Q\left(A_1ZZ^HA_1^H + A_0ZZ^HA_0^H\right)Q^H\right) \\
&= \min_{1\le k\le n} \min_{\substack{Q\in\mathbb{C}^{(n-k+1)\times n},\\ QQ^H=I}} \min_{\substack{Z\in\mathbb{C}^{n\times k},\\ Z^HZ=I}} \\
&\qquad \left\{ \max_{\substack{x\in\mathbb{C}^{n-k+1},\\ x^Hx=1}} x^HQ\left(A_1ZZ^HA_1^H + A_0ZZ^HA_0^H\right)Q^Hx \right\} \\
&\qquad \text{(by Courant-Fischer's Theorem 4.84)} \\
&= \min_{1\le k\le n} \min_{\substack{Z\in\mathbb{C}^{n\times k},\\ Z^HZ=I}} \\
&\qquad \left\{ \min_{\substack{Q\in\mathbb{C}^{(n-k+1)\times n},\\ QQ^H=I}} \max_{\substack{x\in\mathbb{C}^{n-k+1},\\ x^Hx=1}} (Q^Hx)^H\left(A_1ZZ^HA_1^H + A_0ZZ^HA_0^H\right)Q^Hx \right\} \\
&= \min_{1\le k\le n} \min_{\substack{Z\in\mathbb{C}^{n\times k},\\ Z^HZ=I}} \\
&\qquad \left\{ \min_{\substack{\mathcal{Y}\subseteq\mathbb{C}^n,\\ \dim(\mathcal{Y})=n-k+1}} \max_{\substack{y\in\mathcal{Y},\\ y^Hy=1}} y^H\left(A_1ZZ^HA_1^H + A_0ZZ^HA_0^H\right)y \right\} \\
&= \min_{1\le k\le n} \min_{\substack{Z\in\mathbb{C}^{n\times k},\\ Z^HZ=I}} \lambda_k\left(A_1ZZ^HA_1^H + A_0ZZ^HA_0^H\right) \\
&\qquad \text{(again by Courant-Fischer's Theorem 4.84)} \\
&= \min_{1\le k\le n} \min_{\substack{Z\in\mathbb{C}^{n\times k},\\ Z^HZ=I}} \sigma_k\left([A_1Z, A_0Z]\cdot \begin{bmatrix} Z^HA_1^H \\ Z^HA_0^H \end{bmatrix}\right)
\end{aligned}
\tag{4.124}
$$

$$= \min_{1\le k\le n} \min_{\substack{Z \in \mathbb{C}^{n\times k},\\ Z^H Z = I}} \sigma_k^2\left([A_1 Z, A_0 Z]\right), \tag{4.125}$$

where $\lambda_k(\cdot)$ denotes the $k$-th eigenvalue of an $n \times n$ Hermitian matrix with eigenvalues in descending order $\lambda_{\max} = \lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$.

If $p = F$, by (4.123), we have

$$\begin{aligned}
\delta_F^2(A(\lambda)) &= \min_{1\le k\le n} \min_{\substack{Q \in \mathbb{C}^{(n-k+1)\times n},\\ QQ^H = I}} \min_{\substack{Z \in \mathbb{C}^{n\times k},\\ Z^H Z = I}} \|[QA_1Z, QA_0Z]\|_F^2 \\
&= \min_{1\le k\le n} \min_{\substack{Q \in \mathbb{C}^{(n-k+1)\times n},\\ QQ^H = I}} \min_{\substack{Z \in \mathbb{C}^{n\times k},\\ Z^H Z = I}} \operatorname{trace}\left([QA_1Z, QA_0Z] \cdot \begin{bmatrix} Z^H A_1^H Q^H \\ Z^H A_0^H Q^H \end{bmatrix}\right) \\
&= \min_{1\le k\le n} \min_{\substack{Q \in \mathbb{C}^{(n-k+1)\times n},\\ QQ^H = I}} \min_{\substack{Z \in \mathbb{C}^{n\times k},\\ Z^H Z = I}} \\
&\qquad \left\{\operatorname{trace}\left(Q\left(A_1 Z Z^H A_1^H + A_0 Z Z^H A_0^H\right) Q^H\right)\right\}.
\end{aligned} \tag{4.126}$$

Since the trace of a square matrix is equal to the sum of all its eigenvalues, from (4.126) it follows that

$$\begin{aligned}
\delta_F^2(A(\lambda)) &= \min_{1\le k\le n} \min_{\substack{Z \in \mathbb{C}^{n\times k},\\ Z^H Z = I}} \min_{\substack{Q \in \mathbb{C}^{(n-k+1)\times n},\\ QQ^H = I}} \\
&\qquad \left\{(\lambda_1 + \lambda_2 + \cdots + \lambda_{n-k+1})\left(Q\left(A_1 Z Z^H A_1^H + A_0 Z Z^H A_0^H\right) Q^H\right)\right\} \\
&= \min_{1\le k\le n} \min_{\substack{Z \in \mathbb{C}^{n\times k},\\ Z^H Z = I}} \min_{\substack{Q \in \mathbb{C}^{(n-k+1)\times n},\\ QQ^H = I}} \\
&\qquad \Bigg\{ \min_{\substack{\mathcal{X}_j \subseteq \mathbb{C}^{n-k+1}, j = 1, \ldots, n-k+1,\\ \mathcal{X}_1 \supset \mathcal{X}_2 \supset \cdots \supset \mathcal{X}_{n-k+1},\\ \dim(\mathcal{X}_j) = (n-k+1) - j + 1}} \max_{\substack{X = [x_1, x_2, \cdots, x_{n-k+1}], x_j \in \mathcal{X}_j,\\ X^H X = I}} \\
&\qquad \operatorname{trace}\left(X^H Q\left(A_1 Z Z^H A_1^H + A_0 Z Z^H A_0^H\right) Q^H X\right)\Bigg\} \\
&\qquad \text{(by Wielandt's Theorem 4.83)}
\end{aligned} \tag{4.127}$$

$$
\begin{aligned}
&= \min_{1\le k\le n} \min_{\substack{Z\in\mathbb{C}^{n\times k},\\ Z^HZ=I}} \\
&\quad \left\{ \min_{\substack{Q\in\mathbb{C}^{(n-k+1)\times n},\\ QQ^H=I}} \min_{\substack{\mathcal{X}_j\subseteq\mathbb{C}^{n-k+1}, j=1,\dots,n-k+1,\\ \mathcal{X}_1\supset\mathcal{X}_2\supset\cdots\supset\mathcal{X}_{n-k+1},\\ \dim(\mathcal{X}_j)=(n-k+1)-j+1}} \max_{\substack{X=[x_1,x_2,\cdots,x_{n-k+1}], x_j\in\mathcal{X}_j,\\ X^HX=I}} \right. \\
&\quad\quad \left. \mathrm{trace}\left(\left(Q^HX\right)^H\left(A_1ZZ^HA_1^H+A_0ZZ^HA_0^H\right)Q^HX\right)\right\} \\
&= \min_{1\le k\le n} \min_{\substack{Z\in\mathbb{C}^{n\times k},\\ Z^HZ=I}} \\
&\quad \left\{ \min_{\substack{\mathcal{Y}_j\subseteq\mathbb{C}^{n}, j=k,\dots,n,\\ \mathcal{Y}_k\supset\mathcal{Y}_{k+1}\supset\cdots\supset\mathcal{Y}_{n},\\ \dim(\mathcal{Y}_j)=n-j+1}} \max_{\substack{Y=[y_k,y_{k+1},\cdots,y_n], y_j\in\mathcal{Y}_j,\\ Y^HY=I}} \right. \\
&\quad\quad \left. \mathrm{trace}\left(Y^H\left(A_1ZZ^HA_1^H+A_0ZZ^HA_0^H\right)Y\right)\right\} \\
&= \min_{1\le k\le n} \min_{\substack{Z\in\mathbb{C}^{n\times k},\\ Z^HZ=I}} \left(\lambda_k+\cdots+\lambda_n\right)\left(A_1ZZ^HA_1^H+A_0ZZ^HA_0^H\right) \\
&\quad\quad \text{(again by Wielandt's Theorem 4.83)} \\
&= \min_{1\le k\le n} \min_{\substack{Z\in\mathbb{C}^{n\times k},\\ Z^HZ=I}} \left(\lambda_k+\cdots+\lambda_n\right)\left([A_1Z,A_0Z]\cdot\begin{bmatrix} Z^HA_1^H \\ Z^HA_0^H\end{bmatrix}\right) \\
&= \min_{1\le k\le n} \min_{\substack{Z\in\mathbb{C}^{n\times k},\\ Z^HZ=I}} \left(\sigma_k^2+\cdots+\sigma_n^2\right)\left([A_1Z,A_0Z]\right), \qquad (4.128)
\end{aligned}
$$

where $\left(\lambda_k+\cdots+\lambda_n\right)(B)$ denotes the sum of the eigenvalues of the Hermitian matrix $B$ from the $k$-th to the $n$-th if the eigenvalues of $B$ are in the descending order.

Thus, with (4.125) and (4.128), we have finished the proof. □

**Remark 4.85** By the classical result on the distance to the nearest rank-deficient matrix (see Theorem 4.87 in Subsection 4.4.3), we can interpret the quantities $\delta_2$ (and $\delta_F$, respectively) in (4.118) (and in (4.119), respectively) as the nearest distance, in terms of 2-norm (and $F$-norm, respectively), by which $\mathrm{rank}\left([A_1Z,A_0Z]\right)<k$ over all

$k$ and all $Z \in \mathbb{C}^{n\times k}$ with orthonormal columns. Thus, it is clear that the algebraic characterizations for the nearest distances given by (4.118) and (4.119) coincide with the geometrical characterization of square and singular matrix pencils which is presented in Theorem 4.44. $\lozenge$

**Remark 4.86** Let $\hat{k} \in \mathbb{N}$, $1 \leq \hat{k} \leq n$, and let $\hat{Z} \in \mathbb{C}^{n\times \hat{k}}$, $\hat{Z}^H\hat{Z} = I_{\hat{k}}$, be a test matrix. Then, from Theorem 4.82 it immediately follows that for any $Q \in \mathbb{C}^{\hat{k}\times\hat{k}}$, $QQ^H = I_{\hat{k}}$, we have the following upper bounds on the nearest distance to singularity.

$$\begin{aligned}
\delta_2(A(\lambda)) &\leq \sigma_{\hat{k}}\left(\left[A_1\hat{Z}, A_0\hat{Z}\right]\right) = \sigma_{\hat{k}}\left(\left[A_1\hat{Z}Q, A_0\hat{Z}Q\right]\right);\\
\delta_F(A(\lambda)) &\leq \sqrt{\left(\sigma_{\hat{k}}^2 + \cdots + \sigma_n^2\right)\left(\left[A_1\hat{Z}, A_0\hat{Z}\right]\right)} = \sqrt{\left(\sigma_{\hat{k}}^2 + \cdots + \sigma_n^2\right)\left(\left[A_1\hat{Z}Q, A_0\hat{Z}Q\right]\right)}.
\end{aligned}$$

$\lozenge$

### 4.4.3 Special Cases

In this subsection we discuss two special cases of matrix polynomials (in particular, of matrix pencils) in each of which an explicit formula for $\delta(A(\lambda))$ can be determined.

First, let us recall the classical result on the distance to the nearest rank-deficient matrix for a general rectangular matrix.

**Theorem 4.87** *[19, 20] Let $A \in \mathbb{C}^{m\times n}$ have the singular value decomposition (SVD)*

$$\begin{aligned}
A &= U^H \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V, \text{ where}\\
\Sigma &= \mathrm{diag}[\sigma_1, \ldots, \sigma_r],\ \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0,\ r = \mathrm{rank}(A),
\end{aligned}$$

*and let $k < r$. Then, for $p = 2, F$,*

$$\min_{\mathrm{rank}(B)=k} \|A - B\|_p = \|A - A_k\|_p = \begin{cases} \sigma_{k+1}, & p = 2,\\ \sqrt{\sum_{i=k+1}^r \sigma_i^2}, & p = F, \end{cases} \tag{4.129}$$

*where*

$$A_k = U^H \begin{bmatrix} \Sigma_k & 0 \\ 0 & 0 \end{bmatrix} V, \quad \Sigma_k = \mathrm{diag}[\sigma_1, \ldots, \sigma_k].$$

**Case I:** $A(\lambda) = (a_l \lambda^l + a_{l-1}\lambda^{l-1} + \cdots + a_0)A$, where $A \in \mathbb{C}^{n\times n}$.

In this trivial case, the coefficient matrices of $A(\lambda)$ are scalar multiples of one another. It is immediate that if $\Delta A$ is a minimum $p$-norm ($p = 2$ or $F$) perturbation of the matrix $A$ such that $A + \Delta A$ is singular, then the minimum $p$-norm de-regularizing perturbation of the matrix polynomial $A(\lambda)$ is $\Delta A(\lambda) = \sum_{i=0}^{l} \lambda^i \Delta A_i = (a_l\lambda^l + a_{l-1}\lambda^{l-1} + \cdots + a_0)\Delta A$. By Theorem 4.87, for $p = 2, F$, we have $\|\Delta A\|_p = \sigma_{\min}(A)$, and therefore, we have

$$\delta_p(A(\lambda)) = \|[\Delta A_l, \ldots, \Delta A_0]\|_p = \sqrt{|a_l|^2 + |a_{l-1}|^2 + \cdots + |a_0|^2} \cdot \sigma_{\min}(A), \;\; p = 2, F.$$

**Case II:** 2-by-2 Matrix Pencils.

In this case, we restate the Frobenius norm result in [4] and present upper and lower bounds on the distance in the spectral norm.

**Proposition 4.88** *([4], Corollary 3) Let $A_1, A_0 \in \mathbb{C}^{2\times 2}$. Then*

$$\delta_F(\lambda A_1 + A_0) = \min\left\{\sigma_{\min}\left(\begin{bmatrix} A_1 \\ A_0 \end{bmatrix}\right), \sigma_{\min}\left([A_1, A_0]\right)\right\}.$$

**Proof**. See [4]. $\square$

**Proposition 4.89** *Let $A_1, A_0 \in \mathbb{C}^{2\times 2}$. Then*

$$\begin{aligned} & \min\left\{\sigma_{\min}\left([A_1, A_0]\right), \frac{\sqrt{2}}{2}\sigma_{\min}\left(\begin{bmatrix} A_1 \\ A_0 \end{bmatrix}\right)\right\} \\ \leq \;& \delta_2(\lambda A_1 + A_0) \\ \leq \;& \min\left\{\sigma_{\min}\left([A_1, A_0]\right), \sigma_{\min}\left(\begin{bmatrix} A_1 \\ A_0 \end{bmatrix}\right)\right\}. \end{aligned} \tag{4.130}$$

**Proof**. Assume that $\lambda\Delta A_1 + \Delta A_0$ is a minimum 2-norm de-regularizing perturbation of $\lambda A_1 + A_0$, where $\Delta A_1, \Delta A_0 \in \mathbb{C}^{2\times 2}$, and that $\lambda (A_1 + \Delta A_1) + (A_0 + \Delta A_0)$ has the following generalized Schur form (cf. Theorem 4.80)

$$Q\left(\lambda\left(A_1 + \Delta A_1\right) + \left(A_0 + \Delta A_0\right)\right)Z = \lambda \begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{bmatrix} + \begin{bmatrix} s_{11} & s_{12} \\ 0 & s_{22} \end{bmatrix}. \tag{4.131}$$

Since $\lambda (A_1 + \Delta A_1) + (A_0 + \Delta A_0)$ is singular, we have that in (4.131) either $r_{11} = s_{11} = 0$ or $r_{22} = s_{22} = 0$. It follows that either the left null spaces, or the (right) null spaces, or both of $A_1 + \Delta A_1$ and $A_0 + \Delta A_0$ have a nontrivial intersection; or in other words, either $\begin{bmatrix} A_1 + \Delta A_1 \\ A_0 + \Delta A_0 \end{bmatrix}$, or $[A_1 + \Delta A_1, A_0 + \Delta A_0]$, or both are rank deficient. Hence, we have

$$\begin{aligned} \delta_2(\lambda A_1 + A_0) &= \|[\Delta A_1, \Delta A_0]\|_2 \\ &\geq \frac{1}{\sqrt{2}} \left\| \begin{bmatrix} \Delta A_1 \\ \Delta A_0 \end{bmatrix} \right\|_2 \quad \text{(by Lemma 4.74)} \\ &\geq \frac{\sqrt{2}}{2} \sigma_{\min} \left( \begin{bmatrix} A_1 \\ A_0 \end{bmatrix} \right) \quad \text{(by Theorem 4.87)} \end{aligned} \tag{4.132}$$

if $\begin{bmatrix} A_1 + \Delta A_1 \\ A_0 + \Delta A_0 \end{bmatrix}$ is rank deficient, and similarly, by Theorem 4.87, we have

$$\delta_2(\lambda A_1 + A_0) = \|[\Delta A_1, \Delta A_0]\|_2 \geq \sigma_{\min}([A_1, A_0]) \tag{4.133}$$

if $[A_1 + \Delta A_1, A_0 + \Delta A_0]$ is rank deficient. From (4.132) and (4.133) it follows that

$$\delta_2(\lambda A_1 + A_0) \geq \min \left\{ \sigma_{\min}([A_1, A_0]), \frac{\sqrt{2}}{2} \sigma_{\min} \left( \begin{bmatrix} A_1 \\ A_0 \end{bmatrix} \right) \right\}.$$

As for the upper bound of $\delta_2(\lambda A_1 + A_0)$ in (4.130), the proof follows immediately from Proposition 4.71 and Proposition 4.88. □

Unlike the classical result — Theorem 4.87 in the case of $k = (r-1)$ — on the distance to the nearest rank-deficient matrix for matrices, the nearest distance to singularity for matrix polynomials in the spectral norm may be (strictly) less than the nearest distance to singularity in the Frobenius norm, as the following example demonstrates.

**Example 4.90** We investigate the regular matrix pencil

$$A(\lambda) = \lambda A_1 + A_0 := \lambda \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Applying Theorem 4.82 to $A(\lambda)$, we have

$$\delta_2(A(\lambda)) = \min_{k=1,2} \min_{\substack{Z \in \mathbb{C}^{2\times k}, \\ Z^H Z = I}} \sigma_k([A_1 Z, A_0 Z]). \tag{4.134}$$

And applying Proposition 4.88 to $A(\lambda)$, we have

$$\delta_F(A(\lambda)) = \min \left\{ \sigma_{\min} \left( \begin{bmatrix} A_1 \\ A_0 \end{bmatrix} \right), \sigma_{\min}([A_1, A_0]) \right\} = \min\left\{\sqrt{2}, \sqrt{2}\right\} = \sqrt{2}. \tag{4.135}$$

Let $Z = \frac{\sqrt{2}}{2}[1,1]^T \in \mathbb{C}^2$. Then $Z^H Z = 1$, and by (4.134) we have

$$\begin{aligned}
\delta_2(A(\lambda)) &\le \sigma_1\left([A_1 Z, A_0 Z]\right) \\
&= \sigma_1\left(\left[\frac{\sqrt{2}}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \frac{\sqrt{2}}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right]\right) \\
&= \sigma_1\left(\frac{\sqrt{2}}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\right) = 1. \qquad (4.136)
\end{aligned}$$

Applying Proposition 4.89 to $\delta_2(A(\lambda))$, we have

$$\delta_2(A(\lambda)) \ge \min\left\{\sigma_{\min}\left([A_1, A_0]\right), \frac{\sqrt{2}}{2}\sigma_{\min}\left(\begin{bmatrix} A_1 \\ A_0 \end{bmatrix}\right)\right\} = \min\left\{\sqrt{2}, \frac{\sqrt{2}}{2}\cdot\sqrt{2}\right\} = 1. \quad (4.137)$$

Hence, from (4.136) and (4.137) it follows that $\delta_2(A(\lambda)) = 1$. Therefore, by (4.135), we have

$$\delta_2(A(\lambda)) = 1 < \sqrt{2} = \delta_F(A(\lambda)).$$

$\diamondsuit$

### 4.4.4 Lower Bounds on $\delta_p(A(\lambda))$

Since the nearest distance is analogous to the *stability radius* in control theory, we are here only interested in deriving lower bounds on the nearest distance.

**Lower Bounds Using $\sigma_{\min}\left(\alpha^l A_l + \alpha^{l-1}\beta A_{l-1} + \cdots + \beta^l A_0\right)$**

We note that the determinant of a square singular matrix polynomial is identically equal to zero. It is natural to use this information and make use of Theorem 4.87 to obtain lower bounds on the nearest distance $\delta_p(A(\lambda))$, $p = 2, F$. During the process of deduction, we need Lemma 4.92 (below), which describes a relationship between the 2-norm and *unitarily invariant norms*. For the sake of completeness, we first restate the definition of unitarily invariant norms from Stewart and Sun [55].

**Definition 4.91** *([55] p. 74.) Given a norm $\|\cdot\|$ on $\mathbb{C}^{m\times n}$, it is* UNITARILY INVARIANT *if for any $A \in \mathbb{C}^{m\times n}$, and any unitary $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$, it satisfies*

$$\left\|U^H A V\right\| = \|A\|.$$

**Lemma 4.92** *([55] p. 80. Theorem 3.9.) Let $\|\cdot\|$ be a family of unitarily invariant norms, and let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times q}$, where $m, n, q \in \mathbb{N}$. Then*

$$\|AB\| \le \|A\|\|B\|_2 \quad \text{and} \quad \|AB\| \le \|A\|_2\|B\|.$$

We conduct our analysis as follows. Assume that $\Delta A(\lambda) = \sum_{i=0}^{l} \lambda^i \Delta A_i$ is a minimum $p$-norm de-regularizing perturbation of $A(\lambda)$, $p = 2, F$. Since $A(\lambda) + \Delta A(\lambda)$ is singular, we have, for any $\lambda_0 \in \mathbb{C}$,

$$\det(A(\lambda_0) + \Delta A(\lambda_0)) = \det\left(\sum_{i=0}^{l} \lambda_0^i (A_i + \Delta A_i)\right) = 0. \tag{4.138}$$

Let $\lambda_0 = \alpha_0/\beta_0$, where $\alpha_0, \beta_0 \in \mathbb{C}$, $\beta_0 \neq 0$, and

$$\sum_{i=0}^{l} |\alpha_0^{l-i}\beta_0^i|^2 = |\alpha_0^l|^2 + |\alpha_0^{l-1}\beta_0|^2 + \cdots + |\alpha_0\beta_0^{l-1}|^2 + |\beta_0^l|^2 = 1. \tag{4.139}$$

If we substitute $\beta_0\lambda_0$ for $\alpha_0$ in the normalization equation (4.139), then (4.139) implies that

$$|\beta_0| = \begin{cases} \sqrt[2l]{1/(l+1)} & \text{if } |\lambda_0| = 1 \\ \sqrt[2l]{(1-|\lambda_0|^2)/(1-|\lambda_0|^{2(l+1)})} & \text{if } |\lambda_0| \neq 1. \end{cases} \tag{4.140}$$

Obviously, there exist infinitely many pairs $(\alpha_0, \beta_0)$ such that $\alpha_0/\beta_0 = \lambda_0$ and (4.139) is satisfied, provided that $|\beta_0|$ satisfies (4.140). Substituting $\alpha_0/\beta_0$ for $\lambda_0$ in (4.138), we have

$$\det\left(\sum_{i=0}^{l} \alpha_0^i\beta_0^{l-i}(A_i + \Delta A_i)\right) = \det\left(\sum_{i=0}^{l} \alpha_0^i\beta_0^{l-i}A_i + \sum_{i=0}^{l} \alpha_0^i\beta_0^{l-i}\Delta A_i\right) = 0. \tag{4.141}$$

Thus, by Theorem 4.87, (4.141) implies that

$$\begin{aligned} \sigma_{\min}\left(\sum_{i=0}^{l} \alpha_0^i\beta_0^{l-i}A_i\right) &\leq \left\|\sum_{i=0}^{l} \alpha_0^i\beta_0^{l-i}\Delta A_i\right\|_p \\ &= \left\|[\Delta A_l, \Delta A_{l-1}, \ldots, \Delta A_0]\begin{bmatrix} \alpha_0^l I \\ \alpha_0^{l-1}\beta_0 I \\ \vdots \\ \beta_0^l I \end{bmatrix}\right\|_p \\ &\leq \|[\Delta A_l, \Delta A_{l-1}, \ldots, \Delta A_0]\|_p \left\|\begin{bmatrix} \alpha_0^l I \\ \alpha_0^{l-1}\beta_0 I \\ \vdots \\ \beta_0^l I \end{bmatrix}\right\|_2, \end{aligned} \tag{4.142}$$

where $p = 2, F$, and the inequality (4.142) holds due to Lemma 4.92 and the fact that the 2-norm and Frobenius norm are unitarily invariant norms. Note that, by the normalization equation (4.139), we have

$$\begin{bmatrix} \alpha_0^l I \\ \alpha_0^{l-1}\beta_0 I \\ \vdots \\ \beta_0^l I \end{bmatrix}^H \begin{bmatrix} \alpha_0^l I \\ \alpha_0^{l-1}\beta_0 I \\ \vdots \\ \beta_0^l I \end{bmatrix} = \sum_{i=0}^{l} |\alpha_0^i \beta_0^{l-i}|^2 I = I, \tag{4.143}$$

and therefore,

$$\left\| \begin{bmatrix} \alpha_0^l I \\ \alpha_0^{l-1}\beta_0 I \\ \vdots \\ \beta_0^l I \end{bmatrix} \right\|_2 = 1. \tag{4.144}$$

From (4.142) and (4.144) it follows that

$$\sigma_{\min}\left(\sum_{i=0}^{l} \alpha_0^i \beta_0^{l-i} A_i\right) \leq \|[\Delta A_l, \Delta A_{l-1}, \ldots, \Delta A_0]\|_p = \delta_p(A(\lambda)),\ p = 2, F. \tag{4.145}$$

Note that, if $|\alpha_0| = 1$ and $\beta_0 = 0$, then (4.139) still holds, and, by Corollary 4.37, so does (4.145). Since the inequality in (4.145) holds for any pair $(\alpha_0, \beta_0) \in \mathbb{C} \times \mathbb{C}$, immediately, we have the following proposition on a family of lower bounds on $\delta_p(A(\lambda))$.

**Proposition 4.93** *Let $\mathcal{S} = \{(\alpha, \beta) \in \mathbb{C} \times \mathbb{C} \mid \alpha$ and $\beta$ satisfy (4.139)$\}$, and let $\mathcal{T} \subset \mathcal{S}$ be some test set of pairs $(\alpha, \beta)$. Then,*

$$\begin{aligned} \delta_p(A(\lambda)) &\geq \max_{(\alpha,\beta)\in\mathcal{S}} \sigma_{\min}\left(\sum_{i=0}^{l} \alpha^i \beta^{l-i} A_i\right), && (4.146) \\ &\geq \max_{(\alpha,\beta)\in\mathcal{T}\subset\mathcal{S}} \sigma_{\min}\left(\sum_{i=0}^{l} \alpha^i \beta^{l-i} A_i\right), \quad p = 2, F. && (4.147) \end{aligned}$$

In practical computations, we may let the pair $(1, 0)$ be always included in the test set $\mathcal{T}$, and for other elements of $\mathcal{T}$, we can randomly choose distinct numbers $\lambda_1, \lambda_2, \ldots$, and let $\alpha_i := \lambda_i \beta_i$, $i = 1, 2, \ldots$, where

$$\beta_i = \begin{cases} \sqrt[2l]{1/(l+1)} & \text{if } |\lambda_i| = 1; \\ \sqrt[2l]{(1-|\lambda_i|^2)/(1-|\lambda_i|^{2(l+1)})} & \text{if } |\lambda_0| \neq 1, \end{cases} \quad i = 1, 2, \ldots.$$

It should be noted that, given two pairs $(\alpha_1, \beta_1)$ and $(\alpha_2, \beta_2)$ which satisfy the normalization equation (4.139) and $\alpha_1\beta_2 = \alpha_2\beta_1$, if $|\beta_1| = |\beta_2|$, then

$$\sigma_{\min}\left(\sum_{i=0}^{l} \alpha_1^i \beta_1^{l-i} A_i\right) = \sigma_{\min}\left(\sum_{i=0}^{l} \alpha_2^i \beta_2^{l-i} A_i\right) = \begin{cases} |\beta_1|^l \sigma_{\min}\left(\sum_{i=0}^{l} \left(\frac{\alpha_1}{\beta_1}\right)^i A_i\right) & \text{if } \beta_1 \neq 0; \\ \sigma_{\min}(A_l) & \text{if } \beta_1 = 0. \end{cases}$$

For a regular matrix polynomial $A(\lambda)$, if the number of elements test set $\mathcal{T}$ in Proposition 4.93 is sufficiently large, then the lower bound in (4.147) is *positive* (a zero lower bound is trivial for regular matrix polynomials), as the following remark shows.

**Remark 4.94** Let $\tilde{m} \in \mathbb{Z}$ be defined as in (4.86) in Corollary 4.64. If the matrix polynomial $A(\lambda) = \sum_{i=0}^{l} \lambda^i A_i$ is regular, then, by Theorem 4.62, the number of *nonzero roots* of $\det(A(\lambda))$ is strictly less than $\tilde{m}$. Therefore, if $A(\lambda)$ is regular, and if the test set $\mathcal{T}$ contains the pair $(0,1)$ and greater than or equal to $\tilde{m}$ pairs $(\alpha, \beta)$, $\alpha \neq 0$, which are distinct in the sense that for any two pairs $(\alpha_1, \beta_1)$ and $(\alpha_2, \beta_2)$, $\alpha_1\beta_2 \neq \alpha_2\beta_1$, then the lower bound in (4.147) is positive. $\lozenge$

Byers, He, and Mehrmann show in [4] that in the case of regular matrix pencils, it is not difficult to find an example where the lower bound in (4.146) is attainable. In the case of regular matrix polynomials of higher degrees, it is not either. Take, for instance, $A(\lambda)$ in Example 4.69, in which we assume $0 < |\epsilon| < 1$. On the one hand, if we let $\Delta A(\lambda) = \lambda^2 \Delta A_2 + \lambda \Delta A_1 + \Delta A_0$, where

$$\Delta A_2 = \begin{bmatrix} 0 & -\epsilon \\ 0 & 0 \end{bmatrix}, \Delta A_1 = \Delta A_0 = 0,$$

then $A(\lambda) + \Delta A(\lambda)$ is singular. By Definition 4.70, $\delta_p(A(\lambda)) \leq \| [\Delta A_2, \Delta A_1, \Delta A_0] \|_p = |\epsilon|$, where $p = 2, F$. On the other hand, if we let $(\alpha, \beta) = (1, 0)$ and use the lower bound in (4.147), then we have $\delta_p(A(\lambda)) \geq \sigma_{\min}(A_2) = \sigma_{\min}\left(\begin{bmatrix} 0 & \epsilon \\ 1 & 0 \end{bmatrix}\right) = |\epsilon|$, since $|\epsilon| < 1$. Hence, $\delta_p(A(\lambda)) = |\epsilon|$, and therefore by (4.146), we obtain

$$\max_{(\alpha,\beta)\in\mathcal{S}} \sigma_{\min}\left(\alpha^2 A_2 + \alpha\beta A_1 + \beta^2 A_0\right) = \delta_p(A(\lambda)) = |\epsilon|, \ p = 2, F,$$

where $\mathcal{S} = \{(\alpha, \beta) \in \mathbb{C} \times \mathbb{C} \mid |\alpha|^4 + |\alpha|^2|\beta|^2 + |\beta|^4 = 1\}$. Inequality (4.146) also becomes an equality in the case that the coefficient matrices of $A(\lambda)$ are scalar multiples of one another.

We can also find cases where inequality (4.146) holds strictly in the case of Frobenius norm, or in other words, the lower bound in (4.146) is a genuine *bound* for $p = F$, as the next example demonstrates.

**Example 4.95** We investigate the regular matrix pencil

$$A(\lambda) = \lambda A_1 + A_0 := \lambda \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

By Proposition 4.88, we have

$$\begin{aligned} \delta_F(A(\lambda)) &= \min\left\{\sigma_{\min}\left(\begin{bmatrix} A_1 \\ A_0 \end{bmatrix}\right), \sigma_{\min}\left([A_1, A_0]\right)\right\} \\ &= \min\left\{\sqrt{2}, \sqrt{2}\right\} = \sqrt{2}. \end{aligned}$$

For any $(\alpha, \beta) \in \mathcal{S} = \{(\alpha, \beta) \in \mathbb{C} \times \mathbb{C} \mid |\alpha|^2 + |\beta|^2 = 1\}$, we have

$$\begin{aligned}
\sigma_{\min}(\alpha A_1 + \beta A_0) &= \sigma_{\min}\left(\begin{bmatrix} \beta+\alpha & 0 \\ 0 & \beta-\alpha \end{bmatrix}\right) \\
&= \min\{|\beta+\alpha|, |\beta-\alpha|\}. \qquad (4.148)
\end{aligned}$$

Since $|\beta+\alpha| \cdot |\beta-\alpha| = |\beta^2 - \alpha^2| \leq |\alpha|^2 + |\beta|^2 = 1$, we have

$$\min\{|\beta+\alpha|, |\beta-\alpha|\} \leq \sqrt{|\beta+\alpha| \cdot |\beta-\alpha|} \leq 1. \qquad (4.149)$$

Therefore, by (4.148) and (4.149), we obtain that

$$\max_{(\alpha,\beta)\in\mathcal{S}} \sigma_{\min}(\alpha A_1 + \beta A_0) \leq 1 < \sqrt{2} = \delta_F(A(\lambda)).$$

$\Diamond$

In the case of spectral norm, at this writing, we still do not know of an example in which Inequality (4.146) strictly holds.

**Lower Bounds Using $\sigma_{\min}(W_s(A_l, A_{l-1}, \ldots, A_0))$ and $\sigma_{\min}\left(\widehat{W}_s(A_l, A_{l-1}, \ldots, A_0)\right)$**

Another natural way to obtain lower bounds on $\delta_p(A(\lambda))$ is to make use of the characterization described in Proposition 4.78, ignoring that the perturbation required should have the structure and be constrained as in (4.109) and (4.110). In the course of derivation, we need the following lemma, which relates the eigenvalues of a principal submatrix to the eigenvalues of the original *Hermitian* matrix.

**Lemma 4.96** *([55] p. 198. Corollary 4.3.) Let $A \in \mathbb{C}^{n\times n}$ be a Hermitian matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, and let $B$ be a principal submatrix of order $n-k$ of $A$ with eigenvalues $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_{n-k}$. Then*

$$\lambda_i \geq \mu_i \geq \lambda_{i+k}, \quad i = 1, 2, \ldots, n-k.$$

We carry out our deduction as follows. Assume that $\Delta A(\lambda) = \sum_{i=0}^{l} \lambda^i \Delta A_i$ is a minimum $p$-norm de-regularizing perturbation of $A(\lambda)$, $p = 2, F$. Then, by Proposition 4.78, the block matrix

$$\begin{aligned}
& W_s(A_l + \Delta A_l, A_{l-1} + \Delta A_{l-1}, \ldots, A_0 + \Delta A_0) \\
(= \quad & W_s(A_l, A_{l-1}, \ldots, A_0) + W_s(\Delta A_l, \Delta A_{l-1}, \ldots, \Delta A_0))
\end{aligned}$$

is rank deficient for some $s$, where $1 \leq s \leq (n-1)l + 1$. Thus, if we allow unstructured and unconstrained perturbations of $W_s(A_l, A_{l-1}, \ldots, A_0)$, then, by Theorem 4.87, we have

$$\sigma_{\min}(W_s(A_l, A_{l-1}, \ldots, A_0)) \leq \|W_s(\Delta A_l, \Delta A_{l-1}, \ldots, \Delta A_0)\|_p, \; p = 2, F. \qquad (4.150)$$

Analogously, by Proposition 4.78 and Theorem 4.87, we have

$$\sigma_{\min}\left(\widehat{W}_{\hat{s}}\left(A_l, A_{l-1}, \ldots, A_0\right)\right) \leq \left\|\widehat{W}_{\hat{s}}\left(\Delta A_l, \Delta A_{l-1}, \ldots, \Delta A_0\right)\right\|_p, \; p = 2, F, \tag{4.151}$$

for some $\hat{s}$, $1 \leq \hat{s} \leq (n-1)l + 1$.

To get the relation between $\|W_s\left(\Delta A_l, \Delta A_{l-1}, \ldots, \Delta A_0\right)\|_F$ and $\delta_F(A(\lambda))$, we note that

$$\|W_s\left(\Delta A_l, \Delta A_{l-1}, \ldots, \Delta A_0\right)\|_F = \sqrt{s} \cdot \|[\Delta A_l, \ldots, \Delta A_0]\|_F = \sqrt{s} \cdot \delta_F(A(\lambda)). \tag{4.152}$$

Therefore, from (4.150) and (4.152) it follows that

$$\frac{\sigma_{\min}\left(W_s\left(A_l, A_{l-1}, \ldots, A_0\right)\right)}{\sqrt{s}} \leq \delta_F(A(\lambda)), \text{ for some } s, \tag{4.153}$$

where $1 \leq s \leq ((n-1)l + 1)$. Note that for any $s_1, s_2 \in \mathbb{N}$, if $1 \leq s_1 < s_2 \leq ((n-1)l + 1)$, then $W_{s_1}\left(A_l, A_{l-1}, \ldots, A_0\right)$ is a principal submatrix of $W_{s_2}\left(A_l, A_{l-1}, \ldots, A_0\right)$ (cf. (4.32)), and therefore, $W_{s_1}^H\left(A_l, \ldots, A_0\right) \cdot W_{s_1}\left(A_l, \ldots, A_0\right)$ is a principal submatrix of $W_{s_2}^H\left(A_l, \ldots, A_0\right) \cdot W_{s_2}\left(A_l, \ldots, A_0\right)$. Thus, by Lemma 4.96, we have

$$\sigma_{\min}\left(W_{s_2}\left(A_l, A_{l-1}, \ldots, A_0\right)\right) \leq \sigma_{\min}\left(W_{s_1}\left(A_l, A_{l-1}, \ldots, A_0\right)\right). \tag{4.154}$$

Therefore, by (4.153) and (4.154), we finally have

$$\min_s \frac{\sigma_{\min}\left(W_s\left(A_l, A_{l-1}, \ldots, A_0\right)\right)}{\sqrt{s}} = \frac{\sigma_{\min}\left(W_{\tilde{s}}\left(A_l, A_{l-1}, \ldots, A_0\right)\right)}{\sqrt{\tilde{s}}} \leq \delta_F(A(\lambda)), \tag{4.155}$$

where $1 \leq s \leq ((n-1)l + 1)$, and $\tilde{s} = (n-1)l + 1$.

Along the same lines, we obtain that

$$\min_{\hat{s}} \frac{\sigma_{\min}\left(\widehat{W}_s\left(A_l, A_{l-1}, \ldots, A_0\right)\right)}{\sqrt{\hat{s}}} = \frac{\sigma_{\min}\left(\widehat{W}_{\tilde{s}}\left(A_l, A_{l-1}, \ldots, A_0\right)\right)}{\sqrt{\tilde{s}}} \leq \delta_F(A(\lambda)), \tag{4.156}$$

where $1 \leq \hat{s} \leq ((n-1)l + 1)$, and $\tilde{s} = (n-1)l + 1$.

Combining (4.155) and (4.156), we immediately have the following proposition.

**Proposition 4.97** *Let* $\tilde{s} := (n-1)l + 1$. *Then*

$$\frac{1}{\sqrt{\tilde{s}}} \max\left\{\sigma_{\min}\left(W_{\tilde{s}}\left(A_l, A_{l-1}, \ldots, A_0\right)\right), \sigma_{\min}\left(\widehat{W}_{\tilde{s}}\left(A_l, A_{l-1}, \ldots, A_0\right)\right)\right\} \leq \delta_F(A(\lambda)), \tag{4.157}$$

*where* $W_{\tilde{s}}\left(A_l, A_{l-1}, \ldots, A_0\right)$ *and* $\widehat{W}_{\tilde{s}}\left(A_l, A_{l-1}, \ldots, A_0\right)$ *are defined as in (4.32) and (4.34), respectively.*

Since it appears that we can not get a "$\leq$" relation of $\|W_s(\Delta A_l, \Delta A_{l-1}, \ldots, \Delta A_0)\|_2$, or of $\left\|\widehat{W}_s(\Delta A_l, \Delta A_{l-1}, \ldots, \Delta A_0)\right\|_2$, to $c \cdot \delta_2(A(\lambda))$ (where $c$ is a constant), we do not, at this writing, obtain a reasonable lower bound on $\delta_2(A(\lambda))$ which is similar to that on $\delta_F(A(\lambda))$ in (4.157).

Examples in [4] show that the lower bound in (4.157) is usually coarser than that on $\delta_F(A(\lambda))$ obtained in (4.147), which may be regarded as the cost of our disregarding the structure and constraint of the perturbation of $W_s(A_l, A_{l-1}, \ldots, A_0)$. As a matter of fact, at this writing, we still do not know of an example, except for the simplest cases $n = 1$ (in which $A(\lambda)$ degenerates into a scalar polynomial) or $l = 0$ (in which $A(\lambda)$ degenerates into a constant matrix), in which the lower bound in (4.157) is attainable.

# Chapter 5

# Conclusions and Outlook

In this thesis we have presented the theoretical analysis of two interrelated topics: linear differential-algebraic equations of higher-order and the regularity and singularity of matrix polynomials.

In the first part of this thesis, we have directly investigated the mathematical structures of general (including over- and underdetermind) linear higher-order systems of DAEs with constant and variable coefficients. Making use of the algebraic techniques devised in [28, 29, 34] and taking linear second-order systems of DAEs as examples, we have given condensed forms, under strong equivalence transformations, for triples of matrices and triples of matrix-valued functions which are associated with the systems of constant and variable coefficients respectively. It should be noted that in the case of variable coefficients, we have developed a system of invariant quantities and a set of regularity conditions to ensure that the condensed form can be obtained. Based on the condensed forms, we have converted the systems into ordinary-differential-equation part, 'strange' coupled differential-algebraic-equation part, and algebraic-equation part, and designed the differentiation-and-elimination steps to partially decouple the strange part. Inductively conducting such process of transformation and decoupling, we have, finally, converted the original systems into equivalent strangeness-free systems, from which the solution behaviour with respect to solvability, uniqueness of solutions and consistency of initial conditions can be directly read off.

In addition, we have shown that the strangeness-index of systems of DAEs with constant coefficients is well-defined, and can be determined from the right-hand side of the final strangeness-free system. We have also presented that, in the case of a square constant coefficient system of DAEs of higher-order, given necessary and consistent initial conditions, the initial value problem for the system of DAEs is solvable and has a unique solution for any right-hand side $f(t) \in \mathcal{C}^{\mu}([t_0, t_1], \mathbb{C}^n)$ with the strangeness index $\mu$ of the system if and only if the matrix polynomial associated with the system is regular. Note that if one works with such higher-order problems in the traditional theoretical framework of first-order systems of DAEs, then, to get the solvability and uniqueness of solutions, more smoothness of the right-hand side $f(t)$ may be required, namely, $f(t) \in \mathcal{C}^{\hat{\mu}}([t_0, t_1], \mathbb{C}^n)$, where $\hat{\mu} > \mu$ (cf. for example, [17], Chapter 8 and Subsection S1.3.). Whereas, in the case of variable coefficient systems of DAEs of higher-order, it is expected that sufficient and necessary conditions for the existence of the strangeness-index will be investigated in the future.

On the basis of presented theoretical results on linear systems of DAEs of higher-order, we also expect that numerical methods and software packages for the determination of consistent initial conditions and for the computation of solutions to the associated initial value problems will be developed in the future (for numerical treatment of first-order systems, cf. [31, 32, 33, 34]).

In the second part of this thesis, from the point of view of both the theory of matrices and matrix computations, we have discussed the regularity and singularity of matrix polynomials. Several sufficient and necessary conditions for the column- and row- regularity and singularity of rectangular matrix polynomials have been presented. Such conditions have laid a theoretical foundation for the subsequent related investigations. For instance, we have used them to present a geometrical characterization of singular matrix pencils, by which, conversely, the definition of deflating subspaces of a regular matrix pencil has been clarified. We have also presented a canonical form, under equivalence transformations, for $2 \times 2$ singular quadratic matrix polynomials, which clearly demonstrates the geometrical relations between the row (and column) spaces of the coefficient matrices of a $2 \times 2$ singular quadratic matrix polynomial.

In the case of square matrix polynomials, we have investigated the problems of detecting the regularity and singularity and of nearness to singularity for regular matrix polynomials. We have presented an algorithm to check whether or not a given matrix polynomial is regular via the rank information of its matrix coefficients. As a by-product in our investigation, we have also given attainable lower bounds on the algebraic multiplicity of eigenvalues $\infty$ and 0 of a polynomial eigenvalue problem $\left(\sum_{i=0}^{l} \lambda^i A_i\right) x = 0$ if the corresponding matrix polynomial $\sum_{i=0}^{l} \lambda^i A_i$ is regular.

For square and regular matrix polynomials, we have given a definition of the distance, in terms of the spectral and Frobenius matrix norms, to the nearest singular matrix polynomials. Several basic and interesting properties of the distance have been presented. Based on the sufficient and necessary conditions of the regularity of matrix polynomials obtained, a general theoretical characterization of the nearest distance to singularity has been also presented. From the characterization it turns out that the nearness problem is in essence a perturbation-structured and constrained rank-deficiency problem, for which to determine an explicit computable formula appears to be an open problem. Nonetheless, in the case of matrix pencils we have developed a useful characterization, in terms of the singular values of matrices, of the nearest distance, which directly coincides with the obtained geometrical characterization for singular matrix pencils. We have also investigated the nearness problem for two special cases of matrix polynomials, and in particular, presented an example in which the nearest distance to singularity in terms of the spectral norm is less than that in terms of the Frobenius norm. At last, two types of lower bounds on the nearest distance for general regular matrix polynomials have been presented.

In the future we expect that detecting the regularity and singularity and providing information on the nearness to singularity will be realized in those software packages which deal with systems of linear differential-algebraic equations with constant coefficients and polynomial eigenvalue problems.

# Bibliography


[1] U. M. Ascher, L. R. Petzold. *Projected collocation for higher-order higher-index differential-algebraic equations.* J. Comp. Appl. Math. 43 (1992) 243–259.

[2] K. Balla, R. März. *A unified approach to linear differential algebraic algebraic equations and their adjoints.* Z. Anal. Anwendungen, 21(2002)3, 783-802.

[3] K. E. Brenan, S. L. Campbell, and L. R. Petzold. *Numerical Solutions of Initial-Value Problems in Differential-Algebraic Equations.* Classics in Applied Mathematics, Vol. 14, SIAM, 1996.

[4] R. Byers, C. He, V. Mehrmann. *Where is the nearest non-regular pencil?* Lin. Alg. Appl., 285: 81-105, 1998.

[5] E. A. Coddington, R. Carlson. *Linear Ordinary Differential Equations.* SIAM. Philadelphia. 1997.

[6] R. Courant, F. John. *Introduction to Calculus and Analysis I.* Springer-Verlag, New York, Inc. 1989.

[7] S. L. Campbell. *Singular Systems of Differential Equations.* Pitman, Boston, 1980.

[8] S. L. Campbell. *Singular Systems of Differential Equations II.* Pitman, Boston, 1982.

[9] E. K.-W. Chu. *Perturbation of eigenvalues for matrix polynomials via the Bauer–Fike theorems.* SIAM J. Matrix Anal. Appl. 25(2003), pp. 551-573.

[10] E. A. Coddington, N. Levinson. *Theory of Ordinary Differential Equations.* McGraw-Hill Book Company, Inc. 1955.

[11] C. De Boor, H. O. Kreiss. *On the condition of the linear systems associated with discretized BVPs of ODEs.* SIAM J. Numer. Anal., Vol 23, 1986, pp. 936-939.

[12] J. Demmel. *Applied Numerical Linear Algebra.* SIAM Press. 1997.

[13] J. -P. Dedieu, F, Tisseur. *Perturbation theory for homogeneous polynomial eigenvalue problems*. Lin. Alg. Appl., 358: 71-74, 2003.

[14] F. R. Gantmacher. *The Theory of Matrices*. Vol. I. Chelsea Publishing Company, New York, 1959.

[15] F. R. Gantmacher. *The Theory of Matrices*. Vol. II. Chelsea Publishing Company, New York, 1959.

[16] E. Griepentrog, M. Hanke and R. März. *Toward a better understanding of differential algebraic equations (Introductory survey)*. Seminar Notes Edited by E. Griepentrog, M. Hanke and R. März, Berliner Seminar on Differential-Algebraic Equations, 1992.

http://www.mathematik.hu-berlin.de/publ/SB-92-1/s_dae.html

[17] I. Gohberg, P. Lancaster, L. Rodman. *Matrix Polynomials*. Academic Press. 1982.

[18] E. Griepentrog, R. März. *Differential-Algebraic Equations and Their Numerical Treatment*. Teubner-Texte zur Mathematik, Band 88, 1896.

[19] G. H. Golub, C. F. Van Loan. *Matrix Computations, Third Edition*. The Johns Hopkins University Press. 1996.

[20] N. J. Higham. *Matrix nearness problems and applications*. In M. J. C. Gover and S. Barnett, editors, *Applications of Matrix Theory*, pages 1-27, Oxford University Press, 1989.

[21] D. J. Higham, N. J. Higham. *Structured backward error and condition of generalized eigenvalue problems*. SIAM J. Matrix Anal. Appl. 20 (2) (1998) 493-512.

[22] N. J. Higham, F. Tisseur. *More on pseudospectra for polynomial eigenvalue problems and applications in control theory*. Lin. Alg. Appl., 351-352: 435-453, 2002.

[23] N. J. Higham, F. Tisseur. *Bounds for Eigenvalues of Matrix Polynomials*. Lin. Alg. Appl., 358: 5-22, 2003.

[24] N. J. Higham, F. Tisseur, and P. M. Van Dooren. *Detecting a definite Hermitian pair and a hyperbolic or elliptic quadratic eigenvalue problem, and associated nearness problems*. Lin. Alg. Appl., 351-352: 455-474, 2002.

[25] R. A. Horn, C. R. Johnson. *Matrix Analysis*. Cambridge University Press, July, 1990.

[26] T.-M. Hwang, W.-W. Lin and V. Mehrmann. *Numerical solution of quadratic eigenvalue problems with structure-preserving methods.* SIAM J. Sci. Comput. 24:1283-1302, 2003.

[27] P. Kunkel and V. Mehrmann. *Smooth factorizations of matrix valued functions and their derivatives.* Numer. Math. 60: 115-132, 1991.

[28] P. Kunkel and V. Mehrmann. *Canonical forms for linear differential-algebraic equations with variable coefficients.* J. Comp. Appl. Math. 56(1994), 225-251.

[29] P. Kunkel and V. Mehrmann. *A new look at pencils of matrix valued functions.* Lin. Alg. Appl., 212/213: 215-248, 1994.

[30] P. Kunkel and V. Mehrmann. *Local and global invariants of linear differential-algebraic equations and their relation.* Electron. Trans. Numer. Anal. 4: 138-157, 1996.

[31] P. Kunkel and V. Mehrmann. *A new class of discretization methods for the solution of linear differential-algebraic equations.* SIAM J. Numer. Anal. 33: 1941-1961, 1996.

[32] P. Kunkel, V. Mehrmann, W. Rath, and J. Weickert. *A new software package for the solution of linear differential algebraic equations.* SIAM J. Sci. Comput. 18: 115-138, 1997.

[33] P. Kunkel, V. Mehrmann, and Werner Rath. *Analysis and numerical solution of control problems in descriptor form.* Math. Control Signals Systems. 14: 29-61 (2001)

[34] P. Kunkel and V. Mehrmann. *Analytical and Numerical Treatment of Initial Value Problems for Differential-Algerbaic Equations* (in manuscript).

[35] P. Lancaster. *Lambda-Matrices and Vibrating Systems.* Pergamon Press, Oxford, 1966. xiii+196 pp.

[36] P. Lancaster, M. Tismenetsky. *The Theory of Matrices with Applications.* Second Edition. Academic Press. 1985.

[37] R. März. *The index of linear differential algebraic equations with properly stated leading terms.* Results Math. 42(2002), 308-338.

[38] R. März. *Solvability of linear differential algebraic equations with properly stated leading terms.* Humboldt-Universität zu Berlin, Institut für Mathematik, Preprint 02-12.

[39] The MathWorks, Inc.

http://www.mathworks.com/access/helpdesk/help/techdoc/ref/functionlist.shtml

[40] W. S. Martinson, P. I. Barton. *A Differentiation index for partial differential-algebraic equations.* SIAM J. Sci. Comput. 21 (2000), 2295-2315.

[41] D. S. Mackey, N. Mackey, C. Mehl, V. Mehrmann. *Linearization of parlindromic polynomial eigenvalue problems: good vibrations from good linearizations.* (in manuscript), 2003.

[42] T. Maly, L. R. Petzold. *Numerical methods and software for sensitivity analysis of differential-algebraic systems.* Appl. Numer. Math. 20 (1996), 57-79.

[43] N. H. McClamroch. *Singular systems of differential equations as dynamic models for constrained robot systems.* Technical Report RSD-TR-2-86, Univ. of Michigan Robot Systems Division, 1986.

[44] K. Meerbergen, F. Tisseur. *The Quadratic Eigenvalue Problem.* SIAM Review, (43)2:235-286,2001.

[45] C. B. Moler, G. W. Stewart. *An algorithm for generalized matrix eigenvalue problems.* SIAM J. Numer. Anal., 10: 241-56, 1973.

[46] V. Mehrmann, D. Watkins. *Polynomial eigenvalue problems with Hamiltonian structure.* Electron. Trans. Numer. Anal. Vol. 13. 2002

[47] L. R. Petzold. *Differential-algebraic equations are not ODE's.* SIAM J. Sci. & Statist. Comp. 3(1982), 367-384.

[48] P. J. Rabier, W. C. Rheinboldt. *Nonholonomic motion of rigid mechanical systems from a DAE viewpoint.* SIAM, Philadelphia, PA, 2000. viii+140 pp. ISBN: 0-89871-446-X

[49] P. J. Rabier, W. C. Rheinboldt. *Theoretical and numerical analysis of differential-algebraic equations.* Handbook of numerical analysis, Vol. VIII, 183–540, North-Holland, Amsterdam, 2002.

[50] J. Sand. *On implicit Euler for high-order high-index DAEs.* Appl. Numer. Math. 42 (2002), pp. 411-424.

[51] R. Schüpphaus. *Regelungstechnische Analyse und Synthese von Mehrkörpersystemen in Deskriptorform.* Fortschr.-Ber. VDI Reihe 8 Nr. 478. Düsseldorf: VDI-Verlag 1995.

[52] Leonard M. Silverman. *Inversion of multivariable linear systems.* IEEE Trans. on Automatic Control, Vol. AC-14, No. 3, June 1969.

[53] G. W. Stewart. *On the sensitivity of the eigenvalue problem* $Ax = \lambda Bx$. SIAM J. Numer. Anal., 9: 669-686, 1972.

[54] G. W. Stewart. *Error and perturbation bounds for subspaces associated with certain eigenvalue problems.* SIAM Review, 15(4):727-764, Oct., 1973.

[55] G. W. Stewart, J. Sun. *Matrix Perturbation Theory.* Academic Press, New York, 1991.

[56] F. Tisseur. *Backward error and condition of polynomial eigenvalue problems.* Lin. Alg. Appl., 309: 339-361, 2000.

[57] F. Tisseur, N. J. Higham. *Structured pseudospectra for polynomial eigenvalue problems, with applications.* SIAM J. Matrix Anal. Appl. 23 (1) (2001) 187-208.

[58] P. Van Dooren. *The computation of Kronecker's canonical form of a singular pencil.* Lin. Alg. Appl., 27: 103-141, 1979.

[59] P. Van Dooren, P. Dewilde. *The eigenstructure of an arbitrary polynomial matrix: computational aspects.* Lin. Alg. Appl., 50: 545-579, 1983.

CURRICULUM VITAE

## Personal Details


| | |
|---|---|
| Name | Chunchao Shi |
| Nationality | P. R. China |
| Date of Birth | 28th July, 1974 |
| Place of Birth | Shanxi Province, P. R. China |
| Gender | Male |
| Marital status | Single |
| Current address | Sekretariat MA 4-5, Fakultät II: Institut für Mathematik<br>Technische Universität Berlin, Strasse des 17. Juni 136<br>D-10623 Berlin |
| Telephone No. | (030) 314-28036 |
| E-mail | shi@math.TU-Berlin.DE |


## Education

| | |
|---|---|
| Jul., 1991 | Graduated from No. 1 High School of TangShan,<br>Hebei Province, China |
| Sep., 1991-Jul., 1992 | Military training in Dalian Military Academy,<br>Liaoning Province, China |
| Sep., 1992-Jul., 1996 | Department of Mathmatics,<br>Fudan University, Shanghai, China |
| Jul., 1996 | B.S. majoring in Numerical Mathematics and Its Applied Software |
| Sep., 1996-Jul., 1999 | Institute of Mathematics,<br>Fudan University, Shanghai, China |
| Jul., 1999 | M.S. majoring in Numerical Linear Algebra |
| Sep., 1999-Jun. 2000 | Ph. D. Student in Institute of Mathematics,<br>Fudan University, Shanghai, China |
| Oct. 2000-Mar. 2001 | Ph. D. Student in Department of Mathmatics,<br>Technische Universität Chemnitz, Germany |
| Since Apr. 2001 | Ph. D. Student in Department of Mathmatics,<br>Technische Universität Berlin, Germany. |

Berlin, den 09.05.2004