Analysis and Numerical Solution of Generalized Lyapunov Equations

vorgelegt von

Dipl.-Math. Tatjana Stykel

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. G. M. Ziegler
Berichter: Prof. Dr. V. Mehrmann
Berichter: Priv.-Doz. Dr. P. Benner
Gutachter: Prof. Dr. P. Lancaster

Tag der wissenschaftlichen Aussprache: 13. Juni 2002

Berlin 2002
D 83

An erster Stelle möchte ich mich bei Herrn Prof. Dr. Volker Mehrmann für die Anregung zu dieser Arbeit sowie die wertvollen Diskussionen und Hinweise recht herzlich bedanken.

Mein weiterer Dank gilt Herrn Priv.-Doz. Dr. Peter Benner und Prof. Dr. Peter Lancaster für die Begutachtung der Arbeit.

Bedanken möchte ich mich weiterhin bei den Kolleginnen und Kollegen der Arbeitsgruppen Numerische Mathematik an der TU Chemnitz, wo ich meine Arbeit angefangen habe, und an der TU Berlin, wo sie fortgesetzt wurde, für das sehr freundliche Arbeitsklima. Insbesondere danke ich Andreas Steinbrecher, meinem Büronachbarn, für die Beantwortung meiner zahlreichen Fragen.

Nicht zuletzt bedanke ich mich bei meinen Eltern für ihre ständige Unterstützung.

Eidesstattliche Erklärung

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

Berlin, den 20.03.2002 Tatjana Stykel

Zusammenfassung

Diese Arbeit befaßt sich mit der theoretischen Analyse, numerischen Behandlung und Störungstheorie für verallgemeinerte kontinuierliche und diskrete algebraische Lyapunov-Gleichungen. Die Stabilität von singulären Systemen und dazugehörige Eigenwertprobleme werden auch untersucht. Spektralcharakteristiken werden vorgestellt, die die Lage der endlichen Eigenwerte des Matrixbüschels bezüglich der imaginären Achse und des Einheitskreises charakterisieren. Diese Charakteristiken lassen sich zur Schätzung des asymptotischen Verhaltens der Lösungen von singulären Systemen verwenden.

Bei der Lösung von verallgemeinerten Lyapunov-Gleichungen treten einige Schwierigkeiten insbesondere dann auf, wenn eine der Koeffizientenmatrizen singulär ist. In diesem Fall werden verallgemeinerte Lyapunov-Gleichungen mit der speziellen rechten Seite untersucht. Für solche Gleichungen lassen sich die klassischen Stabilitätssätze von Lyapunov nur für Büschel des Indexes höchstens zwei im zeitkontinuierlichen Fall und des Indexes höchstens eins im zeitdiskreten Fall verallgemeinern.

Weiterhin werden projizierte verallgemeinerte kontinuierliche und diskrete Lyapunov-Gleichungen betrachtet, die durch gewisse Projektion der rechten Seite und der Lösung auf die rechten und linken invarianten Unterräume zu den endlichen Eigenwerten des Matrixbüschels entstehen. Für diese Gleichungen werden notwendige und hinreichende Bedingungen der eindeutigen Lösbarkeit vorgestellt, die vom Index des Matrixbüschels unabhängig sind. Es wird gezeigt, dass die projizierten Lyapunov-Gleichungen verwendet werden können um die asymptotische Stabilität der singulären Systeme sowie Steuerbarkeits- und Beobachtbarkeitseigenschaften der Deskriptorsysteme zu charakterisieren. Außerdem sind diese Gleichungen nützlich, die Trägheitssätze für Matrizen auf Matrixbüschel zu erweitern. Schließlich wird gezeigt, dass die Gramschen Matrizen der Steuerbarkeit und Beobachtbarkeit für Deskriptorsysteme als die Lösungen der projizierten Lyapunov-Gleichungen bestimmt werden können.

Die numerische Lösung von verallgemeinerten Lyapunov-Gleichungen wird betrachtet. Die Erweiterungen des Bartels-Stewart-Verfahrens und des Hammarling-Verfahrens auf projizierte Lyapunov-Gleichungen werden vorgestellt. Diese Verfahren basieren auf die Berechnung der GUPTRI-Form des Matrixbüschels.

Die Störungstheorie für verallgemeinerte Lyapunov-Gleichungen wird entwickelt. Es werden die auf Spektralnorm basierenden Konditionszahlen für projizierte verallgemeinerte Lyapunov-Gleichungen eingeführt, die zu Störungsabschätzungen der Lösungen dieser Gleichungen verwendet werden können. Darüber hinaus wird gezeigt, dass diese Konditionszahlen mit den erwähnten Spektralcharakteristiken für die asymptotische Stabilität von singulären Systemen übereinstimmen und sich durch die Lösung von projizierten Lyapunov-Gleichungen mit der Einheitsmatrix in der rechten Seite effizient berechnen lassen.

Die Anwendung der projizierten verallgemeinerten Lyapunov-Gleichungen in der Modellreduktion von Deskriptorsystemen wird ebenso betrachtet. Für Deskriptorsysteme werden die Hankel-Singulärwerte eingeführt und Verallgemeinerungen der Balanced Truncation Verfahren dargestellt.

Notation

$\mathbb{R}$	the field of the real numbers
$\mathbb{R}^- = (\,-\infty,\,0\,)$	the negative real semi-axis
$i = \sqrt{-1}$	the imaginary unit
$\Re e(z)$	the real part of $z \in \mathbb{C}$
$\mathbb{C}$	the field of the complex numbers
$\mathbb{C}^- = \{\, z \in \mathbb{C} \,:\, \Re e(z) < 0 \,\}$	the open left half-plane
$\mathbb{F}^{n,m}$	the space of real ($\mathbb{F} = \mathbb{R}$) or complex ($\mathbb{F} = \mathbb{C}$) matrices of size $n \times m$
$A = [a_{kj}]_{k,j=1}^{n,m}$	a matrix $A \in \mathbb{F}^{n,m}$ with elements a_{kj} in position (k, j)
A^*	the transpose ($A^* = A^T$) of real A or the complex conjugate transpose ($A^* = A^H$) of complex A
A^{-1}	the inverse of A
$A^{-*} = (A^{-1})^*$	the inverse, complex conjugate and transpose of A
$\mathrm{diag}(A_1, \ldots, A_k)$	a block diagonal matrix with $A_j \in \mathbb{F}^{n_j,n_j}$, $j = 1, \ldots, k$
$I = I_n = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$	the identity matrix of order n
$N_n = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}$	a nilpotent matrix of order n in Jordan form
$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & & \vdots \\ a_{n1}B & \cdots & a_{nm}B \end{bmatrix}$	the Kronecker product of matrices $A \in \mathbb{F}^{n,m}$ and $B \in \mathbb{F}^{n,m}$
$\mathrm{vec}(A) = (a_{11}, \ldots, a_{n1}, a_{12}, \ldots, a_{nm})^T$	the vector formed by stacking the columns of $A \in \mathbb{F}^{n,m}$
Π_{n^2}	the vec-permutation matrix of size $n^2 \times n^2$ such that $\mathrm{vec}(A^T) = \Pi_{n^2}\mathrm{vec}(A)$
$\det(A)$	the determinant of $A \in \mathbb{F}^{n,n}$
$\mathrm{rank}(A)$	the rank of $A \in \mathbb{F}^{n,m}$
$\mathrm{trace}(A) = \sum_{j=1}^{n} a_{jj}$	the trace of $A \in \mathbb{F}^{n,n}$
$\mathrm{Ker}\, A = \{\, x \in \mathbb{F}^m \;:\; Ax = 0 \,\}$	the right null space (or kernel) of $A \in \mathbb{F}^{n,m}$
$\mathrm{Im}\, A = \{\, y \in \mathbb{F}^n \;:\; y = Ax,\;\; x \in \mathbb{F}^m \,\}$	the range (or image) of $A \in \mathbb{F}^{n,m}$
$\mathrm{Sp}(A) = \{\, \lambda \in \mathbb{C} \;:\; \det(A - \lambda I) = 0 \,\}$	the set of eigenvalues or the spectrum of $A \in \mathbb{F}^{n,n}$
$\lambda_j(A)$, $\lambda_j(E, A)$	eigenvalues of the matrix A and the pencil $\lambda E - A$

$\sigma_1(A) \geq \ldots \geq \sigma_k(A) \geq 0$	singular values of $A \in \mathbb{F}^{n,m}$, $k = \min(n,m)$
$\sigma_{\min}(A) = \sigma_k(A)$	the smallest singular value of $A \in \mathbb{F}^{n,m}$
$\sigma_{\max}(A) = \sigma_1(A)$	the largest singular value of $A \in \mathbb{F}^{n,m}$
$\langle x, y \rangle = y^* x = \sum_{j=1}^{n} x_j \bar{y}_j$	the inner product in $\mathbb{F}^n$
$\|x\| = \langle x, x \rangle^{1/2} = \left(\sum_{j=1}^{n} \|x_j\|^2 \right)^{1/2}$	the Euclidean vector norm of $x \in \mathbb{F}^n$
$\langle A, B \rangle = \mathrm{trace}(B^* A)$	the inner product in $\mathbb{F}^{n,m}$
$\|A\|_F = \langle A, A \rangle^{1/2} = \left(\sum_{j=1}^{m} \sum_{k=1}^{n} \|a_{kj}\|^2 \right)^{1/2}$	the Frobenius matrix norm of $A \in \mathbb{F}^{n,m}$
$\|A\|_2 = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sigma_{\max}(A)$	the spectral matrix norm of $A \in \mathbb{F}^{n,m}$

Contents

Chapter 1

Introduction

We study the generalized continuous-time algebraic Lyapunov equation (GCALE)

$$E^*XA + A^*XE = -G \tag{1.1}$$

and the generalized discrete-time algebraic Lyapunov equation (GDALE)

$$A^*XA - E^*XE = -G, \tag{1.2}$$

where E, A, G are given matrices and X is an unknown matrix. They are named after the Russian mathematician Alexander Mikhailovitch Lyapunov, who in his doctoral dissertation "The general problem of the stability of motion" in 1892, see [111], presented the stability theory for linear and nonlinear systems. He has shown that the asymptotic behavior of solutions of linear differential equations is closely related to continuous-time Lyapunov matrix equations.

Lyapunov equations arise not only in the stability analysis of differential and difference equations but also in many other applications such as system and control theory [51, 98, 117, 119, 148, 176], eigenvalue problems [62, 99, 113, 116] and partial differential equations [142].

For $E = I$, equations (1.1) and (1.2) are the standard continuous-time and discrete-time Lyapunov equations. In the last century the theory and numerous numerical algorithms were developed for such equations, see [9, 20, 51, 53, 72, 80, 99, 100, 126, 127] and the references therein. The case of nonsingular E has been considered in [17, 34, 101, 117, 125]. However, only little attention has been paid to generalized Lyapunov equations with a singular matrix E [105, 116, 123, 146, 151, 153, 175].

It is known that the GCALE (1.1) has a unique solution for every G if the matrix E is nonsingular and all the eigenvalues of the pencil $\lambda E - A$ have negative real part. The GDALE (1.2) is uniquely solvable for every G if the matrix E is nonsingular and all the eigenvalues of $\lambda E - A$ have modulus smaller than one. However, if E is singular, then the GCALE (1.1) may have no solutions even if all the finite eigenvalues of $\lambda E - A$ lie in the open half-plane and a solution, if it exists, is not unique. Analogous trouble arises in the GDALE (1.2) when both the matrices E and A are singular. Such an equation may have no solutions even if all the finite eigenvalues of the pencil $\lambda E - A$ lie inside the unit circle. Moreover, if the GDALE (1.2) with singular E and A is solvable, the solution is not unique.

To overcome these difficulties various types of generalized Lyapunov equations have been proposed in the literature [11, 105, 116, 153, 154]. Unfortunately, these equations are mostly limited to the case of pencils of index at most one. In this thesis we consider the *projected generalized continuous-time algebraic Lyapunov equation*

$$\begin{aligned} &E^*XA + A^*XE = -P_r^*GP_r, \\ &X = XP_l \end{aligned} \tag{1.3}$$

and the *projected generalized discrete-time algebraic Lyapunov equation*

$$\begin{aligned} &A^*XA - E^*XE = -P_r^*GP_r + \xi(I-P_r)^*G(I-P_r), \\ &P_l^*X = XP_l, \end{aligned} \tag{1.4}$$

with $\xi = -1, 0, 1$. Here P_l and P_r are the spectral projections onto the left and right deflating subspaces of the pencil $\lambda E - A$ corresponding to the finite eigenvalues. For such equations, existence and uniqueness theorems can be stated independently of the index of the pencil $\lambda E - A$. We also discuss applications of equations (1.3) and (1.4) to the study of the asymptotic behavior of solutions of singular systems, the distribution of the generalized eigenvalues of a pencil in the complex plane with respect to the imaginary axis and the unit circle, as well as controllability and observability properties for descriptor systems.

The classical numerical methods for the standard Lyapunov equations ($E = I$) are the Bartels-Stewart method [9], the Hammarling method [72] and the Hessenberg-Schur method [65]. An extension of these methods to generalized Lyapunov equations with nonsingular matrix E was given in [34, 55, 56, 65, 117, 125]. These methods are based on the preliminary reduction of the matrix (matrix pencil) to the (generalized) Schur form [64] or the Hessenberg-Schur form [65], calculation of the solution of the reduced system and back transformation. In this thesis we present a generalization of the Bartels-Stewart and Hammarling methods for the projected generalized Lyapunov equations (1.3) and (1.4).

In numerical problems it is very important to study the sensitivity of the solution to perturbations in the input data and to estimate errors in the computed solution. There are several papers concerned with the perturbation theory and the backward error bounds for standard continuous-time Lyapunov equations, see [61, 74, 75] and references therein. The sensitivity analysis for generalized Lyapunov equations has been presented in [96], where only the case of nonsingular E was considered. In this thesis we discuss the perturbation theory for the projected Lyapunov equations (1.3) and (1.4).

Model reduction is of fundamental importance in modeling and control applications. Often simulation or controller design for large dynamical systems arising from electrical networks and partial differential equations becomes difficult because of storage limits and expensive computations. To overcome these difficulties one can employ model order reduction that consists in an approximation of the dynamical system by a reduced order system. It is required that the approximate system preserve properties of the original system like stability and passivity and it has a small approximation error. Moreover, the computation of the reduced order system should be numerically stable and efficient.

For standard state space systems various model reduction techniques have been proposed such as balanced truncation [102, 119, 129, 137, 156, 164], singular perturbation approximation [94, 107], optimal Hankel norm approximation [58] and Padé approximation [47, 52, 68]. Unfortunately, there is no general approach that can be considered as optimal. Surveys on system approximation and model reduction can be found in [2, 4, 48, 121].

Model reduction of descriptor systems based on the Padé approximation via the Lanczos process has been developed in [47, 52]. Drawbacks of this technique are that there is no approximation error bound for the reduced order system and stability is not necessary preserved. The balanced truncation approach [102, 119, 137, 156, 164] related to the controllability and observability Gramians is free from these disadvantages. Balanced truncation methods for state space systems are based on transforming the dynamical system to a balanced form such that the controllability and observability Gramians become diagonal and equal together with truncation of states that are both difficult to reach and to observe. In this thesis we extend these methods to descriptor systems.

The thesis is organized as follows. Chapter 2 contains some background material that we need in the following. Section 2.1 summarizes some necessary definitions and theorems from matrix analysis. In Section 2.2 we introduce functions of matrix pencils and study some of their properties.

Chapter 3 is devoted to linear continuous-time and discrete-time descriptor systems. In Section 3.1 solvability and stability analysis for continuous-time descriptor systems is presented, while discrete-time descriptor systems are discussed in Section 3.2. We introduce numerical parameters that characterize the property of a pencil $\lambda E - A$ to have all finite eigenvalues in the open left half-plane in the continuous-time case and inside the unit circle in the discrete-time case. In Section 3.3 the different concepts of controllability and observability for descriptor systems are reviewed and equivalent algebraic and geometric characterizations are given.

In Chapter 4 we consider generalized Lyapunov equations. Section 4.1 contains some applications for Lyapunov equations. In Section 4.2 we study the existence and uniqueness of solutions for generalized continuous-time Lyapunov equations with general and special right-hand sides. Special attention will be paid to the projected GCALE (1.3). We also present generalized inertia theorems that give a connection between the signature of the solution of (1.3) and the numbers of eigenvalues of the pencil $\lambda E - A$ in the left and right open half-plane and on the imaginary axis. In Section 4.3 we discuss analogous results for generalized discrete-time Lyapunov equations. Similar to the continuous-time case, we establish a relationship between the signature of the solution of equation (1.4) and the number of eigenvalues of the pencil $\lambda E - A$ inside, outside and on the unit circle. Section 4.4 contains a generalization of the controllability and observability Gramians for descriptor systems that are closely related to the projected generalized Lyapunov equations.

Chapter 5 is concerned with the numerical solution of generalized Lyapunov equations. In Sections 5.1 and 5.1 we describe a generalized Schur-Bartels-Stewart method and a generalized Schur-Hammarling method that can be used to solve the projected generalized

Lyapunov equations (1.3) and (1.4). Numerical aspects and complexity of these methods are presented in Section 5.3. Iterative methods for (generalized) Lyapunov equations are discussed in Section 5.4.

Chapter 6 contains the perturbation theory for generalized Lyapunov equations. In Section 6.1 we review condition numbers and Frobenius norm based condition estimators for deflating subspaces of matrix pencils corresponding to finite eigenvalues. Section 6.2 presents the known sensitivity results for the generalized Lyapunov equations (1.1) and (1.2) with nonsingular E. In Section 6.3 we define a spectral norm based condition number for the projected GCALE (1.3) which can be efficiently computed by solving (1.3) with $G = I$. Using this condition number we derive the perturbation bound for the solution of the projected GCALE (1.3) under perturbations that preserve the deflating subspaces of the pencil $\lambda E - A$ corresponding to the infinite eigenvalues. In Section 6.4 we present the sensitivity analysis for the projected GDALE (1.4) with $\xi = 1$. Section 6.5 contains some results of numerical experiments.

Chapter 7 deals with model reduction for descriptor systems. In Section 7.1 we review some properties of the transfer function and its realizations for descriptor systems. In Section 7.2 we generalize Hankel singular values and study some of their features. Balancing of descriptor systems is treated in Section 7.3. In Section 7.4 we propose an extension of the balanced truncation technique for descriptor systems that leads in a natural way to generalized model reduction algorithms presented in Section 7.5. Section 7.6 contains numerical examples.

In Chapter 8 we give some conclusions. We also point out several open problems that will be investigated in the future.

Chapter 2

Definitions and basic properties

In this chapter we give necessary definitions and present some theorems from matrix analysis that will be used in the sequel. More details can be found in [53, 64, 78, 99, 145].

2.1 Matrices and matrix pencils

A matrix $A \in \mathbb{F}^{n,n}$ is *Hermitian* (*symmetric* for $A \in \mathbb{R}^{n,n}$) if $A = A^*$. The matrix $A \in \mathbb{F}^{n,n}$ is called *positive* (*negative*) *definite on a subspace* $\mathcal{X} \subset \mathbb{F}^n$ if $v^*Av > 0$ ($v^*Av < 0$) for all nonzero $v \in \mathcal{X}$. The matrix $A \in \mathbb{F}^{n,n}$ is called *positive* (*negative*) *definite* and *positive* (*negative*) *semidefinite* if $v^*Av > 0$ ($v^*Av < 0$) and $v^*Av \geq 0$ ($v^*Av \leq 0$), respectively, for all nonzero $v \in \mathbb{F}^n$.

The following matrix decompositions present useful tools in numerical analysis [64, 99, 144].
QR decomposition. Let $A \in \mathbb{F}^{n,n}$. There exist a unitary matrix $Q \in \mathbb{F}^{n,n}$ and an upper triangular matrix $R \in \mathbb{F}^{n,n}$ such that $A = QR$.
Cholesky decomposition. An Hermitian, positive (semi)definite matrix $A \in \mathbb{F}^{n,n}$ can be represented as $A = U_A^*U_A$, where $U_A \in \mathbb{F}^{n,n}$ is an upper triangular *Cholesky factor* of A.
Full rank decomposition. Let $A \in \mathbb{F}^{n,n}$ be an Hermitian, positive semidefinite matrix and $r = \operatorname{rank}(A)$. Then there exists a matrix $R_A \in \mathbb{F}^{r,n}$ of full row rank such that $A = R_A^*R_A$. The matrix R_A is the *full row rank factor* and R_A^* is the *full column rank factor* of A.
Singular value decomposition. Let $A \in \mathbb{F}^{n,m}$ and $r = \operatorname{rank}(A)$. There exist unitary matrices $U \in \mathbb{F}^{n,n}$ and $V \in \mathbb{F}^{m,m}$ such that

$$A = U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V^*,$$

where $\Sigma = \operatorname{diag}\Big(\sigma_1(A), \ldots, \sigma_r(A)\Big)$ is a diagonal matrix with positive, decreasing diagonal elements

$$\sigma_1(A) \geq \sigma_2(A) \geq \ldots \geq \sigma_r(A) > 0$$

that are called the (nonzero) *singular values* of A.

Spectral decomposition. Let $A \in \mathbb{F}^{n,n}$ be Hermitian. Then there exists a unitary matrix $U \in \mathbb{F}^{n,n}$ such that

$$A = U\Lambda U^*,$$

where $\Lambda = \operatorname{diag}\Big(\lambda_1(A), \ldots, \lambda_n(A)\Big)$, and $\lambda_j(A)$ are the *eigenvalues* of A.

Numerically stable algorithms for computing these decompositions can be found in [64, 144, 171].

A matrix pencil $\alpha E - \beta A$ is called *regular* if E and A are square, and $\det(\alpha E - \beta A) \neq 0$ for some $(\alpha, \beta) \in \mathbb{C}^2$. Otherwise, the matrix pencil $\alpha E - \beta A$ is called *singular*. A pair $(\alpha, \beta) \in \mathbb{C}^2 \setminus \{(0,0)\}$ is said to be a *generalized eigenvalue* of $\alpha E - \beta A$ if $\det(\alpha E - \beta A) = 0$. If $\beta \neq 0$, then the pair (α, β) represents a *finite eigenvalue* $\lambda = \alpha/\beta$ of the pencil $\lambda E - A$. The pair $(\alpha, 0)$ represents an *infinite eigenvalue* of $\lambda E - A$. Clearly, the pencil $\lambda E - A$ has an eigenvalue at infinity if and only if the matrix E is singular. The set of all generalized eigenvalues (finite and infinite) of the pencil $\lambda E - A$ is called the *spectrum* of $\lambda E - A$ and denoted by $\mathrm{Sp}(E, A)$.

Vectors $x_1, \ldots, x_k$ form a *right Jordan chain* of the pencil $\lambda E - A$ corresponding to an eigenvalue λ if

$$(\lambda E - A)x_1 = 0, \quad (\lambda E - A)x_2 = -Ex_1, \qquad \ldots \qquad (\lambda E - A)x_k = -Ex_{k-1}. \tag{2.1}$$

Vectors $y_1, \ldots, y_k$ form a *left Jordan chain* of $\lambda E - A$ corresponding to an eigenvalue λ if

$$y_1^*(\lambda E - A) = 0, \quad y_2^*(\lambda E - A) = -y_1^* E, \qquad \ldots \qquad y_k^*(\lambda E - A) = -y_{k-1}^* E.$$

The vectors x_1 and y_1 are called, respectively, *right and left eigenvectors* of the pencil $\lambda E - A$ corresponding to λ.

A subspace $\mathcal{V}_\lambda \subset \mathbb{F}^n$ that is the span of all right (left) Jordan chains corresponding to an eigenvalue λ is called *right* (*left*) *deflating subspace* of $\lambda E - A$ corresponding to λ. Deflating subspaces are a natural generalization of invariant subspaces for the standard eigenproblem $\lambda I - A$ to the generalized eigenproblem $\lambda E - A$.

Let $\mathbf{\Lambda} = \{\lambda_1, \ldots, \lambda_p\}$ be a subset of the spectrum of the pencil $\lambda E - A$, where λ_j are pairwise distinct and let $\mathcal{V}_{\lambda_j}$ be the right (left) deflating subspace of $\lambda E - A$ corresponding to λ_j for $j = 1, \ldots, p$. Then the subspace

$$\mathcal{V}_{\mathbf{\Lambda}} = \mathcal{V}_{\lambda_1} \dot{+} \ldots \dot{+} \mathcal{V}_{\lambda_p}$$

is the *right* (*left*) *deflating subspace of* $\lambda E - A$ *corresponding to* $\mathbf{\Lambda}$. Here $\dot{+}$ denotes the direct sum. Moreover, $\mathbb{F}^n$ admits a decomposition $\mathbb{F}^n = \mathcal{V}_{\mathbf{\Lambda}} \dot{+} \mathcal{V}$, where $\mathcal{V}$ is the right (left) complementary deflating subspace of $\lambda E - A$ corresponding to $\mathrm{Sp}(E, A) \setminus \mathbf{\Lambda}$. A projection P onto the deflating subspace $\mathcal{V}_{\mathbf{\Lambda}}$ along the deflating subspace $\mathcal{V}$ is called the *spectral projection onto* $\mathcal{V}_{\mathbf{\Lambda}}$.

A regular pencil $\lambda E - A$ can be represented in the Weierstrass canonical form that is a special case of the Kronecker canonical form [53, 145]. There exist nonsingular matrices

W and T such that

$$E = W \begin{bmatrix} I_{n_f} & 0 \\ 0 & N \end{bmatrix} T \qquad \text{and} \qquad A = W \begin{bmatrix} J & 0 \\ 0 & I_{n_\infty} \end{bmatrix} T. \tag{2.2}$$

The block J corresponds to the finite eigenvalues and has the form

$$J = \operatorname{diag}(J_{1,1}, J_{1,2}, \ldots, J_{1,m_1}, J_{2,1}, \ldots, J_{2,m_2}, \ldots, J_{k,1}, \ldots, J_{k,m_k}),$$

where

$$J_{j,q} = \begin{bmatrix} \lambda_j & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_j \end{bmatrix}$$

is the Jordan block of order $n_{j,q}$ with $\sum_{j=1}^{k} \sum_{q=1}^{m_j} n_{j,q} = n_f$ and λ_j is a finite eigenvalue of the pencil $\lambda E - A$. The number m_j is called the *geometric multiplicity* of λ_j, the number $a_j = \sum_{q=1}^{m_j} n_{j,q}$ is called the *algebraic multiplicity* of λ_j and n_f is the dimension of the left and right deflating subspaces of $\lambda E - A$ corresponding to the finite eigenvalues. A finite eigenvalue is *simple* if it has the same algebraic and geometric multiplicity. The block N in (2.2) corresponds to the eigenvalue at infinity of the pencil $\lambda E - A$ and has the form $N = \operatorname{diag}(N_{n_1}, \ldots, N_{n_t})$, where

$$N_{n_j} = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}$$

is a nilpotent Jordan block of order n_j. The number $n_\infty = \sum_{j=1}^{t} n_j$ is the *algebraic multiplicity* of the eigenvalue at infinity of $\lambda E - A$ and defines the dimension of the right and left deflating subspaces of $\lambda E - A$ corresponding to the eigenvalue at infinity. The size of the largest nilpotent block, denoted by ν, is called the *index* of the pencil $\lambda E - A$. Clearly, $N^{\nu-1} \neq 0$ and $N^{\nu} = 0$. If the matrix E is nonsingular, then $\lambda E - A$ is of index zero. The pencil $\lambda E - A$ is of index one if and only if it has exactly $n_f = \operatorname{rank}(E)$ finite eigenvalues. The following theorem gives another equivalent characterizations for $\lambda E - A$ to have index at most one.

Theorem 2.1. [91] *The following statements are equivalent.*

1. *The pencil $\lambda E - A$ is regular and of index at most one.*
2. $\operatorname{rank} \begin{bmatrix} E \\ K_{E^*}^* A \end{bmatrix} = \operatorname{rank}\,[\, E,\ AK_E \,] = n$, *where K_E and K_{E^*} are matrices with orthogonal columns spanning the right and left null spaces of E, respectively.*

3. The matrix $K_{E^*}^* A K_E$ *is nonsingular.*

4. $\operatorname{Im} E \dot{+} A \operatorname{Ker} E = \mathbb{F}^n$.

Representation (2.2) defines the decomposition of $\mathbb{F}^n$ into two complementary deflating subspaces of the matrix pencil $\lambda E - A$ corresponding to the finite and infinite eigenvalues. The matrices

$$P_l = W \begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} W^{-1} \qquad \text{and} \qquad P_r = T^{-1} \begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} T \tag{2.3}$$

are spectral projections onto the left and right deflating subspaces of $\lambda E - A$ corresponding to the finite eigenvalues. For simplicity, the deflating subspace of $\lambda E - A$ corresponding to the finite (infinite) eigenvalues we will call the *finite* (*infinite*) *deflating subspace*.

It is well known that computing the Weierstrass canonical form in finite precision arithmetic is, in general, an ill-conditioned problem in the sense that small changes in the data may extremely change the canonical form. Therefore, the Weierstrass canonical form is only of theoretical interest. From a computational point of view, the Generalized UPper TRIangular (GUPTRI) form [41, 42] is more suitable. For a regular pencil $\lambda E - A$ with $E,\, A \in \mathbb{R}^{n,n}$, there exist orthogonal matrices V and U such that

$$E = V \begin{bmatrix} E_f & E_u \\ 0 & E_\infty \end{bmatrix} U^T \qquad \text{and} \qquad A = V \begin{bmatrix} A_f & A_u \\ 0 & A_\infty \end{bmatrix} U^T, \tag{2.4}$$

where the pencil $\lambda E_f - A_f$ is quasi-triangular and has only finite eigenvalues, while the pencil $\lambda E_\infty - A_\infty$ is triangular and all its eigenvalues are infinite. Clearly, the matrices E_f and A_∞ are nonsingular, and E_∞ is nilpotent. The GUPTRI form is a special case of the generalized Schur form for regular pencils [64, 145] and can also be extended to singular pencils [41, 42]. The numerical computation of the GUPTRI form and the generalized Schur form of a matrix pencil has been intensively studied and various methods have been proposed, see [10, 41, 42, 64, 169] and the references therein. A comparison of the different algorithms can be found in [41].

2.2 Generalized resolvent and functions of matrix pencils

Let $\lambda E - A$ be a regular matrix pencil. Consider a *generalized resolvent* $(\lambda E - A)^{-1}$ which is a rational matrix-valued function of a complex variable λ defined on $\mathbb{C} \setminus \operatorname{Sp}(E, A)$. At an eigenvalue $\lambda_j(E, A)$ (finite or infinite) of algebraic multiplicity a_j the generalized resolvent has a pole of order a_j. For any $\lambda,\, \mu \notin \operatorname{Sp}(E, A)$, the *generalized resolvent equation*

$$(\lambda E - A)^{-1} - (\mu E - A)^{-1} = (\mu - \lambda)(\lambda E - A)^{-1} E (\mu E - A)^{-1} \tag{2.5}$$

holds.

The generalized resolvent $(\lambda E - A)^{-1}$ has the following Laurent expansion at infinity

$$(\lambda E - A)^{-1} = \lambda^{-1} \sum_{k=-\infty}^{\infty} F_k \lambda^{-k}, \tag{2.6}$$

where the coefficients F_k have the form

$$F_k = \begin{cases} T^{-1} \begin{bmatrix} J^k & 0 \\ 0 & 0 \end{bmatrix} W^{-1}, & k = 0, 1, 2 \ldots, \\ T^{-1} \begin{bmatrix} 0 & 0 \\ 0 & -N^{-k-1} \end{bmatrix} W^{-1}, & k = -1, -2, \ldots \end{cases} \tag{2.7}$$

with W, T, J and N as in (2.2), see [103]. Note that $F_k = 0$ for all $k < -\nu$, where ν is the index of the pencil $\lambda E - A$. The following theorem gives some useful properties of the matrices F_k.

Theorem 2.2. *Let the matrices F_k be as in* (2.7) *and let the projections P_r and P_l be as in* (2.3). *Then*

$$F_j E F_k = F_k E F_j = F_j A F_k = F_k A F_j = 0 \quad \text{for } j < 0,\ k \geq 0, \tag{2.8}$$

$$F_j E F_k = F_k E F_j = \begin{cases} F_{j+k}, & j, k \geq 0, \\ -F_{j+k}, & j, k < 0, \end{cases} \tag{2.9}$$

$$F_j A F_k = F_k A F_j = \begin{cases} F_{j+k+1}, & j, k \geq 0, \\ -F_{j+k+1}, & j, k < 0, \end{cases}$$

$$E F_k A = A F_k E \quad \text{for all } k,$$

$$F_0 E = P_r, \qquad -F_{-1} A = I - P_r, \tag{2.10}$$

$$E F_0 = P_l, \qquad -A F_{-1} = I - P_l.$$

Moreover,

$$\begin{aligned} E F_k &= A F_{k-1} + \delta_{0,k} I, \\ F_k E &= F_{k-1} A + \delta_{0,k} I, \end{aligned} \tag{2.11}$$

where $\delta_{j,k}$ is the Kronecker delta.

Proof. See [11, 113]. □

Similarly to the matrix case [99], we may define a *function of a matrix pencil* [39, 63, 149] as follows.

Definition 2.3. Let $\lambda E - A$ be a regular pencil. Let Γ be a closed Jordan curve such that the finite spectrum of $\lambda E - A$ lies inside Γ. If f is a function that is analytic inside Γ and continuous on Γ, then the function $f(E, A)$ of the pencil $\lambda E - A$ is defined via

$$f(E, A) = \frac{1}{2\pi i} \oint_{\Gamma} f(\lambda)(\lambda E - A)^{-1}\, d\lambda. \tag{2.12}$$

Equation (2.12) is a matrix pencil version of Cauchy's integral formula [135]. From complex function theory [135] it follows that the integral (2.12) does not depend on the particular choice of the curve Γ. For $E = I$, we have that $f(I, A) = f(A)$ is a classical *function of the matrix* A [99]. If the matrix E is nonsingular, then

$$f(E, A) = f(E^{-1}A)E^{-1} = E^{-1}f(AE^{-1}).$$

Remark 2.4. Note that $f(E, A)$ is a matrix but not a matrix pencil.

Example 2.5. Since the exponential function $e^{\lambda t}$ of the complex variable λ is analytic everywhere on $\mathbb{C}$, we may define the *exponential function of the pencil* $\lambda E - A$ via

$$\exp(t, E, A) = \frac{1}{2\pi i}\oint_{\Gamma} e^{\lambda t}(\lambda E - A)^{-1}\, d\lambda, \tag{2.13}$$

where Γ is a closed Jordan curve that encloses the finite spectrum of $\lambda E - A$. This function is a generalization of the matrix exponential e^{At} [99].

Some familiar properties of scalar functions and functions of matrices [63, 99] can be extended to matrix pencils.

Lemma 2.6. *Let* $\mathbf{\Lambda}$ *be a subset of the finite spectrum of a regular pencil* $\lambda E - A$ *and let* $\Gamma_{\mathbf{\Lambda}}$ *be a closed Jordan curve enclosing* $\mathbf{\Lambda}$. *Then the matrices*

$$P_{l,\mathbf{\Lambda}} = \frac{1}{2\pi i}\oint_{\Gamma_{\mathbf{\Lambda}}} E(\lambda E - A)^{-1}\, d\lambda \tag{2.14}$$

and

$$P_{r,\mathbf{\Lambda}} = \frac{1}{2\pi i}\oint_{\Gamma_{\mathbf{\Lambda}}} (\lambda E - A)^{-1}E\, d\lambda \tag{2.15}$$

are spectral projections (*known as Riesz projections*) *onto the left and right deflating subspaces of the pencil* $\lambda E - A$ *corresponding to* $\mathbf{\Lambda}$.

Proof. See [63, Theorem IV.1.1]. □

Lemma 2.7 (Generalized Hamilton-Cayley theorem). *Let* $\chi(\lambda) = \det(\lambda E - A)$ *be the characteristic polynomial of a regular pencil* $\lambda E - A$. *Then* $\chi(E, A) = 0$.

Proof. Let Γ be a closed Jordan curve enclosing the finite spectrum of $\lambda E - A$. Then the function $\chi(\lambda)(\lambda E - A)^{-1}$ is analytic everywhere on $\mathbb{C}$ and, hence, by Cauchy's theorem [135] we have

$$\chi(E, A) = \frac{1}{2\pi i}\oint_{\Gamma} \chi(\lambda)(\lambda E - A)^{-1}\, d\lambda = 0.$$

□

Lemma 2.8. *Let $\lambda E - A$ be a regular pencil and let Γ be a closed Jordan curve such that all finite eigenvalues of $\lambda E - A$ lie inside Γ. Assume that functions f and g are continuous on Γ and analytic inside Γ. Then*

$$\begin{aligned} &(i) \quad (f+g)(E,A) = f(E,A) + g(E,A), && (2.16)\\ &(ii) \quad (af)(E,A) = af(E,A) \qquad \text{for all} \quad a \in \mathbb{C}, && (2.17)\\ &(iii) \quad (fg)(E,A) = f(E,A)\,E\,g(E,A) = g(E,A)\,E\,f(E,A). && (2.18) \end{aligned}$$

Proof. Clearly, the functions $f+g$, af and fg are continuous on the curve Γ and analytic inside Γ. Equations (2.16) and (2.17) are obvious. To prove (2.18), see [39, Lemma 1]. □

Lemma 2.9. *Let Γ be a closed Jordan curve enclosing the finite spectrum of a regular pencil $\lambda E - A$ and let the matrices F_k be as in (2.7). Then*

$$F_k = \frac{1}{2\pi i}\oint_\Gamma \lambda^k(\lambda E - A)^{-1}\,d\lambda \qquad \text{for} \quad k \geq 0. \tag{2.19}$$

Moreover, if the origin is inside Γ, then

$$F_k = -\frac{1}{2\pi i}\oint_\Gamma \lambda^k(\lambda E - A)^{-1}\,d\lambda \qquad \text{for} \quad k < 0. \tag{2.20}$$

Proof. Using the Weierstrass canonical form (2.2) of the pencil $\lambda E - A$ we obtain

$$\frac{1}{2\pi i}\oint_\Gamma \lambda^k(\lambda E - A)^{-1}\,d\lambda = W^{-1}\left(\frac{1}{2\pi i}\oint_\Gamma \begin{bmatrix} \lambda^k(\lambda I - J)^{-1} & 0 \\ 0 & \lambda^k(\lambda N - I)^{-1} \end{bmatrix} d\lambda\right) T^{-1}.$$

Since all eigenvalues of J lie inside the curve Γ and N is nilpotent, we have

$$\frac{1}{2\pi i}\oint_\Gamma \lambda^k(\lambda I - J)^{-1}d\lambda = J^k, \qquad \frac{1}{2\pi i}\oint_\Gamma \lambda^k(\lambda N - I)^{-1}d\lambda = 0 \qquad \text{for} \quad k \geq 0.$$

Furthermore, if the origin is inside Γ, then

$$-\frac{1}{2\pi i}\oint_\Gamma \lambda^k(\lambda I - J)^{-1}d\lambda = 0, \qquad -\frac{1}{2\pi i}\oint_\Gamma \lambda^k(\lambda N - I)^{-1}d\lambda = -N^{k-1} \qquad \text{for} \quad k < 0.$$

Thus, (2.19) and (2.20) hold. □

Corollary 2.10. *Let $\lambda E - A$ be a regular pencil and let F_k be as in (2.7). Consider a polynomial $p(\lambda) = a_0 + a_1\lambda + \ldots + a_m\lambda^m$. Then $p(E,A) = a_0F_0 + a_1F_1 + \ldots + a_mF_m$.*

Proof. The result follows from (2.12), (2.16) and (2.19). □

Chapter 3

Linear descriptor systems

Consider a linear time-invariant continuous-time system

$$\begin{array}{rcl} E\dot{x}(t) & = & Ax(t) + Bu(t), \quad x(0) = x^0, \\ y(t) & = & Cx(t), \end{array} \tag{3.1}$$

and a linear time-invariant discrete-time system

$$\begin{array}{rcl} Ex_{k+1} & = & Ax_k + Bu_k, \quad x_0 = x^0, \\ y_k & = & Cx_k, \end{array} \tag{3.2}$$

where E, $A \in \mathbb{F}^{n,n}$, $B \in \mathbb{F}^{n,m}$, $C \in \mathbb{F}^{p,n}$, $x(t)$, $x_k \in \mathbb{F}^n$ are *state vectors*, $u(t)$, $u_k \in \mathbb{F}^m$ are *control inputs*, $y(t)$, $y_k \in \mathbb{F}^p$ are *outputs* and $x^0 \in \mathbb{F}^n$ is an *initial value*.

If $E = I_n$, then systems (3.1) and (3.2) are called *standard state space systems*. Such systems have been extensively studied, see, e.g., [88, 93, 176] and the references therein. Systems (3.1) and (3.2) with singular E are known in the literature as *descriptor systems* [103, 117, 174], *singular systems* [30, 31, 35, 36], *differential-algebraic equations* [21, 132], *generalized state space systems* [83] and *implicit linear systems* [6, 105]. These equations arise in many different applications such as electrical circuits [21, 30, 31, 69, 70], multibody systems [45, 132, 138], chemical engineering [21, 97], (semi)discretization of partial differential equations [19, 21, 170], economic systems [110] and others.

In this chapter we present some basic concepts of control theory for continuous-time and discrete-time descriptor systems (3.1) and (3.2). We consider existence and uniqueness of solutions of these systems as well as the stability theory. Various types of controllability and observability for descriptor systems are defined and equivalent algebraic and geometric characterizations are given.

3.1 Continuous-time descriptor systems

3.1.1 Solvability and the fundamental solution matrix

In this subsection we review some of the results [30, 36] on the existence and uniqueness of solutions of the continuous-time descriptor system (3.1).

Let $\lambda E - A$ be a regular pencil in Weierstrass canonical form (2.2) and let the matrices

$$W^{-1}B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \qquad \text{and} \qquad CT^{-1} = [\, C_1, \;\; C_2] \tag{3.3}$$

be partitioned in blocks conformally to E and A. Under the coordinate transformation

$$\begin{bmatrix} z(t) \\ w(t) \end{bmatrix} = Tx(t),$$

system (3.1) is decoupled in the *slow* system

$$\dot{z}(t) = Jz(t) + B_1 u(t), \quad z(0) = z^0, \tag{3.4}$$

and the *fast* system

$$N\dot{w}(t) = w(t) + B_2 u(t), \quad w(0) = w^0, \tag{3.5}$$

with $y(t) = C_1 z(t) + C_2 w(t)$ and $\begin{bmatrix} z^0 \\ w^0 \end{bmatrix} = Tx^0$. Systems (3.4) and (3.5) are called also *dynamic* and *algebraic parts* of (3.1), respectively.

Equation (3.4) has a unique solution for any input $u(t)$ and initial value $z^0 \in \mathbb{F}^{n_f}$. This solution has the form

$$z(t) = e^{tJ} z^0 + \int_0^t e^{(t-\tau)J} B_1 u(\tau)\, d\tau.$$

A unique solution of equation (3.5) is given by

$$w(t) = -\sum_{k=0}^{\nu-1} N^k B_2 u^{(k)}(t), \tag{3.6}$$

where ν is the index of the pencil $\lambda E - A$. We see from (3.6) that for the existence of a classical smooth solution $x(t)$, it is necessary that the input $u(t)$ is sufficiently smooth. Moreover, (3.6) shows that not for all initial conditions $x(0) = x^0$ system (3.1) is solvable. The initial value x^0 has to be *consistent*, that is, it must belong to the *set of consistent initial conditions* given by

$$\mathcal{X}_c^0 = \left\{ T^{-1} \begin{bmatrix} z^0 \\ w^0 \end{bmatrix} \quad : \quad z^0 \in \mathbb{F}^{n_f}, \;\; w^0 = -\sum_{k=0}^{\nu-1} N^k B_2 u^{(k)}(0) \right\}. \tag{3.7}$$

Thus, if the pencil $\lambda E - A$ is regular, $x^0 \in \mathcal{X}_c^0$ and $u(t)$ is ν times continuously differentiable, then system (3.1) has a unique, continuously differentiable solution $x(t)$ [30, 36]. We will often denote the solution of (3.1) by $x(t, x^0, u)$ to show explicitly the dependence on the initial value x^0 and the input $u(t)$.

Similarly to the standard case ($E = I$), e.g., [61], we can define a fundamental solution matrix for the descriptor system (3.1).

Definition 3.1. A matrix-valued function $\mathcal{F}(t)$ defined for all $t \in \mathbb{R}$ is called *fundamental solution matrix* of the continuous-time descriptor system (3.1) if it is continuously differentiable and satisfies the initial value problem

$$\begin{aligned} E\dot{\mathcal{F}}(t) &= A\mathcal{F}(t), \\ E\mathcal{F}(0) &= P_l, \end{aligned} \tag{3.8}$$

where P_l is the projection onto the left finite deflating subspace of the pencil $\lambda E - A$.

It should be noted that the introduced fundamental solution matrix $\mathcal{F}(t)$ differs by a left multiple factor E from the fundamental solution matrix considered in [67, 147].

The following theorem discusses existence and uniqueness of $\mathcal{F}(t)$.

Theorem 3.2. *Let $\lambda E - A$ be a regular pencil. Then there exists a unique fundamental solution matrix $\mathcal{F}(t)$ of system* (3.1) *that has the form*

$$\mathcal{F}(t) = \frac{1}{2\pi i}\oint_\Gamma e^{\lambda t}(\lambda E - A)^{-1} d\lambda, \tag{3.9}$$

where Γ is a closed Jordan curve enclosing the finite eigenvalues of the pencil $\lambda E - A$.

Proof. Consider the exponential function $\exp(t) = \exp(t, E, A)$ as in (2.13). Substituting this function in (3.8), we obtain

$$E\frac{d}{dt}\exp(t) - A\exp(t) = \frac{1}{2\pi i}\oint_\Gamma e^{\lambda t}(\lambda E - A)(\lambda E - A)^{-1}\, d\lambda = \frac{1}{2\pi i} I \oint_\Gamma e^{\lambda t} d\lambda = 0.$$

Moreover, it follows from (2.13) and (2.14) that

$$E\exp(0) = \frac{1}{2\pi i}\oint_\Gamma E(\lambda E - A)^{-1}\, d\lambda = P_l.$$

Thus, the fundamental solution matrix of (3.1) exists and is given by (3.9).

In order to prove the uniqueness of the fundamental solution matrix, we consider the homogeneous initial value problem

$$E\dot{\mathcal{F}}(t) = A\mathcal{F}(t), \qquad E\mathcal{F}(0) = 0. \tag{3.10}$$

Using the Weierstrass canonical form (2.2) for the regular pencil $\lambda E - A$ we obtain that (3.10) has only the trivial solution $\mathcal{F}(t) \equiv 0$. Let us now suppose that there exist two fundamental solution matrices $\mathcal{F}_1(t)$ and $\mathcal{F}_2(t)$. Then their difference $\mathcal{F}(t) = \mathcal{F}_1(t) - \mathcal{F}_2(t)$ satisfying (3.10) is identically equal to zero, i.e., $\mathcal{F}_1(t) = \mathcal{F}_2(t)$. □

It follows from Lemmas 2.6 and 2.8 that

$$\begin{aligned} &\mathcal{F}(t)P_l = \mathcal{F}(t) = P_r\mathcal{F}(t), \\ &\mathcal{F}(t)EP_r = \mathcal{F}(t)E = P_r\mathcal{F}(t)E, \\ &P_lE\mathcal{F}(t) = E\mathcal{F}(t) = E\mathcal{F}(t)P_l. \end{aligned}$$

Taking into account (2.2), we can rewrite the fundamental solution matrix $\mathcal{F}(t)$ in (3.9) as

$$\mathcal{F}(t) = T^{-1} \begin{bmatrix} e^{tJ} & 0 \\ 0 & 0 \end{bmatrix} W^{-1}. \tag{3.11}$$

Moreover, if the pencil $\lambda E - A$ has no finite eigenvalues on the imaginary axis, then $\mathcal{F}(t)$ has the following integral representations

$$\mathcal{F}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} P_r (i\omega E - A)^{-1}\, d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} (i\omega E - A)^{-1} P_l \, d\omega. \tag{3.12}$$

These immediately follow from (3.11) and the identity

$$e^{tJ} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} (i\omega I - J)^{-1} d\omega,$$

see, e.g., [61].

Remark 3.3. The fundamental solution matrix $\mathcal{F}(t)$ is closely related to the exponential relation introduced in [13, 14]. For a real matrix pencil $\lambda E - A$, a *linear relation* $(E\backslash A)$ is defined via

$$\big(E\backslash A\big) = \left\{\ (x, v) \in \mathbb{R}^n \times \mathbb{R}^n \quad : \quad Ev = Ax\ \right\}.$$

In terms of linear relations, the continuous-time singular system

$$E\dot{x}(t) = Ax(t) \tag{3.13}$$

can be rewritten as $(x(t), \dot{x}(t)) \in \big(E\backslash A\big)$. Moreover, $x(t)$ is the solution of system (3.13) if and only if

$$(x(t_0), x(t)) \in \exp\big(E\backslash(A(t - t_0))\big),$$

where $t_0 \in \mathbb{R}$ and

$$\exp\big(E\backslash(A(t - t_0))\big) = \sum_{k=0}^{\infty} \frac{(t - t_0)^k}{k!} \big(E\backslash A\big)^k$$

is the *exponential relation*, see [14] for details. On the other hand, the solution of (3.13) has the form $x(t) = \mathcal{F}(t - t_0)Ex(t_0)$ or, equivalently, $(x(t_0), x(t)) \in \big(I\backslash(\mathcal{F}(t - t_0)E)\big)$. Thus, we obtain that $\exp\big(E\backslash(A(t - t_0))\big) = \big(I\backslash(\mathcal{F}(t - t_0)E)\big)$.

Using the fundamental solution matrix $\mathcal{F}(t)$ and the matrices F_k as in (2.7), the classical solution $x(t, x^0, u)$ of the descriptor system (3.1) can be written as

$$x(t, x^0, u) = T^{-1} \begin{bmatrix} z(t) \\ w(t) \end{bmatrix} = \mathcal{F}(t)Ex^0 + \int_0^t \mathcal{F}(t - \tau)Bu(\tau)\, d\tau + \sum_{k=0}^{\nu-1} F_{-k-1} B u^{(k)}(t).$$

If the initial condition x^0 is inconsistent or the input $u(t)$ is not sufficiently smooth (for example, in most control problems u(t) is only piecewise continuous), then the solution of the continuous-time descriptor system (3.1) may have impulsive modes [35, 36]. Such a solution exists in the distributional sense and has the form

$$\begin{aligned} x(t, x^0, u) &= \mathcal{F}(t)Ex^0 + \int_0^t \mathcal{F}(t-\tau)Bu(\tau)\,d\tau \\ &\quad + \sum_{k=1}^{\nu-1} \delta^{(k-1)}(t)F_{-k}Ex^0 + \sum_{k=0}^{\nu-1} F_{-k-1}Bu^{(k)}(t), \end{aligned} \tag{3.14}$$

where $\delta(t)$ is the *Dirac delta function*, $\delta^{(k)}(t)$ and $u^{(k)}(t)$ are distributional derivatives [43]. It follows from (3.14) that system (3.1) has no impulsive solutions for every piecewise continuous input $u(t)$ if and only if $x^0 \in \operatorname{Ker} E$ and $F_{-k-1}B = 0$ for $k > 0$. Moreover, impulsive solutions in (3.1) do not arise if the pencil $\lambda E - A$ is of index at most one.

3.1.2 Stability

In this subsection we discuss the asymptotic behavior of solutions of the descriptor system (3.1) with $u(t) \equiv 0$. There exist various types of stability for ordinary differential equations such as exponential stability, Lyapunov stability, asymptotic stability, uniform stability, internal and external stability, see [61, 71, 88, 111].

The following definitions describe Lyapunov stability for the continuous-time singular system (3.13).

Definition 3.4. The trivial solution $x(t) \equiv 0$ of (3.13) is *stable in the sense of Lyapunov* or *Lyapunov stable* if

(i) for all $x^0 \in \mathbb{F}^n$ the initial value problem

$$\begin{aligned} E\dot{x}(t) - Ax(t) &= 0, \\ P_r\,(x(0) - x^0) &= 0 \end{aligned} \tag{3.15}$$

has a solution $x(t, x^0) \in \operatorname{Im} P_r$ defined on $[0, \infty)$;

(ii) for every $\varepsilon > 0$ there exists a $\delta = \delta(\varepsilon) > 0$ such that $\|x(t, x^0)\| < \varepsilon$ for all $t \geq 0$ and for all $x^0 \in \mathbb{F}^n$ with $\|P_r x^0\| < \delta$.

Definition 3.5. The trivial solution $x(t) \equiv 0$ of (3.13) is *asymptotically stable* if it is Lyapunov stable and if there is a $\delta_0 > 0$ such that for the solution $x(t, x^0)$ of (3.15) with $\|P_r x^0\| < \delta_0$ we have that $x(t, x^0) \to 0$ as $t \to \infty$.

Remark 3.6. Note that the Lyapunov stability does not depend on the special choice of the projection P_r which can be replaced by any matrix M with the property that $\operatorname{Ker} M = \operatorname{Ker} P_r$. This fact is an immediate consequence of the relations $MP_r = M$ and $P_r = P_r M^+ M$, where the matrix M^+ denotes the Moore-Penrose inverse of M, see [32].

The following theorem is well known and gives a necessary and sufficient condition for the trivial solution of (3.13) to be asymptotically stable, see [36, 67, 123].

Theorem 3.7. *Let $\lambda E - A$ be a regular pencil. The trivial solution $x(t) \equiv 0$ of equation* (3.13) *is asymptotically stable if and only if all the finite eigenvalues of $\lambda E - A$ lie in the open left half-plane.*

We now consider the problem to determine via a numerical method whether all the finite eigenvalues of a regular pencil $\lambda E - A$ lie in the open left half-plane. This problem arises also in the study of the asymptotic properties of stationary solutions of autonomous quasilinear and nonlinear differential-algebraic equations [114, 155] and nonautonomous differential-algebraic equations with constant linear part and small nonlinearity [115].

Definition 3.8. A matrix pencil $\lambda E - A$ is called *c-stable* if it is regular and all the finite eigenvalues of $\lambda E - A$ lie in the open left half-plane.

It is known that the generalized eigenvalue problem as well as the standard eigenvalue problem may be ill-conditioned in the sense that eigenvalues may change strongly even under small perturbations in E and A [145, 171]. Consider the following example.

Example 3.9. Let $E = I_{20}$ and

$$A_\varepsilon = \begin{bmatrix} -1 & 10 & & \\ & \ddots & \ddots & \\ & & \ddots & 10 \\ \varepsilon & & & -1 \end{bmatrix}.$$

All eigenvalues of A_0 are -1 and lie in the open left half-plane. However, if $\varepsilon = 10^{-18}$, then the matrix A_ε has an eigenvalue $\lambda = \sqrt[20]{10} - 1$ in the right half-plane.

Recently the concept of ε-pseudospectra and spectral portraits [60, 157] was developed to better understand the influence of perturbations on the spectrum of matrices and matrix pencils, see also [62, 76, 158, 159] and references therein. The application of the ε-pseudospectra in the study of the asymptotic stability of differential equations arising in computational fluid dynamics can be found in [49, 160, 162].

Another possible approach to investigate the asymptotic behavior of solutions of linear ordinary differential equations without explicitly computing the eigenvalues is the consideration of so-called dichotomy parameters that characterize numerically the property of matrices to have all eigenvalues in the open left half-plane and that are efficiently computable [22, 23, 59, 61]. Analogous parameters were introduced in [147, 149] for equation (3.13).

Consider a matrix

$$H_c = \int_0^\infty \mathcal{F}^*(t)\mathcal{F}(t)\, dt, \tag{3.16}$$

where $\mathcal{F}(t)$ is the fundamental solution matrix as in (3.11). If the pencil $\lambda E - A$ is c-stable, that is, $\Re e(\lambda_j(J)) \le -\zeta < 0$, then the estimate

$$\|e^{tJ}\|_2 \le c(n_f)\Big(\frac{\|J\|_2}{\zeta}\Big)^{n_f-1} e^{-t\zeta/2}$$

holds [61]. Here $c(n_f)$ is a constant that depends on n_f only. Then from (3.11) we have the estimate

$$\|\mathcal{F}(t)\|_2 \le \|T^{-1}\|_2\|W^{-1}\|_2\|e^{tJ}\|_2 \le c(n_f)\|T^{-1}\|_2\|W^{-1}\|_2\Big(\frac{\|J\|_2}{\zeta}\Big)^{n_f-1}e^{-t\zeta/2}, \qquad (3.17)$$

and, hence, the integral (3.16) is convergent. The matrix H_c can be used to compute the maximum $\mathbb{L}_2$-norm of the solution $x(t,x^0)=\mathcal{F}(t)Ex^0$ of the initial value problem (3.15). We have

$$\|E^*H_cE\|_2 = \max_{\|v\|=1}\int_0^\infty \|\mathcal{F}(t)Ev\|^2dt = \max_{\|P_rx^0\|=1}\int_0^\infty \|x(t,x^0)\|^2dt.$$

We introduce

$$\kappa_c(E,A) = 2\|\big(EP_r + A(I-P_r)\big)^{-1}A\|_2\|E^*H_cE\|_2, \qquad (3.18)$$

where P_r is as in (2.3). It follows from (3.17) that if the pencil $\lambda E - A$ is c-stable, then $\kappa_c(E,A)$ is bounded. We set $\kappa_c(E,A)=\infty$ if $\lambda E - A$ has at least one finite eigenvalue with nonnegative real part.

It is interesting that the parameter $\kappa_c(E,A)$ can be used for pointwise estimation of the solution of problem (3.15). We will develop a similar technique as in [61].

Theorem 3.10. *Let $x(t,x^0)$ be a solution of the initial value problem* (3.15). *Then*

$$\|x(t,x^0)\| \le \sqrt{\kappa_c(E,A)}\,e^{-t\|(EP_r+A(I-P_r))^{-1}A\|_2/\kappa_c(E,A)}\|P_rx^0\|. \qquad (3.19)$$

Proof. If $\kappa_c(E,A)=\infty$ then inequality (3.19) is fulfilled. Assume that $\kappa_c(E,A)<\infty$ and consider for $t\ge 0$ the matrix-valued function

$$Y(t) = \int_t^\infty \mathcal{F}^*(\tau)\mathcal{F}(\tau)ds.$$

It follows from Lemma 2.8 with $f(\lambda)=e^{t\lambda}$ and $g(\lambda)=e^{\tau\lambda}$ that

$$\mathcal{F}(t+\tau) = \mathcal{F}(t)E\mathcal{F}(\tau) = \mathcal{F}(\tau)E\mathcal{F}(t).$$

Then

$$Y(t) = \int_t^\infty \mathcal{F}^*(\tau)\mathcal{F}(\tau)d\tau = \mathcal{F}^*(t)E^*\left(\int_0^\infty \mathcal{F}^*(\tau)\mathcal{F}(\tau)d\tau\right)E\mathcal{F}(t) = \mathcal{F}^*(t)E^*H_cE\mathcal{F}(t).$$

Differentiating the matrix $Y(t)$, we obtain

$$\frac{d}{dt}Y(t) = -\mathcal{F}^*(t)\mathcal{F}(t).$$

For an arbitrary vector $v\in\mathbb{F}^n$ we have the estimate

$$\frac{d}{dt}\langle Y(t)v,v\rangle = -\langle \mathcal{F}(t)v,\mathcal{F}(t)v\rangle \le -\frac{\langle E^*H_cE\mathcal{F}(t)v,\mathcal{F}(t)v\rangle}{\|E^*H_cE\|_2} = -\frac{\langle Y(t)v,v\rangle}{\|E^*H_cE\|_2},$$

which implies that

$$\frac{d}{dt}\Big(e^{t/\|E^*H_cE\|_2}\langle Y(t)v, v\rangle\Big) \le 0,$$

and, consequently,

$$\begin{aligned}\langle \mathcal{F}^*(t)E^*H_cE\mathcal{F}(t)v, v\rangle &= \langle Y(t)v, v\rangle \le e^{-t/\|E^*H_cE\|_2}\langle Y(0)v, v\rangle \\ &= e^{-t/\|E^*H_cE\|_2}\langle H_cP_lv, P_lv\rangle \\ &= e^{-t/\|E^*H_cE\|_2}\langle H_cv, v\rangle.\end{aligned} \tag{3.20}$$

Furthermore, it is not difficult to verify that

$$\mathcal{F}(t)E = e^{t(EP_r+A(I-P_r))^{-1}A}P_r.$$

Then, taking into account that $\|e^{t(EP_r+A(I-P_r))^{-1}A}P_rv\| \ge e^{-|t|\|(EP_r+A(I-P_r))^{-1}A\|_2}\|P_rv\|$, see [61, p. 24], we have

$$\begin{aligned}\langle E^*H_cEv, v\rangle &= \int_0^\infty \|\mathcal{F}(t)Ev\|^2dt \ge \|P_rv\|^2\int_0^\infty e^{-2t\|(EP_r+A(I-P_r))^{-1}A\|_2}dt \\ &= \frac{\|P_rv\|^2}{2\|(EP_r+A(I-P_r))^{-1}A\|_2}.\end{aligned} \tag{3.21}$$

Substituting in (3.21) the vector $v = \mathcal{F}(t)Ex^0$ we obtain that

$$\|x(t,x^0)\|^2 = \|\mathcal{F}(t)Ex^0\|^2 \le 2\|\big(EP_r + A(I-P_r)\big)^{-1}A\|_2\langle E^*H_cE\mathcal{F}(t)Ex^0, \mathcal{F}(t)Ex^0\rangle.$$

Finally, from (3.20) with $v = Ex^0$ we have

$$\|x(t,x^0)\|^2 \le \kappa_c(E,A)\, e^{-2t\|(EP_r+A(I-P_r))^{-1}A\|_2/\kappa_c(E,A)}\|P_rx^0\|^2.$$

□

The following example shows that the estimate (3.19) is sharp.

Example 3.11. Consider the system

$$E_\varepsilon\dot{x}(t) = A_\varepsilon x(t) \tag{3.22}$$

with

$$E_\varepsilon = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad A_\varepsilon = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -\varepsilon & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

For $0 < \varepsilon < 1$, the general solution of (3.22) is $x(t,x^0) = e^{-t}P_rx^0$ and, hence, the trivial solution of (3.22) is asymptotically stable. We have $\kappa_c(E_\varepsilon, A_\varepsilon) = 1$ and from (3.19) it follows that $\|x(t,x^0)\| \le e^{-t}\|P_rx^0\|$. However, for $\varepsilon = 0$ the pencil $\lambda E_\varepsilon - A_\varepsilon$ is singular, i.e., under a perturbation of norm ε the trivial solution of (3.22) is not asymptotically stable.

From Theorem 3.10 we obtain some useful consequences.

Corollary 3.12. *Let $\kappa_c(E,A)$ be as in* (3.18). *The trivial solution of equation* (3.13) *is asymptotically stable if and only if $\kappa_c(E,A)$ is bounded.*

Proof. If $\kappa_c(E,A)$ is bounded, then by (3.19) the trivial solution of (3.13) is asymptotically stable. On the other hand, by Theorem 3.7 it follows from the asymptotic stability of (3.13) that $\kappa_c(E,A) < \infty$. □

Corollary 3.13. *Let $\mathcal{F}(t)$ be a fundamental solution matrix of* (3.1). *Then*

$$\|\mathcal{F}(t)E\|_2 \leq \sqrt{\kappa_c(E,A)}\, e^{-t\|(EP_r+A(I-P_r))^{-1}A\|_2/\kappa_c(E,A)}. \tag{3.23}$$

Proof. The result follows from the proof of Theorem 3.10. □

Corollary 3.14. *Let P_r be the spectral projection onto the right finite deflating subspace of a regular pencil $\lambda E - A$. Then*

$$\|P_r\|_2 \leq \sqrt{\kappa_c(E,A)}. \tag{3.24}$$

Proof. Since $P_r = \mathcal{F}(0)E$, bound (3.24) immediately follows from (3.23). □

From (3.19) it is also possible to derive a weaker but more practical bound for the solution $x(t,x^0)$ of (3.15). Indeed, from $\|E^*H_cE\|_2 \leq \|E\|_2^2\|H_c\|_2$ and (3.19) we obtain the estimate

$$\begin{aligned}\|x(t,x^0)\| &\leq \sqrt{2\|E\|_2^2\|H_c\|_2\|(EP_r+A(I-P_r))^{-1}\|_2\|A\|_2}\, e^{-t/(2\|E\|_2^2\|H_c\|_2)}\|P_rx^0\|\\ &= \sqrt{\kappa_{c,2}(E,A)\|E\|_2\|(EP_r+A(I-P_r))^{-1}\|_2}\, e^{-t\|A\|_2/(\|E\|_2\kappa_{c,2}(E,A))}\|P_rx^0\|,\end{aligned} \tag{3.25}$$

where $\kappa_{c,2}(E,A) = 2\|E\|_2\|A\|_2\|H_c\|_2$.

Despite of the fact that bound (3.25) may overestimate the solution $x(t,x^0)$ of (3.15), the parameter $\kappa_{c,2}(E,A)$ also characterizes the behavior of $x(t,x^0)$ at infinity. Moreover, $\kappa_{c,2}(E,A)$, in contrast to $\kappa_c(E,A)$, may be more useful to evaluate the "quality" of the asymptotic stability. We see in Example 3.11, that $\kappa_{c,2}(E_\varepsilon,A_\varepsilon) = \varepsilon^{-2} \to \infty$ as $\varepsilon \to 0$ and, hence, (3.22) approaches to an unstable system.

Note that for $E = I$ both parameters $\kappa_c(E,A)$ and $\kappa_{c,2}(E,A)$ coincide with the parameter æ(A) introduced in [22, 61] to study the asymptotic stability of linear ordinary differential equations.

To compute the parameters $\kappa_c(E,A)$ and $\kappa_{c,2}(E,A)$ we need the matrix H_c. The numerical computation of this matrix will be discussed in Section 6.3.

3.2 Discrete-time descriptor systems

In this section we study the discrete-time descriptor system (3.2).

3.2.1 Solvability

Let a regular pencil $\lambda E - A$ be in Weierstrass canonical form (2.2) and let the matrices B and C be as in (3.3). Then (3.2) is equivalent to the decoupled system of equations

$$
\begin{aligned}
z_{k+1} &= Jz_k + B_1 u_k, \quad z_0 = z^0, && (3.26)\\
Nw_{k+1} &= w_k + B_2 u_k, \quad w_0 = w^0, && (3.27)
\end{aligned}
$$

with $y_k = C_1 z_k + C_2 w_k$. Here

$$
\begin{bmatrix} z_k \\ w_k \end{bmatrix} = Tx_k, \qquad \begin{bmatrix} z^0 \\ w^0 \end{bmatrix} = Tx^0.
$$

Equation (3.26) has a unique forward solution z_k, $k \geq 0$, for any input u_k and initial value $z^0 \in \mathbb{F}^{n_f}$. This solution is given by

$$
z_k = J^k z^0 + \sum_{j=0}^{k-1} J^{k-j-1} B_1 u_j, \qquad k \geq 0.
$$

The unique solution of (3.27) has the form

$$
w_k = -\sum_{j=0}^{\nu-1} N^j B_2 u_{k+j}, \qquad k \geq 0. \tag{3.28}
$$

Thus, if the pencil $\lambda E - A$ is regular and the initial value x^0 belongs to the *set of consistent initial conditions*

$$
\mathcal{X}_d^0 = \left\{ T^{-1} \begin{bmatrix} z^0 \\ w^0 \end{bmatrix} \quad : \quad z^0 \in \mathbb{F}^{n_f},\ w^0 = -\sum_{j=0}^{\nu-1} N^j B_2 u_j \right\}.
$$

then the discrete-time descriptor system (3.2) has a unique solution x_k for all $k \geq 0$. Using the fundamental matrices F_k as in (2.7), this solution can be written as

$$
x_k = F_k E x^0 + \sum_{j=0}^{k+\nu-1} F_{k-j-1} B u_j, \qquad k \geq 0.
$$

We see that to determine x_k we need not only past inputs u_j, $j \leq k$, but also future inputs u_j, $k < j \leq k + \nu - 1$, see [36] for details. This concept is often called *noncausality* of discrete-time descriptor systems. For the *causal* descriptor system (3.2), the state x_k is determined completely by the initial vector x^0 and control inputs u_0, u_1, …, u_k. Clearly, system (3.2) is causal if the pencil $\lambda E - A$ is of index at most one.

3.2.2 Stability

In this subsection we discuss the stability of the singular difference equation

$$Ex_{k+1} = Ax_k. \tag{3.29}$$

First some notions of stability for such an equation are presented.

Definition 3.15. The trivial solution $x_k \equiv 0$ of (3.29) is called *stable in the sense of Lyapunov* or *Lyapunov stable* if

(i) for all $x^0 \in \mathbb{F}^n$ the initial value problem

$$\begin{array}{rcl} Ex_{k+1} - Ax_k & = & 0, \\ P_r\,(x_0 - x^0) & = & 0 \end{array} \tag{3.30}$$

has a unique solution $x_k \in \operatorname{Im} P_r$ defined for $k \geq 0$;

(ii) for every $\varepsilon > 0$ there exists a $\delta = \delta(\varepsilon) > 0$ such that $\|x_k\| < \varepsilon$ for all $k \geq 0$ and for all $x^0 \in \mathbb{F}^n$ with $\|P_r x^0\| < \delta$.

Definition 3.16. The trivial solution $x_k \equiv 0$ of (3.29) is called *asymptotically stable* if it is Lyapunov stable and if there is a $\delta_0 > 0$ such that for the solution x_k of (3.30) with $\|P_r x^0\| < \delta_0$ we have that $x_k \to 0$ as $k \to \infty$.

The following theorem gives a necessary and sufficient condition for the trivial solution of (3.29) to be asymptotically stable, see [36, 153] for details.

Theorem 3.17. *Let $\lambda E - A$ be a regular pencil. The trivial solution $x_k \equiv 0$ of equation* (3.29) *is asymptotically stable if and only if all finite eigenvalues of $\lambda E - A$ lie inside the unit circle.*

It should be noted that although the infinite eigenvalues lie outside the unit circle they, in contrast to the finite eigenvalues of modulus not less than 1, do not affect the behavior at infinity of solutions of (3.29).

Definition 3.18. A matrix pencil $\lambda E - A$ is called *d-stable* if it is regular and all the finite eigenvalues of $\lambda E - A$ lie inside the unit circle.

The problem of the distribution of eigenvalues of the pencil $\lambda E - A$ with respect to the unit circle has been considered in [62, 112, 113].

Similar to the continuous-time case, as a numerical parameter characterizing the d-stability of the pencil $\lambda E - A$ we take

$$\kappa_{d,2}(E, A) = \left(\|E\|_2^2 + \|A\|_2^2\right)\, \|H_d\|_2, \tag{3.31}$$

where the matrix H_d has the form

$$H_d = \sum_{k=-\infty}^{\infty} F_k^* F_k \tag{3.32}$$

and the matrices F_k are given in (2.7). If the pencil $\lambda E - A$ is d-stable, then series (3.32) is convergent and $\kappa_{d,2}(E, A)$ is bounded. We set $\kappa_{d,2}(E, A) = \infty$ if $\lambda E - A$ has at least one finite eigenvalues of modulus not less than 1. Note that $\kappa_{d,2}(E, A)$ slightly differs from the unit circle dichotomy parameter introduced in [112].

The following theorem gives bounds on the spectral norm of the matrices $F_k E$ with $k \geq 0$ and $F_k A$ with $k < 0$.

Theorem 3.19. *Let the matrices F_k be as in* (2.7). *Then*

$$\|F_k E\|_2 \leq \sqrt{\kappa_{d,2}(E,A)} \left(\frac{\kappa_{d,2}(E,A)}{1 + \kappa_{d,2}(E,A)} \right)^{k/2}, \qquad k \geq 0, \tag{3.33}$$

$$\|F_k A\|_2 \leq \sqrt{\kappa_{d,2}(E,A)} \left(\frac{\kappa_{d,2}(E,A)}{1 + \kappa_{d,2}(E,A)} \right)^{(-k-1)/2}, \qquad k < 0. \tag{3.34}$$

Proof. Note that

$$\kappa_{d,2}(E,A) \geq \|E\|_2^2 \|H_d\|_2 \geq \|E^* H_d E\|_2 \geq \|F_0 E\|_2^2 = \|P_r\|_2^2 \geq 1,$$
$$\kappa_{d,2}(E,A) \geq \|A\|_2^2 \|H_d\|_2 \geq \|A^* H_d A\|_2 \geq \|F_{-1} A\|_2^2 = \|I - P_r\|_2^2 \geq 1.$$

Using (2.8)-(2.10) for every vector $v \in \mathbb{F}^n$ and every $k > 0$ we obtain that

$$\begin{aligned}
\langle E^* H_d E F_k E v, F_k E v \rangle &= \langle E^* H_d E F_{k-1} E v, F_{k-1} E v \rangle - \langle F_{k-1} E v, F_{k-1} E v \rangle \\
&= \left(1 - \frac{\langle F_{k-1} E v, F_{k-1} E v \rangle}{\langle E^* H_d E F_{k-1} E v, F_{k-1} E v \rangle} \right) \langle E^* H_d E F_{k-1} E v, F_{k-1} E v \rangle \\
&\leq \left(1 - \frac{1}{\|E\|_2^2 \|H_d\|_2} \right) \langle E^* H_d E F_{k-1} E v, F_{k-1} E v \rangle \leq \ldots \\
&\leq \left(1 - \frac{1}{\|E\|_2^2 \|H_d\|_2} \right)^k \langle E^* H_d E F_0 E v, F_0 E v \rangle \\
&\leq \|E\|_2^2 \|H_d\|_2 \left(1 - \frac{1}{\|E\|_2^2 \|H_d\|_2} \right)^k \|P_r v\|^2.
\end{aligned}$$

From this estimate it immediately follows that

$$\|F_k E\|_2 = \max_{v \neq 0} \frac{\|F_k E v\|}{\|v\|} = \max_{P_r v \neq 0} \frac{\|F_k E v\|}{\|P_r v\|} \leq \sqrt{\kappa_{d,2}(E,A)} \left(\frac{\kappa_{d,2}(E,A)}{1 + \kappa_{d,2}(E,A)} \right)^{k/2}$$

for all $k \geq 0$. Furthermore, for $v \in \mathbb{F}^n$ and $k < -1$, we have

$$\begin{aligned}
\langle A^* H_d A F_k A v, F_k A v \rangle &= \langle A^* H_d A F_{k+1} A v, F_{k+1} A v \rangle - \langle F_{k+1} A v, F_{k+1} A v \rangle \\
&\leq \left(1 - \frac{1}{\|A\|_2^2 \|H_d\|_2} \right) \langle A^* H_d A F_{k+1} A v, F_{k+1} A v \rangle \\
&\leq \frac{\kappa_{d,2}(E,A)}{1 + \kappa_{d,2}(E,A)} \langle A^* H_d A F_{k+1} A v, F_{k+1} A v \rangle.
\end{aligned}$$

Hence,

$$\begin{aligned}
\|F_kA\|_2^2 &= \max_{v\neq 0}\frac{\|F_kAv\|^2}{\|v\|^2} = \max_{\|(I-P_r)v\|=1}\|F_kAv\|^2 = \|F_kAv_0\|^2 \\
&\le \frac{\kappa_{d,2}(E,A)}{1+\kappa_{d,2}(E,A)}\langle A^*H_dAF_{k+1}Av_0, F_{k+1}Av_0\rangle \le \dots \\
&\le \left(\frac{\kappa_{d,2}(E,A)}{1+\kappa_{d,2}(E,A)}\right)^{-k-1}\langle A^*H_dAF_{-1}Av_0, F_{-1}Av_0\rangle \\
&\le \kappa_{d,2}(E,A)\left(\frac{\kappa_{d,2}(E,A)}{1+\kappa_{d,2}(E,A)}\right)^{-k-1}.
\end{aligned}$$

Thus, for all $k<0$, estimate (3.34) holds. □

From Theorem 3.19 we obtain the following bound for the solution of (3.30).

Corollary 3.20. *Let x_k be a solution of the initial value problem* (3.30). *Then*

$$\|x_k\| \le \sqrt{\kappa_{d,2}(E,A)}\left(\frac{\kappa_{d,2}(E,A)}{1+\kappa_{d,2}(E,A)}\right)^{k/2}\|P_rx^0\|, \qquad k\ge 0. \tag{3.35}$$

Proof. Since the solution of (3.30) has the form $x_k = F_kEx^0$ for all $k\ge 0$, bound (3.35) immediately follows from (2.9), (2.10) and (3.33). □

As a consequence of Theorem 3.17 and Corollary 3.20 we have the following result.

Corollary 3.21. *Let $\kappa_{d,2}(E,A)$ be as in* (3.31). *The trivial solution of equation* (3.29) *is asymptotically stable if and only if $\kappa_{d,2}(E,A)$ is bounded.*

The numerical computation of the matrix H_d and the parameter $\kappa_{d,2}(E,A)$ will be discussed in Section 6.4.

3.3 Controllability and observability for descriptor systems

In this section we give a survey of the existing concepts of controllability and observability for descriptor systems that will be used in the sequel. In contrast to standard state space systems, for descriptor systems, there are several different notions of controllability and observability. Unfortunately, there is no uniform terminology in the literature on this subject, see [24, 35, 36, 105, 166, 174] and references therein.

Definition 3.22. Systems (3.1) and (3.2) are called *completely controllable* (*C-controllable*) if

$$\operatorname{rank}[\alpha E-\beta A,\ B] = n \quad \text{for all } (\alpha,\beta)\in\mathbb{C}^2\backslash\{(0,0)\}. \tag{3.36}$$

Systems (3.1) and (3.2) are called *controllable on a reachable set* (*R-controllable*) if

$$\operatorname{rank}[\lambda E-A,\ B] = n \quad \text{for all finite } \lambda\in\mathbb{C}\,. \tag{3.37}$$

Systems (3.1) and (3.2) are called *controllable at infinity* (*I-controllable*) if

$$\operatorname{rank}[E,\ AK_E,\ B] = n, \tag{3.38}$$

where the columns of K_E span the null space of E.
Systems (3.1) and (3.2) are called *strongly controllable* (*S-controllable*) if (3.37) and (3.38) are satisfied.

The C-controllability implies that for any given initial and final states x^0, $x_f \in \mathbb{F}^n$, there exists an admissible control input that transfers the system from x^0 to x_f in finite time. This notion follows [24, 174] and it is consistent with the definition of *controllability* given in [35, 36].

The conception of R-controllability comes from [36] and conforms to the controllability in [35, 166]. It was shown in [36] that the reachable set for a descriptor system is nothing else than the solution space. The R-controllability ensures that for any consistent initial state x^0 and final state x_f from the solution space, there exists an admissible control input that transfers the system from x^0 to x_f in finite time. In the case of $E = I$, the R-controllability coincides with the C-controllability and is the usual controllability of state space systems [88].

In the continuous-time case the I-controllability is also known as *impulse controllability* [35, 36] and means that impulsive modes in the solution can be excluded by a suitable linear state feedback control. In other words, for every initial vector x^0 there exists a feedback control $u(t) = F_c x(t) + v(t)$ with a feedback matrix $F_c \in \mathbb{F}^{m,n}$ and a new smooth control input $v(t) \in \mathbb{F}^m$ such that the closed-loop system

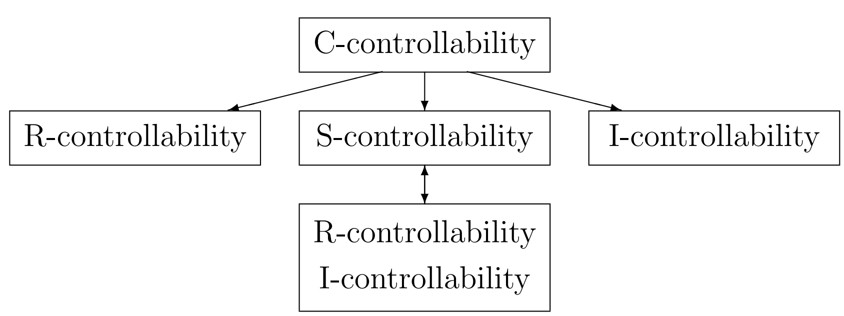
$$E\dot{x}(t) = (A + BF_c)x(t) + Bv(t), \qquad x(0) = x^0$$

has no impulsive solutions. In the discrete-time case the I-controllability implies that for every initial vector x^0 one can find a feedback control $u_k = F_c x_k + v_k$ with a feedback matrix $F_c \in \mathbb{F}^{m,n}$ and a new control input $v_k \in \mathbb{F}^m$ such that the closed-loop system

$$Ex_{k+1} = (A + BF_c)x_k + Bv_k, \qquad x_0 = x^0$$

is causal [36]. Note that the descriptor systems (3.1) and (3.2) with the pencil $\lambda E - A$ of index at most one are I-controllable.

The relationship between various controllability concepts is presented in the following diagram.



Since controllability of the descriptor systems (3.1) and (3.2) depends only on the matrices E, A and B, we will say that the triplet (E, A, B) is C(R, I, S)-controllable if system (3.1) or (3.2) is C(R, I, S)-controllable.

Definition 3.23. Let F_k be as in (2.7). The matrices

$$\mathbf{C}_+ = [\, F_0B,\ \ldots,\ F_kB,\ \ldots\,] \qquad \text{and} \qquad \mathbf{C}_- = [\,\ldots,\ F_{-k}B,\ \ldots,\ F_{-1}B\,] \tag{3.39}$$

are called the *proper* and *improper controllability matrices* in the continuous-time case and the *causal* and *noncausal controllability matrices* in the discrete-time case. The matrix

$$\mathbf{C} = [\,\mathbf{C}_-,\ \mathbf{C}_+\,] = [\,\ldots,\ F_{-k}B,\ \ldots,\ F_{-1}B,\ F_0B,\ \ldots,\ F_kB,\ \ldots\,]$$

is the *controllability matrix* of the descriptor systems (3.1) and (3.2).

The following theorem gives equivalent algebraic and geometric characterizations of different concepts of controllability for descriptor systems.

Theorem 3.24. *Consider the descriptor systems* (3.1) *and* (3.2) *with a regular pencil* $\lambda E - A$ *as in* (2.2) *and the matrices* B *and* C *as in* (3.3).

1. *The following statements are equivalent:*

   (a) *the triplet* (E, A, B) *is R-controllable;*

   (b) $\operatorname{rank}[\,\lambda I - J,\ B_1] = n_f$ *for all* $\lambda \in \mathbb{C}$*;*

   (c) $\operatorname{Im}(\lambda E - A) + \operatorname{Im} B = \mathbb{F}^n$ *for all* $\lambda \in \mathbb{C}$*;*

   (d) $\operatorname{Im}(\lambda I - J) + \operatorname{Im} B_1 = \mathbb{F}^{n_f}$ *for all* $\lambda \in \mathbb{C}$*;*

   (e) $\operatorname{rank}\big[\, F_0B,\ F_1B,\ \ldots,\ F_{n_f-1}B\,\big] = n_f$*;*

   (f) $\operatorname{rank}[\,B_1,\ JB_1,\ \ldots,\ J^{n_f-1}B_1\,] = n_f$*;*

   (g) $\operatorname{Im}\mathbf{C}_+ = \operatorname{Im} P_r$*;*

   (h) $\operatorname{Im}\,[\,B_1,\ JB_1,\ \ldots,\ J^{n_f-1}B_1\,] = \mathbb{F}^{n_f}$.

2. *The following statements are equivalent:*

   (a) *the triplet* (E, A, B) *is I-controllable;*

   (b) $\operatorname{rank}[\,N,\ K_N,\ B_2] = n_\infty$*, where the columns of* K_N *form a basis of* $\operatorname{Ker} N$*;*

   (c) $\operatorname{Ker} E + \operatorname{Im}(F_{-1}E) + \operatorname{Im}(F_{-1}B) = \operatorname{Ker} P_r$*;*

   (d) $\operatorname{Ker} N + \operatorname{Im} N + \operatorname{Im} B_2 = \mathbb{F}^{n_\infty}$*;*

   (e) $\operatorname{Ker} E + \operatorname{Im}\,[\,F_{-\nu}B,\ F_{-\nu+1}B,\ \ldots,\ F_{-1}B] = \operatorname{Ker} P_r$*;*

   (f) $\operatorname{Ker} N + \operatorname{Im}\,[\,B_2,\ NB_2,\ \ldots,\ N^{\nu-1}B_2] = \mathbb{F}^{n_\infty}$*;*

   (g) $\operatorname{Im} F_{-2} = \operatorname{Im}\,[\,F_{-\nu}B,\ F_{-\nu+1}B,\ \ldots,\ F_{-2}B]$*;*

   (h) $\operatorname{Im} N = \operatorname{Im}\,[\,NB_2,\ N^2B_2,\ \ldots,\ N^{\nu-1}B_2]$*;*

(i) $\operatorname{rank}\begin{bmatrix} E & 0 & 0 \\ A & E & B \end{bmatrix} = n + \operatorname{rank}(E)$*;*

(j) $\operatorname{rank}\begin{bmatrix} N & 0 & 0 \\ I & N & B_2 \end{bmatrix} = n_\infty + \operatorname{rank}(N)$*;*

(k) there exists a feedback matrix $F_c \in \mathbb{F}^{m,n}$ *such that the pencil* $\lambda E - (A + BF_c)$ *is regular and of index one;*

3. The following statements are equivalent:

(a) the triplet (E, A, B) *is C-controllable;*

(b) $\operatorname{rank}[\,\lambda E - A,\ B\,] = n$ *for all* $\lambda \in \mathbb{C}$ *and* $\operatorname{rank}[\,E,\ B] = n$*;*

(c) $\operatorname{rank}[\,\lambda I - J,\ B_1\,] = n_f$ *for all* $\lambda \in \mathbb{C}$ *and* $\operatorname{rank}[\,N,\ B_2\,] = n_\infty$*;*

(d) $\operatorname{Im}(\lambda E - A) + \operatorname{Im} B = \mathbb{F}^n$ *for all* $\lambda \in \mathbb{C}$ *and* $\operatorname{Im} E + \operatorname{Im} B = \mathbb{F}^n$*;*

(e) $\operatorname{Im}(\lambda I - J) + \operatorname{Im} B_1 = \mathbb{F}^{n_f}$ *for all* $\lambda \in \mathbb{C}$ *and* $\operatorname{Im} N + \operatorname{Im} B_2 = \mathbb{F}^{n_\infty}$*;*

(f) $\operatorname{rank}\left[F_0 B,\ F_1 B,\ \ldots,\ F_{n_f-1} B\right] = n_f$, $\operatorname{rank}[F_{-\nu} B,\ F_{-\nu+1} B,\ \ldots,\ F_{-1} B\,] = n_\infty$*;*

(g) $\operatorname{rank}[\,B_1,\ JB_1,\ \ldots,\ J^{n_f-1} B_1\,] = n_f$ *and* $\operatorname{rank}[\,B_2,\ NB_2,\ \ldots,\ N^{\nu-1} B_2\,] = n_\infty$*;*

(h) $\operatorname{Im} \mathbf{C}_+ = \operatorname{Im} P_r$ *and* $\operatorname{Im} \mathbf{C}_- = \operatorname{Ker} P_r$.

Proof. See [35, 36, 83, 174]. □

Observability is a dual concept of controllability.

Definition 3.25. Systems (3.1) and (3.2) are called *completely observable* (*C-observable*) if

$$\operatorname{rank}\begin{bmatrix} \alpha E - \beta A \\ C \end{bmatrix} = n \quad \text{for all } (\alpha, \beta) \in \mathbb{C}^2 \backslash \{(0,0)\}. \tag{3.40}$$

Systems (3.1) and (3.2) are called *observable on the reachable set* (*R-observable*) if

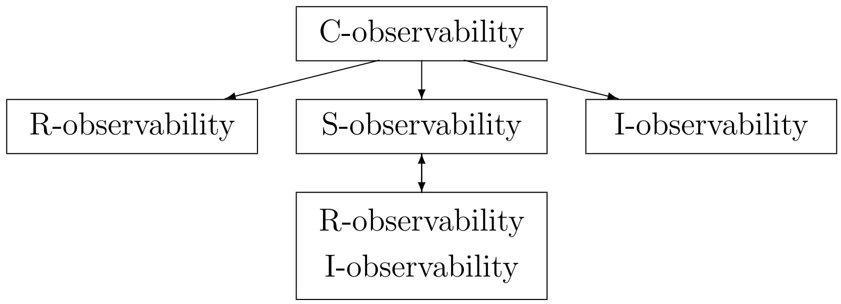
$$\operatorname{rank}\begin{bmatrix} \lambda E - A \\ C \end{bmatrix} = n \quad \text{for all finite } \lambda \in \mathbb{C}\,. \tag{3.41}$$

System (3.1) and (3.2) are called *observable at infinity* (*I-observable*) if

$$\operatorname{rank}\begin{bmatrix} E \\ K_{E^*}^* A \\ C \end{bmatrix} = n, \tag{3.42}$$

where the columns of K_{E^*} span the null space of E^*.
Systems (3.1) and (3.2) are called *strongly observable* (*S-observable*) if (3.41) and (3.42) are satisfied.

The relationship between various observability concepts is presented in the following diagram.



Observability of the descriptor systems (3.1) and (3.2) depends only on the matrices E, A and C. Therefore, the triplet (E, A, C) will be referred to as C(R, I, S)-observable if system (3.1) or (3.2) is C(R, I, S)-observable.

Definition 3.26. Let F_k be as in (2.7). The matrices

$$\mathbf{O}_+ = \begin{bmatrix} CF_0 \\ \vdots \\ CF_k \\ \vdots \end{bmatrix} \qquad \text{and} \qquad \mathbf{O}_- = \begin{bmatrix} \vdots \\ CF_{-k} \\ \vdots \\ CF_{-1} \end{bmatrix} \tag{3.43}$$

are called the *proper* and *improper observability matrices* of the continuous-time descriptor system (3.1) and the *causal* and *noncausal observability matrices* of the discrete-time descriptor system (3.2). The matrix

$$\mathbf{O} = \begin{bmatrix} \mathbf{O}_- \\ \mathbf{O}_+ \end{bmatrix}$$

is the *observability matrix* of the descriptor systems (3.1) and (3.2).

The following theorem gives equivalent algebraic and geometric characterizations of different concepts of observability for descriptor systems.

Theorem 3.27. *Consider the descriptor systems* (3.1) *and* (3.2) *with a regular pencil* $\lambda E - A$ *as in* (2.2) *and the matrices* B *and* C *as in* (3.3).

1. *The following statements are equivalent:*

   (a) the triplet (E, A, C) *is R-observable;*

   (b) $\operatorname{rank}\,[\,\lambda I - J^*,\ C_1^*] = n_f$ *for all* $\lambda \in \mathbb{C}$*;*

   (c) $\operatorname{Ker}(\lambda E - A) \bigcap \operatorname{Ker} C = 0$ *for all* $\lambda \in \mathbb{C}$*;*

   (d) $\operatorname{Ker}(\lambda I - J) \bigcap \operatorname{Ker} C_1 = 0$ *for all* $\lambda \in \mathbb{C}$*;*

   (e) $\operatorname{rank}\left[\, F_0^* C^*,\ F_1^* C^*,\ \ldots,\ F_{n_f-1}^* C^* \,\right] = n_f$*;*

(f) $\operatorname{rank}[\,C_1^*,\ J^*C_1^*,\ \ldots,\ (J^{n_f-1})^*C_1^*\,] = n_f$;

(g) $\operatorname{Ker}\mathbf{O}_+ \bigcap \operatorname{Im} P_l = 0$;

(h) $\operatorname{Ker}\,[\,C_1^*,\ J^*C_1^*,\ \ldots,\ (J^{n_f-1})^*C_1^*\,] = 0$.

2. *The following statements are equivalent:*

   (a) *the triplet (E, A, C) is I-observable;*

   (b) $\operatorname{rank}[\,N^*,\ K_{N^*},\ C_2^*] = n_\infty$, *where the columns of K_{N^*} span* $\operatorname{Ker} N^*$;

   (c) $\operatorname{Ker}(EF_{-1}) \bigcap \operatorname{Im} E \bigcap \operatorname{Ker}(CF_{-1}) = \operatorname{Im} P_l$;

   (d) $\operatorname{Ker} N \bigcap \operatorname{Im} N \bigcap \operatorname{Ker} C_2 = 0$;

   (e) $\bigcap_{k=1}^{\nu} \operatorname{Ker}(CF_{-k}) \bigcap \operatorname{Im} E = \operatorname{Im} P_l$;

   (f) $\bigcap_{k=0}^{\nu-1} \operatorname{Ker}(C_2N^k) \bigcap \operatorname{Im} N = 0$;

   (g) $\operatorname{Ker} F_{-2} = \bigcap_{k=1}^{\nu} \operatorname{Ker}(CF_{-k})$;

   (h) $\operatorname{Ker} N = \bigcap_{k=1}^{\nu-1} \operatorname{Ker}(C_2N^k)$;

   (i) $\operatorname{rank}\begin{bmatrix} E^* & 0 & 0 \\ A^* & E^* & C^* \end{bmatrix} = n + \operatorname{rank}(E)$;

   (j) $\operatorname{rank}\begin{bmatrix} N^* & 0 & 0 \\ I & N^* & C_2^* \end{bmatrix} = n_\infty + \operatorname{rank}(N)$;

   (k) *there exists a feedback matrix $F_o \in \mathbb{F}^{n,p}$ such that the pencil $\lambda E - (A + F_oC)$ is regular and of index one;*

3. *The following statements are equivalent:*

   (a) *the triplet (E, A, C) is C-observable;*

   (b) $\operatorname{rank}[\,\lambda E^* - A^*,\ C^*\,] = n$ *for all $\lambda \in \mathbb{C}$ and* $\operatorname{rank}[\,E^*,\ C^*] = n$;

   (c) $\operatorname{rank}[\,\lambda I - J^*,\ C_1^*\,] = n_f$ *for all $\lambda \in \mathbb{C}$ and* $\operatorname{rank}[\,N^*,\ C_2^*\,] = n_\infty$;

   (d) $\operatorname{Ker}(\lambda E - A) \bigcap \operatorname{Ker} C = 0$ *for all $\lambda \in \mathbb{C}$ and* $\operatorname{Ker} E \bigcap \operatorname{Ker} C = 0$;

   (e) $\operatorname{Ker}(\lambda I - J) \bigcap \operatorname{Ker} C_1 = 0$ *for all $\lambda \in \mathbb{C}$ and* $\operatorname{Ker} N \bigcap \operatorname{Ker} C_2 = 0$;

   (f) $\operatorname{rank}\Big[F_0^*C^*,\ F_1^*C^*,\ \ldots,\ F_{n_f-1}^*C^*\Big] = n_f$ *and* $\operatorname{rank}\Big[F_{-\nu}^*C^*,\ \ldots,\ F_{-1}^*C^*\Big] = n_\infty$;

   (g) $\operatorname{rank}[C_1^*,\ J^*C_1^*,\ \ldots,\ (J^{n_f-1})^*C_1^*] = n_f$, $\operatorname{rank}[C_2^*,\ N^*C_2^*,\ \ldots,\ (N^{\nu-1})^*C_2^*] = n_\infty$;

   (h) $\operatorname{Ker}\mathbf{O}_+ \bigcap \operatorname{Im} P_l = 0$ *and* $\operatorname{Ker}\mathbf{O}_- \bigcap \operatorname{Ker} P_l = 0$.

Proof. See [35, 36, 83, 174]. □

Here we have considered the stability and various concepts of controllability and observability for the continuous-time and discrete-time descriptor systems. In the next chapter we will show how these properties of descriptor systems can be characterized in terms of solutions of generalized Lyapunov equations.

Chapter 4

Generalized Lyapunov equations

Generalized continuous-time algebraic Lyapunov equations (GCALEs)

$$E^*XA + A^*XE = -G \tag{4.1}$$

and generalized discrete-time algebraic Lyapunov equations (GDALEs)

$$A^*XA - E^*XE = -G \tag{4.2}$$

arise in many fields of mathematics and engineering such as stability analysis for differential and difference equations [53, 61, 123, 148, 149], problems of spectral dichotomy [62, 113, 116] and control theory [11, 58, 98, 117, 119, 176].

For $E = I_n$, equations (4.1) and (4.2) are the standard continuous-time and discrete-time Lyapunov equations (the latter is also known as the Stein equation). These equations have been the topic of numerous publications, see [9, 51, 53, 72, 99] and the references therein. The case of nonsingular E has been considered in [17, 34, 125]. However, many applications of descriptor systems lead to generalized Lyapunov equations with a singular matrix E, see [11, 116, 120, 148, 149, 153]. In this chapter we study the existence and uniqueness of solutions of generalized Lyapunov equations with general and special right-hand sides.

The classical stability and inertia theorems [20, 29, 33, 37, 108, 122, 172, 173] characterize connections between the signatures of solutions of standard Lyapunov equations and the numbers of eigenvalues of a matrix in the left and right open half-planes and on the imaginary axis in the continuous-time case and inside, outside and on the unit circle in the discrete-time case. A brief survey of matrix inertia theorems and their applications has been presented in [38]. In this chapter we extend these theorems to matrix pencils.

4.1 Applications for generalized Lyapunov equations

In this section we present some applications for generalized Lyapunov equations.

Stability analysis

It is well known that the asymptotic behavior of solutions of differential and difference equations is closely related to the analysis of Lyapunov equations [61, 74, 123, 151, 147, 153].

Consider the continuous-time singular system (3.13). The trivial solution of (3.13) is asymptotically stable if there exists a matrix X that is Hermitian, positive definite on the subspace $\operatorname{Im} P_l$ and satisfies the GCALE (4.1), where G is Hermitian and positive definite on $\operatorname{Im} P_r$, see [120, 123]. For such a solution X, the matrix E^*XE is Hermitian, positive definite on the subspace $\operatorname{Im} P_r$, and we obtain that

$$\mathbf{V}(t) := x^*(t)E^*XEx(t) > 0, \qquad t \in [\,0, \infty)$$

for all nonzero solutions $x(t) \in \operatorname{Im} P_r$ of equation (3.13). Moreover, we have

$$\dot{\mathbf{V}}(t) = x^*(t)(A^*XE + E^*XA)x(t) = -x^*(t)Gx(t) < 0.$$

The quadratic form $\mathbf{V}(t)$ is the *generalized Lyapunov function* for system (3.13).

Similarly, the trivial solution of the discrete-time singular system (3.29) is asymptotically stable if there exists a matrix X that is Hermitian, positive definite on the subspace $\operatorname{Im} P_l$ and satisfies the GDALE (4.2), where G is Hermitian, positive definite on $\operatorname{Im} P_r$ [123, 151]. In this case a quadratic form $\mathbf{V}_k := x_k^*E^*XEx_k$ presents the *generalized Lyapunov function* for system (3.29). We have that $\mathbf{V}_k > 0$ for all nonzero solutions $x_k \in \operatorname{Im} P_r$ of (3.29) and

$$\mathbf{V}_{k+1} - \mathbf{V}_k = x_k^*(A^*XA - E^*XE)x_k = -x_k^*Gx_k < 0.$$

These results are generalizations of the known connection between the standard Lyapunov equations and the standard state space differential/difference equations [50, 61, 74].

Linear-quadratic optimal control

Consider the linear-quadratic optimal regulator control problems:

Minimize the cost functional

$$\mathcal{J}_c(x^0, u) = \frac{1}{2}\int_0^\infty \Big(y^*(t)Qy(t) + u^*(t)Ru(t)\Big)\,dt \tag{4.3}$$

subject to the continuous-time descriptor system

$$\begin{array}{rcl} E\dot{x}(t) & = & Ax(t) + Bu(t), \quad x(t) = x^0, \\ y(t) & = & Cx(t), \end{array} \tag{4.4}$$

and

Minimize the cost functional

$$\mathcal{J}_d(x^0, u) = \frac{1}{2}\sum_{k=0}^\infty \Big(y_k^*Qy_k + u_k^*Ru_k\Big) \tag{4.5}$$

subject to the discrete-time descriptor system

$$\begin{array}{rcl} Ex_{k+1} & = & Ax_k + Bu_k, \quad x_0 = x^0, \\ y_k & = & Cx_k, \end{array} \tag{4.6}$$

where E, $A \in \mathbb{F}^{n,n}$, $B \in \mathbb{F}^{m,n}$, $C \in \mathbb{F}^{p,n}$, $Q \in \mathbb{F}^{p,p}$ is Hermitian and $R \in \mathbb{F}^{m,m}$ is Hermitian, positive definite.

Under stabilizability and detectability conditions, see [117], the optimal solution of the continuous-time minimization problem (4.3)-(4.4) is given by

$$u_{opt}(t) = -R^{-1}B^*XEx(t),$$

where X is an Hermitian, positive semidefinite solution of the generalized continuous-time Riccati equation

$$E^*XA + A^*XE - E^*XBR^{-1}B^*XE + C^*QC = 0. \tag{4.7}$$

In the discrete-time case, the optimal solution of the minimization problem (4.5)-(4.6) is given by

$$(u_{opt})_k = -(R + B^*XB)^{-1}B^*XAx_k,$$

where X is an Hermitian, positive semidefinite solution of the generalized discrete-time Riccati equation

$$A^*XA - E^*XE - A^*XB(R + B^*XB)^{-1}B^*XA + C^*QC = 0, \tag{4.8}$$

see [12, 117] for details. The generalized Riccati equations (4.7) and (4.8) can be solved by Newton's method [12, 98, 117].

Algorithm 4.1.1. *Newton's method for the continuous-time Riccati equation*
Input: *Matrices E, A, B, C, Q, R and a starting stabilizing guess X_0, E is nonsingular.*
Output: *An approximate solution X_{k+1} of the generalized Riccati equation* (4.7)*.*
FOR $k = 0, 1, 2, \ldots$
1. $A_k = A - BR^{-1}B^*X_kE$.
2. $R_k = E^*X_kA_k + A_k^*X_kE + E^*X_kBR^{-1}B^*X_kE + C^*QC$.
3. *Solve the GCALE* $\quad E^*Y_kA_k + A_k^*Y_kE = -R_k$.
4. $X_{k+1} = X_k + Y_k$.

END FOR

Algorithm 4.1.2. *Newton's method for the discrete-time Riccati equation*
Input: *Matrices E, A, B, C, Q, R and a starting stabilizing guess X_0, E is nonsingular.*
Output: *An approximate solution X_{k+1} of the generalized Riccati equation* (4.8)*.*
FOR $k = 0, 1, 2, \ldots$
1. $K_k = (R + B^*X_kB)^{-1}B^*X_kA$.
2. $A_k = A - BK_k$.
3. $R_k = A_k^*X_kA_k - E^*X_kE + K_k^*RK_k + C^*QC$.
4. *Solve the GDALE* $\quad A_k^*Y_kA_k - E^*Y_kE = -R_k$.
5. $X_{k+1} = X_k + Y_k$.

END FOR

We see that in each iteration step of Algorithms 4.1.1 and 4.1.2 we need to solve generalized Lyapunov equations.

The Lyapunov equations arise also in many other fields of control theory such as system balancing [102, 119], $\mathbb{H}_\infty$ control [58, 176] and model reduction [4, 48, 119, 137, 150].

4.2 Generalized continuous-time Lyapunov equations

In this section we present general results concerning the solution of the GCALE

$$E^*XA + A^*XE = -G, \tag{4.9}$$

where E, A, $G \in \mathbb{F}^{n,n}$ are given matrices and $X \in \mathbb{F}^{n,n}$ is an unknown matrix.

4.2.1 General case

Consider a *continuous-time Lyapunov operator* $\mathcal{L}_c : \mathbb{F}^{n,n} \to \mathbb{F}^{n,n}$ given by

$$\mathcal{L}_c(X) := E^*XA + A^*XE. \tag{4.10}$$

The GCALE (4.9) can be written in the operator form

$$\mathcal{L}_c(X) = -G. \tag{4.11}$$

If $x = \mathrm{vec}(X)$ and $g = \mathrm{vec}(G)$ are vectors of order n^2 obtained by stacking the columns of the matrices X and G, respectively, then we can also rewrite the GCALE (4.9) as a linear system

$$\mathbf{L}_c x = -g, \tag{4.12}$$

where the $n^2 \times n^2$-matrix

$$\mathbf{L}_c = E^T \otimes A^* + A^T \otimes E^* \tag{4.13}$$

is the matrix representation of the continuous-time Lyapunov operator $\mathcal{L}_c$, see, e.g., [78]. In this case we may apply the theory of linear systems [53, 99] to determine conditions for the existence and uniqueness of solutions of the GCALE (4.9).

Theorem 4.1. [99] *Let* $\mathbf{L}_c$ *be as in* (4.13) *and let* $x = \mathrm{vec}(X)$, $g = \mathrm{vec}(G)$. *The GCALE* (4.9) *has a solution if and only if* $\mathrm{rank}\,[\,\mathbf{L}_c, g\,] = \mathrm{rank}\,\mathbf{L}_c$. *There exists a unique solution of* (4.9) *if and only if the matrix* $\mathbf{L}_c$ *is nonsingular.*

Note that already for moderately large n the matrix $\mathbf{L}_c$ is very large. Therefore, the equivalent formulation (4.12) for the GCALE (4.9) is only of theoretical interest.

The GCALE (4.9) is a special case of the generalized Sylvester equation

$$BXA - FXE = -G, \tag{4.14}$$

where A, B, E, F, $G \in \mathbb{F}^{n,n}$ are given matrices and $X \in \mathbb{F}^{n,n}$ is an unknown matrix. The following theorem gives the necessary and sufficient conditions for unique solvability of equation (4.14).

Theorem 4.2. [34] *The generalized Sylvester equation* (4.14) *has a unique solution* X *if and only if the pencils* $\lambda B - F$ *and* $\lambda E - A$ *are regular and they have no common eigenvalues.*

As a consequence of Theorem 4.2 we have the following necessary and sufficient conditions for the existence and uniqueness of solutions of the GCALE (4.9) in terms of the spectrum of the pencil $\lambda E - A$.

Theorem 4.3. [125] *Let* $\lambda E - A$ *be a regular pencil with eigenvalues* $\{\lambda_1, \ldots, \lambda_n\}$ *counted according to their multiplicities. The GCALE* (4.9) *has a unique solution for every matrix* G *if and only if all eigenvalues of the pencil* $\lambda E - A$ *are finite and* $\lambda_j + \overline{\lambda}_k \neq 0$ *for all* $j,\, k = 1, \ldots, n$.

The GCALE (4.9) is said to be *regular* if it is uniquely solvable. For such an equation, the finiteness of the eigenvalues of $\lambda E - A$ implies the nonsingularity of E, while the condition $\lambda_j + \overline{\lambda}_k \neq 0$ implies that the pencil $\lambda E - A$ has no eigenvalues on the imaginary axis and, hence, the matrix A is also nonsingular. The GCALE (4.9) is called *non-degenerate* if both matrices E and A are nonsingular. Otherwise, the GCALE (4.9) is called *degenerate*.

The non-degenerate GCALE (4.9) is equivalent to standard Lyapunov equations

$$\begin{aligned} XAE^{-1} + (AE^{-1})^* X &= -E^{-*} G E^{-1}, \\ (EA^{-1})^* X + XEA^{-1} &= -A^{-*} G A^{-1}. \end{aligned} \tag{4.15}$$

In this case classical Lyapunov theorems [53] on the existence and uniqueness of positive definite solutions of these equations can be extended to the GCALE (4.9).

Theorem 4.4. *Let* $\lambda E - A$ *be a regular pencil. If all eigenvalues of* $\lambda E - A$ *are finite and lie in the open left half-plane, then for every Hermitian, positive* (*semi*)*definite matrix* G, *the GCALE* (4.9) *has a unique Hermitian, positive* (*semi*)*definite solution* X. *Conversely, if there exist Hermitian, positive definite matrices* X *and* G *satisfying* (4.9), *then all eigenvalues of the pencil* $\lambda E - A$ *are finite and lie in the open left half-plane.*

The degenerate GCALE (4.9) is *singular* in the sense that it may have no solutions even if all finite eigenvalues of the pencil $\lambda E - A$ have negative real part. Since E and A play a symmetric role in (4.9), in the sequel we will assume that the matrix E is singular.

Example 4.5. The GCALE (4.9) with

$$E = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad A = -I_2, \qquad G = I_2$$

has no solutions.

Even if a solution of the degenerate GCALE (4.9) exists, it is not unique. Indeed, if X is a solution of the degenerate GCALE (4.9), then for any nonzero vector $v \in \operatorname{Ker} E^*$, the matrix $X + vv^*$ satisfies (4.9) as well.

The GCALE (4.9) is closely related to the study of the asymptotic properties of solutions of the differential-algebraic equation (3.13), e.g., [120, 123, 147]. The following theorem gives sufficient conditions for the pencil $\lambda E - A$ to be c-stable or, equivalently, for the trivial solution of (3.13) to be asymptotically stable.

Theorem 4.6. *Let P_l and P_r be the spectral projections onto the left and right finite deflating subspaces of a regular pencil $\lambda E - A$ and let G be a matrix that is Hermitian, positive definite on the subspace $\operatorname{Im} P_r$. If the GCALE (4.9) has a solution X which is Hermitian and positive definite on $\operatorname{Im} P_l$, then the pencil $\lambda E - A$ is c-stable.*

Proof. Let the pencil $\lambda E - A$ be in Weierstrass canonical form (2.2) and let the Hermitian matrix

$$X = W^{-*} \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{12}^* & Y_{22} \end{bmatrix} W^{-1} \tag{4.16}$$

satisfy the GCALE (4.9). If X is positive definite on $\operatorname{Im} P_l$, then Y_{11} in (4.16) is positive definite, and, hence, the matrix

$$E^* X E = T^* \begin{bmatrix} Y_{11} & Y_{12} N \\ N^* Y_{12}^* & N^* Y_{22} N \end{bmatrix} T$$

is Hermitian, positive definite on the subspace $\operatorname{Im} P_r$.

Let $v \neq 0$ be an eigenvector of the pencil $\lambda E - A$ corresponding to a finite eigenvalue λ, that is, $\lambda E v = A v$ and $v \in \operatorname{Im} P_r$. Multiplication of (4.9) on the right and left by v and v^*, respectively, gives

$$\begin{aligned} -v^* G v &= v^*(E^* X A + A^* X E) v = \lambda\, v^* E^* X E v + \overline{\lambda}\, v^* E^* X E v \\ &= 2\,\Re e(\lambda)\, v^* E^* X E v. \end{aligned} \tag{4.17}$$

Since G and $E^* X E$ are positive definite on $\operatorname{Im} P_r$, we obtain that $\Re e(\lambda) < 0$, i.e., all finite eigenvalues of the pencil $\lambda E - A$ lie in the open left half-plane. □

Example 4.5 demonstrates that the c-stability of the pencil $\lambda E - A$ does not imply the solvability of the degenerate GCALE (4.9).

It follows from (4.17) that the condition for X to be positive definite on $\operatorname{Im} P_l$ can be replaced by the assumption that X is positive semidefinite on $\mathbb{F}^n$. Thus, we obtain the following result.

Corollary 4.7. *Let P_r be the spectral projection onto the right finite deflating subspace of a regular pencil $\lambda E - A$ and let G be a matrix that is Hermitian, positive definite on $\operatorname{Im} P_r$. If the GCALE (4.9) has an Hermitian, positive semidefinite solution X, then $\lambda E - A$ is c-stable.*

4.2.2 Special right-hand side: index 1 and 2 cases

Consider the generalized continuous-time Lyapunov equation

$$E^*XA + A^*XE = -E^*GE. \tag{4.18}$$

Such an equation has been studied first in [104, 116]. The presence of E in the right-hand side guarantees the solvability of the GCALE (4.18) under assumptions that the pencil $\lambda E - A$ is c-stable and its index does not exceed two.

Theorem 4.8. *Let $\lambda E - A$ be a regular pencil of index at most two. If $\lambda E - A$ is c-stable, then for every matrix G, the GCALE* (4.18) *has a solution. For all solutions X of* (4.18)*, the matrix E^*XE is unique. Moreover, if G is positive definite, then every solution X of* (4.18) *is positive definite on* $\operatorname{Im} P_l$.

Proof. Let the pencil $\lambda E - A$ be in Weierstrass canonical form (2.2), where the eigenvalues of J lie in the open left half-plane and $N^2 = 0$ by assumption. Let the matrices

$$W^*GW = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \qquad \text{and} \qquad W^*XW = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \tag{4.19}$$

be partitioned in blocks conformally to E and A. Then from (4.18) we have

$$Y_{11}J + J^*Y_{11} = -W_{11}, \tag{4.20}$$

$$Y_{12} + J^*Y_{12}N = -W_{12}N, \tag{4.21}$$

$$N^*Y_{21}J + Y_{21} = -N^*W_{21}, \tag{4.22}$$

$$N^*Y_{22} + Y_{22}N = -N^*W_{22}N. \tag{4.23}$$

Since all eigenvalues of J have negative real part, the standard Lyapunov equation (4.20) has a unique solution Y_{11} for every W_{11}, see [53]. Taking into account that the matrices J^{-*} and $-N$ have disjoint spectra, equations (4.21) and (4.22) are uniquely solvable [99] and their solutions are given by $Y_{12} = -W_{12}N$ and $Y_{21} = -N^*W_{21}$. Equation (4.23) has a (nonunique) solution for every W_{22}. For example, the matrix $Y_{22} = -\frac{1}{2}(N^*W_{22} + W_{22}N)$ satisfies (4.23).

Thus, every solution of the GCALE (4.18) has the form

$$X = W^{-*} \begin{bmatrix} Y_{11} & -W_{12}N \\ -N^*W_{21} & Y_{22} \end{bmatrix} W^{-1}, \tag{4.24}$$

where Y_{11} and Y_{22} satisfy equations (4.20) and (4.23), respectively. Multiplying equation (4.23) on the right by the matrix N we obtain that $N^*Y_{22}N = 0$ holds for every solution Y_{22} of (4.23). Since equation (4.20) has the unique solution Y_{11}, the matrix

$$E^*XE = T^* \begin{bmatrix} Y_{11} & -W_{12}N^2 \\ -(N^*)^2W_{21} & N^*Y_{22}N \end{bmatrix} T = T^* \begin{bmatrix} Y_{11} & 0 \\ 0 & 0 \end{bmatrix} T \tag{4.25}$$

is uniquely defined for all solutions X of (4.18). If G is positive definite, then also W_{11} is positive definite and, hence, the solution Y_{11} of (4.20) is positive definite. Then X in (4.24) is positive definite on $\operatorname{Im} P_l$. □

Note that the assumption for $\lambda E - A$ to be of index at most two is important, since otherwise the GCALE (4.18) may have no solutions even if the pencil $\lambda E - A$ is c-stable. To understand better what happens if the index of the pencil is increased from two to three, we consider the following example.

Example 4.9. Let $A = -I_n$, $G = I_n$, $X = [x_{ij}]_{i,j=1}^n$ and $E = N_n$. Taking these matrices with $n = 2$ in (4.18), we have the equation

$$\begin{bmatrix} 0 & x_{11} \\ x_{11} & x_{12} + x_{21} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

which has the solution set

$$\left\{ X = \begin{bmatrix} 0 & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \quad : \quad x_{12} + x_{21} = 1 \right\}.$$

For $n = 3$ we obtain the equation

$$\begin{bmatrix} 0 & x_{11} & x_{12} \\ x_{11} & x_{12} + x_{21} & x_{13} + x_{22} \\ x_{21} & x_{22} + x_{31} & x_{23} + x_{32} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

which has no solution.

If G is Hermitian, then (4.18) has Hermitian as well as non-Hermitian solutions (see, Example 4.9), while the matrix E^*XE is Hermitian for every solution X of (4.18). If G is positive definite on $\mathbb{F}^n$, then E^*XE is positive semidefinite on $\mathbb{F}^n$ and positive definite on $\operatorname{Im} P_r$.

Remark 4.10. Theorem 4.8 still holds if the matrix G in the GCALE (4.18) is positive definite only on the subspace $\operatorname{Im} P_l$.

The converse of Theorem 4.8 also holds.

Theorem 4.11. *Let $\lambda E - A$ be a regular pencil and let G be an Hermitian, positive definite matrix. If the GCALE* (4.18) *has a solution X, then the pencil $\lambda E - A$ is of index at most two. Moreover, if X is Hermitian, positive definite on* $\operatorname{Im} P_l$*, then $\lambda E - A$ is c-stable.*

Proof. Let the pencil $\lambda E - A$ be in Weierstrass canonical form (2.2) and let the matrices G and X as in (4.19) satisfy the GCALE (4.18). Then equations (4.20)–(4.23) are fulfilled.

Let the matrices

$$Y_{22} = \begin{bmatrix} \check{Y}_{11} & \cdots & \check{Y}_{1t} \\ \vdots & \ddots & \vdots \\ \check{Y}_{t1} & \cdots & \check{Y}_{tt} \end{bmatrix} \qquad \text{and} \qquad W_{22} = \begin{bmatrix} \check{W}_{11} & \cdots & \check{W}_{1t} \\ \vdots & \ddots & \vdots \\ \check{W}_{t1} & \cdots & \check{W}_{tt} \end{bmatrix}$$

be partitioned in blocks conformally to $N = \mathrm{diag}(N_{n_1}, \ldots, N_{n_t})$, where N_{n_j} is a nilpotent Jordan block of order n_j. In this case equation (4.23) is equivalent to the system of matrix equations

$$N_{n_p}^* \check{Y}_{pq} + \check{Y}_{pq} N_{n_q} = -N_{n_p}^* \check{W}_{pq} N_{n_q}, \qquad p, q = 1, \ldots, t. \tag{4.26}$$

Since G is Hermitian and positive definite, also all $\check{W}_{jj}$ are Hermitian and positive definite. Assume that the index of the pencil $\lambda E - A$ is larger than two, i.e., there exists a block N_{n_k} with $n_k > 2$. Let $\check{Y}_{kk} = [y_{ij}]_{i,j=1}^{n_k}$ and $\check{W}_{kk} = [w_{ij}]_{i,j=1}^{n_k}$. It is easy to verify that

$$\begin{array}{rcll} (N_{n_k}^* \check{Y}_{kk})_{ij} & = & x_{i-1,j}, & i, j = 1, 2, \ldots, n_k, \\ (\check{Y}_{kk} N_{n_k})_{ij} & = & x_{i,j-1}, & i, j = 1, 2, \ldots, n_k, \\ (N_{n_k}^* \check{W}_{kk} N_{n_k})_{ij} & = & w_{i-1,j-1}, & i, j = 1, 2, \ldots, n_k, \end{array}$$

where we have set

$$y_{0j} = y_{j0} = w_{0j} = w_{j0} = w_{00} = 0, \qquad j = 1, 2, \ldots, n_k. \tag{4.27}$$

It follows from (4.26) for $p = q = k$ that

$$y_{i-1\,j} + y_{i\,j-1} = -w_{i-1\,j-1}, \qquad i, j = 1, 2, \ldots, n_k. \tag{4.28}$$

Hence, by (4.27) we obtain $y_{1\,j-1} = y_{j-1\,1} = 0$ for all $j = 2, \ldots, n_k$. Then it follows from (4.28) that $w_{11} = -y_{12} - y_{21} = 0$ which contradicts the positive definiteness of $\check{W}_{kk}$. Thus, the index of the pencil $\lambda E - A$ is at most two.

Taking into account that $E^* G E$ is Hermitian, positive definite on $\mathrm{Im}\, P_r$ and X is Hermitian, positive definite on $\mathrm{Im}\, P_l$, we have from Theorem 4.6 that all finite eigenvalues of the pencil $\lambda E - A$ lie in the open left half-plane. □

If we replace the condition for the solution X of (4.18) to be positive definite on $\mathrm{Im}\, P_l$ by the assumption that X is positive semidefinite on $\mathbb{F}^n$, then we obtain the following result.

Corollary 4.12. *Let $\lambda E - A$ be a regular pencil and let G be an Hermitian, positive definite matrix. The GCALE* (4.18) *has an Hermitian, positive* (*semi*)*definite solution X if and only if the index of the pencil $\lambda E - A$ is at most one and $\lambda E - A$ is c-stable.*

Proof. If the pencil $\lambda E - A$ is of index at most one and c-stable, then from the proof of Theorem 4.8 we obtain that the matrix

$$X = W^{-*} \begin{bmatrix} Y_{11} & 0 \\ 0 & Y_{22} \end{bmatrix} W^{-1}$$

satisfies the GCALE (4.18). Here Y_{11} is a unique Hermitian, positive definite solution of (4.20) and Y_{22} is an arbitrary Hermitian, positive (semi)definite matrix. In this case X is the Hermitian, positive (semi)definite solution of (4.18).

Assume now that the GCALE (4.18) with Hermitian, positive definite G has an Hermitian, positive (semi)definite solution X. Then equations (4.20)–(4.23) are fulfilled. It follows from Theorem 4.11 that the index of $\lambda E - A$ is at most two and $N^*Y_{22}N = 0$, where Y_{22} is Hermitian, positive (semi)definite. Hence, $Y_{22}N = 0$ and $N^*W_{22}N = 0$. Since W_{22} is Hermitian, positive definite, we get $N = 0$. Thus, $\lambda E - A$ is of index at most one. Furthermore, by Corollary 4.7 the pencil $\lambda E - A$ is c-stable. □

4.2.3 Projected continuous-time Lyapunov equations

As we have seen above, the presence of the eigenvalue at infinity in the pencil $\lambda E - A$ may be a reason for the unsolvability of the GCALE (4.9). A consideration of the GCALE (4.18) with a special right-hand side is only partially useful since for such a equation, the existence theorems can be stated only for pencils of index at most two. To overcome this difficulty we consider the following generalized continuous-time Lyapunov equation

$$E^*XA + A^*XE = -P_r^*GP_r, \tag{4.29}$$

where P_r is the spectral projection onto the right finite deflating subspace of the pencil $\lambda E - A$. The following theorem gives necessary and sufficient conditions for the existence of solutions of the GCALE (4.29). Note that these conditions are independent of the index of $\lambda E - A$.

Theorem 4.13. *Let $\lambda E - A$ be a regular pencil with finite eigenvalues $\{\lambda_1, \ldots \lambda_{n_f}\}$ counted according to their multiplicities and let P_r and P_l be the spectral projections as in (2.3). The GCALE* (4.29) *has a solution for every matrix G if and only if $\lambda_j + \overline{\lambda}_k \neq 0$ for all $j, k = 1, \ldots, n_f$. Moreover, if the solution X of* (4.29) *satisfies $X = XP_l$, then it is unique.*

Proof. Let the pencil $\lambda E - A$ be in Weierstrass canonical form (2.2), where J has eigenvalues $\{\lambda_1, \ldots \lambda_{n_f}\}$, and let the matrices

$$T^{-*}GT^{-1} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \qquad \text{and} \qquad W^*XW = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \tag{4.30}$$

be partitioned in blocks accordingly to E and A. We have from (4.29) the decoupled system of equations

$$Y_{11}J + J^*Y_{11} = -T_{11}, \tag{4.31}$$
$$Y_{12} + J^*Y_{12}N = 0, \tag{4.32}$$
$$N^*Y_{21}J + Y_{21} = 0, \tag{4.33}$$
$$N^*Y_{22} + Y_{22}N = 0. \tag{4.34}$$

The Lyapunov equation (4.31) has a unique solution Y_{11} for every matrix T_{11} if and only if $\lambda_j + \overline{\lambda}_k \neq 0$ for any two eigenvalues λ_j and λ_k of J, see [99]. The homogeneous equations (4.32) and (4.33) are uniquely solvable [99] and have the trivial solutions $Y_{12} = 0$ and $Y_{21} = 0$. Equation (4.34) is not uniquely solvable. It follows from $X = XP_l$ that $Y_{22} = 0$. Thus, the solution of (4.29) together with $X = XP_l$ is unique and given by

$$X = W^{-*}\begin{bmatrix} Y_{11} & 0 \\ 0 & 0 \end{bmatrix} W^{-1}, \tag{4.35}$$

where Y_{11} satisfies (4.31). □

From the proof of Theorem 4.13 it follows that the solution of the GCALE (4.29) is not unique, since equation (4.34) has many solutions. As usual for linear systems we may resolve the nonuniqueness of the solution by requiring extra conditions. This may be the solution of minimum norm, or we may choose the nonunique part Y_{22} to be zero. In terms of the original data the latter requirement is expressed as $X = XP_l$. In the following a system of matrix equations

$$\begin{aligned} &E^*XA + A^*XE = -P_r^*GP_r, \\ &X = XP_l \end{aligned} \tag{4.36}$$

will be called *projected generalized continuous-time algebraic Lyapunov equation.*

As a consequence of Corollary 4.7 and Theorem 4.13 we obtain generalizations of classical Lyapunov stability theorems [53, 99] for the projected GCALE (4.36).

Corollary 4.14. *Let $\lambda E - A$ be a regular pencil and let P_r and P_l be the spectral projections onto the right and left finite deflating subspaces of $\lambda E - A$. If there exist an Hermitian, positive definite matrix G and an Hermitian, positive semidefinite matrix X satisfying the projected GCALE* (4.36)*, then the pencil $\lambda E - A$ is c-stable.*

Proof. The result immediately follows from Corollary 4.7. □

Corollary 4.15. *Let $\lambda E - A$ be a regular pencil and let P_r and P_l be the spectral projections onto the right and left finite deflating subspaces of $\lambda E - A$, respectively. If the pencil $\lambda E - A$ is c-stable, then the projected GCALE* (4.36) *has a unique solution for every matrix G. This solution is given by*

$$X = \frac{1}{2\pi}\int_{-\infty}^{\infty}(i\omega E - A)^{-*}P_r^*GP_r(i\omega E - A)^{-1}d\omega. \tag{4.37}$$

If G is Hermitian, then the solution X is Hermitian. If G is positive (*semi*)*definite, then X is positive semidefinite.*

Proof. If $\lambda E - A$ is c-stable, then by Theorem 4.13 the projected GCALE (4.36) is uniquely solvable for every matrix G. The solution X is given by (4.35), where Y_{11} satisfies equation (4.31) and has the form

$$Y_{11} = \frac{1}{2\pi}\int_{-\infty}^{\infty}(i\omega I - J)^{-*}T_{11}(i\omega I - J)^{-1}d\omega.$$

Therefore, (4.37) holds. Clearly, if G is Hermitian and positive (semi)definite, then this solution X is Hermitian and positive semidefinite. □

Remark 4.16. Note that if $\lambda E - A$ is c-stable and if G is positive definite, then the solution X of the projected GCALE (4.36) is positive definite on the subspace $\operatorname{Im} P_l$ and the matrix E^*XE is positive definite on the subspace $\operatorname{Im} P_r$.

Remark 4.17. It follows from Corollaries 4.14 and 4.15 that if the projected GCALE (4.36) has an Hermitian, positive semidefinite solution for some Hermitian, positive definite matrix G, then (4.36) is uniquely solvable for every G.

Remark 4.18. The assertions of Corollaries 4.14, 4.15 and Remarks 4.16, 4.17 remain valid if the matrix G is positive definite only on the subspace $\operatorname{Im} P_r$.

In Table 4.1 we review the generalized continuous-time Lyapunov equations with different right-hand sides discussed above.

4.2.4 Inertia with respect to the imaginary axis

The projected GCALE (4.36) can be used to generalize some matrix inertia theorems [20, 29, 33, 37, 108, 122, 172] for matrix pencils.

First we recall the definition of an inertia with respect to the imaginary axis for matrices.

Definition 4.19. The *inertia of a matrix A with respect to the imaginary axis* (*c-inertia*) is defined by the triplet of integers

$$\operatorname{In}_c(A) = \{\ \pi_-(A),\ \pi_+(A),\ \pi_0(A)\ \},$$

where $\pi_-(A)$, $\pi_+(A)$ and $\pi_0(A)$ denote the numbers of eigenvalues of A with negative, positive and zero real part, respectively, counting multiplicities.

Taking into account that a matrix pencil may have finite as well as infinite eigenvalues, the c-inertia for matrices can be generalized for regular pencils as follows.

Definition 4.20. The *c-inertia of a regular matrix pencil $\lambda E - A$* is defined by the quadruple of integers

$$\operatorname{In}_c(E, A) = \{\ \pi_-(E, A),\ \pi_+(E, A),\ \pi_0(E, A),\ \pi_\infty(E, A)\ \},$$

where $\pi_-(E, A)$, $\pi_+(E, A)$ and $\pi_0(E, A)$ denote the numbers of the finite eigenvalues of $\lambda E - A$ counted with their algebraic multiplicities with negative, positive and zero real part, respectively, and $\pi_\infty(E, A)$ denotes the number of infinite eigenvalues of $\lambda E - A$.

Clearly, $\pi_-(E, A) + \pi_+(E, A) + \pi_0(E, A) + \pi_\infty(E, A) = n$ is the size of E and A. If the matrix E is nonsingular, then $\pi_\infty(E, A) = 0$ and $\pi_\varrho(E, A) = \pi_\varrho(AE^{-1}) = \pi_\varrho(E^{-1}A)$, where ϱ is $-$, $+$ and 0. A c-stable pencil $\lambda E - A$ has the c-inertia $\operatorname{In}_c(E, A) = \{\, n_f, 0, 0, n_\infty \,\}$, where n_f and n_∞ are the dimensions of the finite and infinite deflating subspaces of $\lambda E - A$.

The following theorems give connections between the c-inertia of the pencil $\lambda E - A$ and the c-inertia of the Hermitian solution X of the projected GCALE (4.36).

Theorem 4.21. *Let $\lambda E - A$ be a regular pencil and let G be an Hermitian, positive definite matrix. If the projected GCALE* (4.36) *has an Hermitian solution X, then*

$$\begin{array}{ll} \pi_-(E, A) = \pi_+(X), & \pi_+(E, A) = \pi_-(X), \\ \pi_0(E, A) = 0, & \pi_\infty(E, A) = \pi_0(X). \end{array} \tag{4.38}$$

right-hand side $-G$	$E^*XA + A^*XE = -G$	
	$X = X^* > 0$ on $\mathbb{F}^n$, unique	$X = X^* \geq 0$ on $\mathbb{F}^n$, unique
$G = G^* > 0$ on $\mathbb{F}^n$ E is nonsingular	$\Longleftrightarrow$ c-stable	$\Longleftrightarrow$ c-stable
$G = G^* \geq 0$ on $\mathbb{F}^n$ E is nonsingular		$\Longleftarrow$ c-stable
right-hand side $-G$	$E^*XA + A^*XE = -G$	
	$X = X^* > 0$ on $\operatorname{Im} P_l$	$X = X^* \geq 0$ on $\mathbb{F}^n$
$G = G^* > 0$ on $\operatorname{Im} P_r$	$\Longrightarrow$ c-stable	$\Longrightarrow$ c-stable
right-hand side $-E^*GE$	$E^*XA + A^*XE = -E^*GE$	
	$X = X^* > 0$ on $\operatorname{Im} P_l$	$X = X^* \geq 0$ on $\mathbb{F}^n$
$G = G^* > 0$ on $\mathbb{F}^n$	$\Longleftrightarrow$ c-stable index at most 2	$\Longleftrightarrow$ c-stable index at most 1
$G = G^* > 0$ on $\operatorname{Im} P_l$	$\Longrightarrow$ c-stable	$\Longrightarrow$ c-stable
	$\Longleftarrow$ c-stable index at most 2	$\Longleftarrow$ c-stable index at most 1
right-hand side $-P_r^*GP_r$	$E^*XA + A^*XE = -P_r^*GP_r, \quad X = XP_l$	
	$X = X^* > 0$ on $\operatorname{Im} P_l$, unique	$X = X^* \geq 0$ on $\mathbb{F}^n$, unique
$G = G^* > 0$ on $\mathbb{F}^n$	$\Longleftrightarrow$ c-stable	$\Longleftrightarrow$ c-stable
$G = G^* > 0$ on $\operatorname{Im} P_r$	$\Longleftrightarrow$ c-stable	$\Longleftrightarrow$ c-stable
$G = G^* \geq 0$ on $\mathbb{F}^n$		$\Longleftarrow$ c-stable

Table 4.1: Generalized continuous-time Lyapunov equations with different right-hand sides

Conversely, if $\pi_0(E,A)=0$*, then there exist an Hermitian matrix* X *and an Hermitian, positive definite matrix* G *such that the GCALE in* (4.36) *is fulfilled and the c-inertia identities* (4.38) *hold.*

Proof. Since the Hermitian solution X of the projected GCALE (4.36) has the form (4.35), where the Hermitian matrix Y_{11} satisfies the Lyapunov equation (4.31) with the Hermitian, positive definite matrix T_{11}, it follows from the Sylvester law of inertia [29] and the main inertia theorem [122, Theorem 1] that

$$\begin{aligned}
&\pi_-(E,A)=\pi_-(J)=\pi_+(Y_{11})=\pi_+(X),\\
&\pi_+(E,A)=\pi_+(J)=\pi_-(Y_{11})=\pi_-(X),\\
&\pi_0(E,A)=\pi_0(J)=\pi_0(Y_{11})=0,\\
&\pi_\infty(E,A)=\pi_\infty(E,A)+\pi_0(Y_{11})=\pi_0(X).
\end{aligned}$$

Assume now that $\pi_0(E,A)=0$. Then $\pi_0(J)=0$, and by the main inertia theorem [122, Theorem 1] there exists an Hermitian matrix Y_{11} such that $T_{11}=-(Y_{11}J+J^*Y_{11})$ is Hermitian, positive definite and

$$\pi_-(J)=\pi_+(Y_{11}),\qquad \pi_+(J)=\pi_-(Y_{11}),\qquad \pi_0(J)=\pi_0(Y_{11})=0.$$

In this case the Hermitian matrices

$$X=W^{-*}\begin{bmatrix} Y_{11} & 0\\ 0 & 0\end{bmatrix}W^{-1}\qquad\text{and}\qquad G=T^*\begin{bmatrix} T_{11} & 0\\ 0 & I\end{bmatrix}T$$

satisfy the GCALE in (4.36), G is positive definite and the c-inertia identities (4.38) hold. □

Consider now the case when the matrix G is Hermitian, positive semidefinite.

Theorem 4.22. *Let* $\lambda E-A$ *be a regular pencil and let* X *be an Hermitian solution of the projected GCALE* (4.36) *with an Hermitian, positive semidefinite matrix* G*.*

1. *If* $\pi_0(E,A)=0$*, then* $\pi_-(X)\le\pi_+(E,A)$ *and* $\pi_+(X)\le\pi_-(E,A)$*.*
2. *If* $\pi_0(X)=\pi_\infty(E,A)$*, then* $\pi_+(E,A)\le\pi_-(X)$ *and* $\pi_-(E,A)\le\pi_+(X)$*.*

Proof. The result immediately follows if we apply the matrix inertia theorems [33, Lemma 1 and Lemma 2] to equation (4.31). □

As a consequence of Theorem 4.22 we obtain a generalization of Theorem 4.21 for the case that G is Hermitian, positive semidefinite.

Corollary 4.23. *Let* $\lambda E-A$ *be a regular pencil and let* G *be an Hermitian, positive semidefinite matrix. Assume that the projected GCALE* (4.36) *has an Hermitian solution* X*. If* $\pi_0(E,A)=0$ *and* $\pi_0(X)=\pi_\infty(E,A)$*, then the c-inertia identities* (4.38) *hold.*

Similar to the matrix case [99, 108, 173], the c-inertia identities (4.38) can also be derived using observability conditions for the continuous-time descriptor system (3.1). The following corollary shows that in the case of an Hermitian, positive semidefinite matrix $G = C^*C$, the conditions $\pi_\infty(E, A) = \pi_0(X)$ and $\pi_0(E, A) = 0$ in Corollary 4.23 may be replaced by the assumption for the triplet (E, A, C) to be R-observable.

Corollary 4.24. *Let $\lambda E - A$ be a regular pencil. If there exists an Hermitian matrix X satisfying the projected GCALE*

$$E^*XA + A^*XE = -P_r^*C^*CP_r, \qquad X = XP_l, \tag{4.39}$$

and if the triplet (E, A, C) is R-observable, then the c-inertia identities (4.38) hold.

Proof. Let $\lambda E - A$ be in Weierstrass canonical form (2.2) and let $CT^{-1} = [\, C_1,\ C_2 \,]$ be partitioned in blocks conformally to E and A. Then the Hermitian solution of the projected GCALE (4.39) has the form (4.35), where Y_{11} satisfies the Lyapunov equation

$$Y_{11}J + J^*Y_{11} = -C_1^*C_1. \tag{4.40}$$

Since (E, A, C) is R-observable, by Theorem 3.27 the matrix $\begin{bmatrix} \lambda I - J \\ C_1 \end{bmatrix}$ has full column rank for all $\lambda \in \mathbb{C}$. In this case the solution X_{11} of (4.40) is nonsingular and the matrix J has no eigenvalues on the imaginary axis [99, Theorem 13.1.4]. Hence, $\pi_0(E, A) = 0$ and $\pi_0(X) = \pi_0(X_{11}) + \pi_\infty(E, A) = \pi_\infty(E, A)$. The remaining relations in (4.38) immediately follow from Corollary 4.23. □

The following corollary gives connections between c-stability of the pencil $\lambda E - A$, the R-observability of the triplet (E, A, C) and the existence of an Hermitian solution of the projected GCALE (4.39).

Corollary 4.25. *Consider the statements*

1. *the pencil $\lambda E - A$ is c-stable,*
2. *the triplet (E, A, C) is R-observable,*
3. *the projected GCALE (4.39) has a unique solution X which is Hermitian, positive definite on the subspace $\operatorname{Im} P_l$.*

Any two of these statements together imply the third.

Proof. '1 and 2 ⇒ 3' and '2 and 3 ⇒ 1' can be obtained from Corollaries 4.15 and 4.24.

'1 and 3 ⇒ 2'. Suppose that (E, A, C) is not R-observable. Then there exists $\lambda_0 \in \mathbb{C}$ and a vector $v \neq 0$ such that

$$\begin{bmatrix} \lambda_0 E - A \\ C \end{bmatrix} v = 0.$$

We obtain that v is an eigenvector of the pencil $\lambda E - A$ corresponding to the finite eigenvalue λ_0. Hence $\Re e\, \lambda_0 > 0$ and $v \in \operatorname{Im} P_r$. Moreover, we have $Cv = 0$. On the other hand, it follows from the Lyapunov equation in (4.39) that

$$-\|Cv\|^2 = v^*(E^*XA + A^*XE)v = 2(\Re e\, \lambda_0)v^*E^*XEv.$$

and, hence, $Cv \neq 0$. Thus, the triplet (E, A, C) is R-observable. □

Corollary 4.25 generalizes the stability result of Corollary 4.14 to the case that $G = C^*C$ is Hermitian, positive semidefinite. We see that weakening the assumption for G to be positive semidefinite requires the additional R-observability condition. Moreover, Corollary 4.25 gives necessary and sufficient conditions for (E, A, C) to be R-observable.

It is natural to ask what happens if the triplet (E, A, C) is not R-observable. Consider the proper observability matrix $\mathbf{O}_+$ as in (3.43). By Theorem 3.27 the triplet (E, A, C) is R-observable if and only if $\operatorname{rank}(\mathbf{O}_+) = n - \pi_\infty(E, A)$. Using the Weierstrass canonical form (2.2) and the matrix inertia theorem [108] we obtain the following c-inertia inequalities.

Theorem 4.26. *Let $\lambda E - A$ be a regular pencil and let X be an Hermitian solution of the projected GCALE* (4.39). *Assume that* $\operatorname{rank}(\mathbf{O}_+) < n - \pi_\infty(E, A)$. *Then*

$$\begin{aligned} |\pi_-(E, A) - \pi_+(X)| &\le n - \pi_\infty(E, A) - \operatorname{rank}(\mathbf{O}_+), \\ |\pi_+(E, A) - \pi_-(X)| &\le n - \pi_\infty(E, A) - \operatorname{rank}(\mathbf{O}_+). \end{aligned} \tag{4.41}$$

Proof. The result follows by applying the matrix inertia theorem from [108] to equation (4.31). □

Other matrix inertia theorems concerning the matrix c-inertia and the rank of the observability matrix [20, 140] can be generalized for matrix pencils in the same way.

Remark 4.27. By duality of controllability and observability conditions analogies of Corollaries 4.24, 4.25 and Theorem 4.26 can be proved for the dual projected GCALE

$$EXA^* + AXE^* = -P_l BB^* P_l^*, \qquad X = P_r X. \tag{4.42}$$

4.3 Generalized discrete-time Lyapunov equations

In this section we study the GDALE

$$A^*XA - E^*XE = -G, \tag{4.43}$$

where E, A, $G \in \mathbb{F}^{n,n}$ are given matrices and $X \in \mathbb{F}^{n,n}$ is unknown matrix.

4.3.1 General case

Consider a *discrete-time Lyapunov operator* $\mathcal{L}_d : \mathbb{F}^{n,n} \to \mathbb{F}^{n,n}$ of the form

$$\mathcal{L}_d(X) := A^*XA - E^*XE. \tag{4.44}$$

Similarly to the continuous-time case, the GDALE (4.43) can be written in the operator form $\mathcal{L}_d(X) = -G$ or as the linear system

$$\mathbf{L}_d x = -g, \tag{4.45}$$

where $x = \mathrm{vec}(X)$, $g = \mathrm{vec}(G)$ and the $n^2 \times n^2$-matrix

$$\mathbf{L}_d = A^T \otimes A^* - E^T \otimes E^* \tag{4.46}$$

is the matrix representation of the discrete-time Lyapunov operator $\mathcal{L}_d$, see [78]. Applying the theory of linear systems [53, 99] to (4.45), we obtain the following necessary and sufficient conditions for the GDALE (4.43) to be solvable and to have a unique solution.

Theorem 4.28. [99] *Let* $\mathbf{L}_d$ *be as in* (4.46) *and let* $x = \mathrm{vec}(X)$, $g = \mathrm{vec}(G)$. *The GDALE* (4.43) *has a solution if and only if* $\mathrm{rank}\,[\,\mathbf{L}_d, g\,] = \mathrm{rank}\,\mathbf{L}_d$. *There exists a unique solution of* (4.43) *if and only if the matrix* $\mathbf{L}_d$ *is nonsingular.*

The GDALE (4.43) is a special case of the generalized Sylvester equation (4.14) with $B = A^*$ and $F = E^*$. Then from Theorem 4.2 we have necessary and sufficient conditions for unique solvability of equation (4.43) in terms of the spectrum of the pencil $\lambda E - A$.

Theorem 4.29. [125] *Let* $\lambda E - A$ *be a regular pencil with eigenvalues* $\{\lambda_1, \ldots, \lambda_n\}$ *counted according to their multiplicities. The GDALE* (4.43) *has a unique solution for every matrix* G *if and only if at least one of the matrices* E *and* A *is nonsingular and* $\lambda_j \overline{\lambda}_k \neq 1$ *for all finite eigenvalues* λ_j *and* λ_k *of* $\lambda E - A$.

The GDALE (4.43) is said to be *regular* if it has a unique solution for every G. For the regular GDALE (4.43), the singularity of one of the matrices E and A implies the nonsingularity of the other and it follows from the condition $\lambda_j \overline{\lambda}_k \neq 1$ that the pencil $\lambda E - A$ has no eigenvalues on the unit circle. Unlike the continuous-time case, the GDALE (4.43) is called *non-degenerate* if at least one of the matrices E and A is nonsingular, and the GDALE (4.43) is called *degenerate* if both the matrices E and A are singular.

The non-degenerate GDALE (4.43) is equivalent to standard discrete-time Lyapunov equations

$$(AE^{-1})^* X A E^{-1} - X = -E^{-*} G E^{-1} \tag{4.47}$$

or

$$X - (EA^{-1})^* X E A^{-1} = -A^{-*} G A^{-1}. \tag{4.48}$$

Then the classical Lyapunov theorems [53] on the existence and uniqueness of positive definite solutions of (4.47) or (4.48) can be generalized to equation (4.43).

Theorem 4.30. *Let* $\lambda E - A$ *be a regular pencil. If all eigenvalues of* $\lambda E - A$ *are finite and lie inside the unit circle, then for every Hermitian, positive* (*semi*)*definite matrix* G, *the GDALE* (4.43) *has a unique Hermitian, positive* (*semi*)*definite solution* X. *Conversely, if there exist Hermitian, positive definite matrices* X *and* G *satisfying* (4.43), *then all eigenvalues of the pencil* $\lambda E - A$ *are finite and lie inside the unit circle.*

In contrast to the continuous-time case, the GDALE (4.43) with singular E and positive definite G has a unique negative definite solution X if and only if the matrix A is nonsingular and all eigenvalues of the pencil $\lambda E - A$ lie outside the unit circle or, equivalently, the eigenvalues of the reciprocal pencil $E - \mu A$ are finite and lie inside the unit circle. However, as the following example demonstrates, if both matrices E and A are singular, then the degenerate GDALE (4.43) may have no solutions although all finite eigenvalues of $\lambda E - A$ lie inside the unit circle.

Example 4.31. The GDALE (4.43) with

$$E = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \qquad G = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

is not solvable.

Even if a solution of the degenerate GCALE (4.43) exists, it is not unique. Indeed, if X satisfies the degenerate GDALE (4.43), then for any nonzero vectors $z \in \operatorname{Ker} E^*$ and $v \in \operatorname{Ker} A^*$, the matrix $X + zv^* + vz^*$ also satisfies (4.43).

Analogous to the continuous-time case, the GDALE (4.43) can be used to investigate the asymptotic solution behavior of system (3.29). The following theorem gives sufficient conditions for the pencil $\lambda E - A$ to be d-stable, that is, sufficient conditions for the trivial solution of (3.29) to be asymptotically stable.

Theorem 4.32. *Let P_l and P_r be the spectral projections onto the left and right finite deflating subspaces of a regular pencil $\lambda E - A$ and let G be a matrix that is Hermitian, positive definite on the subspace* $\operatorname{Im} P_r$. *If the GDALE* (4.43) *has a solution X which is Hermitian, positive definite on the subspace* $\operatorname{Im} P_l$, *then the pencil $\lambda E - A$ is d-stable.*

Proof. As in the continuous-time case we have that if X is positive definite on $\operatorname{Im} P_l$, then E^*XE is positive definite on $\operatorname{Im} P_r$. Let $v \neq 0$ be an eigenvector of the pencil $\lambda E - A$ corresponding to a finite eigenvalue λ, that is, $\lambda Ev = Av$ and $v \in \operatorname{Im} P_r$. Multiplying the GDALE (4.43) on the right and left by v and v^* we obtain from

$$\begin{aligned} -v^*Gv &= v^*(A^*XA - E^*XE)v = \lambda\overline{\lambda}\, v^*E^*XEv - v^*E^*XEv \\ &= (|\lambda|^2 - 1)v^*E^*XEv \end{aligned} \tag{4.49}$$

that $|\lambda| < 1$, i.e., all finite eigenvalues of the pencil $\lambda E - A$ lie inside the unit circle. □

It follows from (4.49) that the condition for X to be positive definite on $\operatorname{Im} P_l$ can be replaced by the assumption that X is positive semidefinite on $\mathbb{F}^n$.

Corollary 4.33. *Let P_r be the spectral projection onto the right finite deflating subspace of a regular pencil $\lambda E - A$ and let G be a matrix that is Hermitian, positive definite on* $\operatorname{Im} P_r$. *If the GDALE* (4.43) *has an Hermitian, positive semidefinite solution X, then the pencil $\lambda E - A$ is d-stable.*

Example 4.31 shows that d-stability of the pencil $\lambda E - A$ does not imply the existence of solutions of the degenerate GDALE (4.43) for every G.

It is well known that standard continuous-time and discrete-time Lyapunov equations are related via a *Cayley transformation* for matrices defined by $\mathfrak{C}(A) = (A - I)^{-1}(A + I)$, see, e.g., [99]. A *generalized Cayley transformation* for matrix pencils given by

$$\mathfrak{C}(E, A) = \lambda(A - E) - (E + A) \tag{4.50}$$

allows us to state a similar connection between generalized Lyapunov equations in continuous-time and discrete-time cases [118]. Indeed, X is a solution of the GDALE (4.43) if and only if X satisfies the GCALE

$$\mathcal{E}^* X \mathcal{A} + \mathcal{A}^* X \mathcal{E} = -2G, \tag{4.51}$$

where $\lambda\mathcal{E} - \mathcal{A} = \lambda(A - E) - (E + A)$ is the Cayley-transformed pencil.

The following theorem gives a relationship between the eigenvalues of the pencils $\lambda E - A$ and $\lambda\mathcal{E} - \mathcal{A}$, see [118] for details.

Proposition 4.34. *1. Consider the generalized Cayley transformation* (4.50) *for the $\lambda E - A$ associated with the GCALE* (4.9). *Then*

- *(a) the finite eigenvalues of $\lambda E - A$ in the open left and right half-plane are mapped to eigenvalues inside and outside the unit circle, respectively, and the eigenvalue $\lambda = 1$ is mapped to ∞;*
- *(b) the finite eigenvalues on the imaginary axis are mapped to eigenvalues on the unit circle;*
- *(c) the infinite eigenvalues of $\lambda E - A$ are mapped to $\lambda = 1$.*

2. Consider the generalized Cayley transformation (4.50) *for the pencil $\lambda E - A$ associated with the GDALE* (4.43). *Then*

- *(a) the finite eigenvalues of $\lambda E - A$ inside and outside the unit circle are mapped to eigenvalues in the open left and right half-plane, respectively;*
- *(b) the finite eigenvalues on the unit circle except $\lambda = 1$ are mapped to eigenvalues on the imaginary axis and the eigenvalue $\lambda = 1$ is mapped to ∞;*
- *(c) the infinite eigenvalues of $\lambda E - A$ are mapped to $\lambda = 1$.*

Thus, in the case of a nonsingular matrix E we obtain from Proposition 4.34 that the matrix pencil $\lambda E - A$ is c-stable (d-stable) if and only if the Cayley-transformed pencil $\lambda\mathcal{E} - \mathcal{A}$ is d-stable (c-stable). However, if E is singular, then this assertion does not hold any more, since infinite eigenvalues of a c-stable pencil are mapped under the generalized Cayley transformation to an eigenvalue $\lambda = 1$ on the unit circle and infinite eigenvalues of a d-stable pencil are mapped to an eigenvalue $\lambda = 1$ in the right half-plane. Therefore, we consider the generalized Lyapunov equations in the continuous-time and discrete-time case separately.

4.3.2 Special right-hand side: index 1 case

Analogous to the continuous-time case, we consider the generalized discrete-time Lyapunov equation

$$A^*XA - E^*XE = -E^*GE. \tag{4.52}$$

Such an equation has been considered previously in [116, 153]. The following theorem gives sufficient conditions for the existence of solutions of the GDALE (4.52), where both the matrices E and A are singular.

Theorem 4.35. *Let $\lambda E - A$ be a d-stable pencil. If $\lambda E - A$ is of index one or if the zero eigenvalue of $\lambda E - A$ is simple, then for every matrix G, the degenerate GDALE* (4.52) *has a solution X. If G is Hermitian, then* (4.52) *has an Hermitian solution.*

Proof. We may assume without loss of generality that the pencil $\lambda E - A$ is in Weierstrass canonical form

$$E = W \begin{bmatrix} I & & \\ & I & \\ & & N \end{bmatrix} T, \qquad A = W \begin{bmatrix} J_1 & & \\ & J_2 & \\ & & I \end{bmatrix} T,$$

where the matrix J_1 is nonsingular with all eigenvalues inside the unit circle and the matrix J_2 has zero eigenvalues only. Let the matrices

$$W^*GW = \begin{bmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \\ W_{31} & W_{32} & W_{33} \end{bmatrix} \quad \text{and} \quad W^*XW = \begin{bmatrix} Y_{11} & Y_{12} & Y_{13} \\ Y_{21} & Y_{22} & Y_{23} \\ Y_{31} & Y_{32} & Y_{33} \end{bmatrix} \tag{4.53}$$

be partitioned in blocks conformally to E and A. Then from (4.52) we have

$$J_i^* Y_{ij} J_j - Y_{ij} = -W_{ij}, \qquad i,j = 1,2, \tag{4.54}$$
$$J_i^* Y_{i3} - Y_{i3} N = -W_{i3} N, \qquad i = 1,2, \tag{4.55}$$
$$Y_{3j} J_j - N^* Y_{3j} = -N^* W_{3j}, \qquad j = 1,2, \tag{4.56}$$
$$Y_{33} - N^* Y_{33} N = -N^* W_{33} N. \tag{4.57}$$

Since all eigenvalues of J_1 lie inside the unit circle and J_2, N are nilpotent, the standard Lyapunov equations (4.54) and (4.57) have unique solutions for every right-hand side, see [99]. Equations (4.55) with $i = 1$ and (4.56) with $j = 1$ are uniquely solvable for every W_{13} and W_{31}, since J_1 and N have no common eigenvalues. Moreover, if $W_{31} = W_{13}^*$, then $Y_{31} = Y_{13}^*$.

Consider equations (4.55) with $i = 2$ and (4.56) with $j = 2$. If the index of $\lambda E - A$ is one, i.e., $N = 0$, then these equations have trivial solutions for every W_{23} and W_{32}. If the zero eigenvalues of $\lambda E - A$ are simple, i.e., $J_2 = 0$, then these equations have solutions $Y_{23} = W_{23}$ and $Y_{32} = W_{32}$, respectively. Clearly, if G is Hermitian, then the GDALE (4.52) has an Hermitian solution. □

Note that if the index of the pencil $\lambda E - A$ is larger than one and if $\lambda E - A$ has a zero eigenvalue which is not simple, then as the following example shows, the GDALE (4.52) may have no solutions.

Example 4.36. For $X = [x_{ij}]_{i,j=1}^4$, $G = [g_{ij}]_{i,j=1}^4$ and

$$E = \begin{bmatrix} I_2 & 0 \\ 0 & N_2 \end{bmatrix}, \qquad A = \begin{bmatrix} N_2 & 0 \\ 0 & I_2 \end{bmatrix},$$

we have

$$\begin{aligned} A^*XA - E^*XE &= \begin{bmatrix} -x_{11} & -x_{12} & 0 & -x_{13} \\ -x_{21} & x_{11}-x_{22} & x_{13} & x_{14}-x_{23} \\ 0 & x_{31} & x_{33} & x_{34} \\ -x_{31} & x_{41}-x_{32} & x_{43} & x_{44}-x_{33} \end{bmatrix} \\ &= -E^*GE = -\begin{bmatrix} g_{11} & g_{12} & 0 & g_{13} \\ g_{21} & g_{22} & 0 & g_{23} \\ 0 & 0 & 0 & 0 \\ g_{31} & g_{32} & 0 & g_{33} \end{bmatrix}. \end{aligned}$$

If $g_{13} \neq 0$ or $g_{31} \neq 0$, then this equation has no solution.

The following theorem gives necessary and sufficient conditions for the GDALE (4.52) to have an Hermitian, positive semidefinite solution.

Theorem 4.37. *Let $\lambda E - A$ be a regular pencil and let G be an Hermitian, positive definite matrix. The GDALE* (4.52) *has an Hermitian, positive semidefinite solution X if and only if the pencil $\lambda E - A$ is of index at most one and it is d-stable.*

Proof. From the proof of Theorem 4.35 we have that if the d-stable pencil $\lambda E - A$ is of index at most one, then the matrix

$$X = W^{-*} \begin{bmatrix} Y_{11} & Y_{12} & 0 \\ Y_{21} & Y_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} W^{-1}$$

satisfies the GDALE (4.52), where

$$\begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} = \sum_{j=0}^{\infty} \begin{bmatrix} J_1^* & 0 \\ 0 & J_2^* \end{bmatrix}^j \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix}^j.$$

Since G is Hermitian and positive definite, X is Hermitian and positive semidefinite.

Conversely, assume that an Hermitian, positive definite matrix G and an Hermitian, positive semidefinite matrix X satisfy equation (4.52). Let G and X be as in (4.53). Then Y_{33} is an Hermitian, positive semidefinite solution of equation (4.57). This solution is given by

$$Y_{33} = -\sum_{j=1}^{\nu-1} (N^*)^j W_{33} N^j,$$

where ν is the index of the pencil $\lambda E - A$. For every nonzero vector v, we have

$$0 \leq v^* Y_{33} v = -\sum_{j=1}^{\nu-1} v^* (N^*)^j W_{33} N^j v \leq 0.$$

Hence, $v^* N^* W_{33} N v = 0$. Since W_{33} is Hermitian and positive definite, we obtain that $N = 0$.

Since G is Hermitian, positive definite, the matrix $E^* G E$ is Hermitian, positive definite on $\operatorname{Im} P_r$. Then by Corollary 4.33 the pencil $\lambda E - A$ is d-stable. □

Remark 4.38. If the d-stable pencil $\lambda E - A$ is of index at most one and if G is Hermitian, positive definite on $\mathbb{F}^n$ (or positive definite only on $\operatorname{Im} P_l$), then equation (4.52) has a (nonunique) Hermitian solution which is positive definite on $\operatorname{Im} P_l$ and positive semidefinite on $\mathbb{F}^n$. In this case for all solutions X of (4.52), the matrix $E^* X E$ is unique, positive definite on $\operatorname{Im} P_r$ and positive semidefinite on $\mathbb{F}^n$.

In Table 4.2 we review the generalized discrete-time Lyapunov equations discussed above.

right-hand side $-G$	$A^*XA - E^*XE = -G$	
	$X = X^* > 0$ on $\mathbb{F}^n$, unique	$X = X^* \geq 0$ on $\mathbb{F}^n$, unique
$G = G^* > 0$ on $\mathbb{F}^n$ E is nonsingular	$\Longleftrightarrow$ d-stable	$\Longleftrightarrow$ d-stable
$G = G^* \geq 0$ on $\mathbb{F}^n$ E is nonsingular		$\Longleftarrow$ d-stable

right-hand side $-G$	$A^*XA - E^*XE = -G$	
	$X = X^* > 0$ on $\operatorname{Im} P_l$	$X = X^* \geq 0$ on $\mathbb{F}^n$
$G = G^* > 0$ on $\operatorname{Im} P_r$	$\Longrightarrow$ d-stable	$\Longrightarrow$ d-stable

right-hand side $-E^*GE$	$A^*XA - E^*XE = -E^*GE$	
	$X = X^* > 0$ on $\operatorname{Im} P_l$	$X = X^* \geq 0$ on $\mathbb{F}^n$
$G = G^* > 0$ on $\mathbb{F}^n$	$\Longleftarrow$ d-stable index at most 1	$\Longleftrightarrow$ d-stable index at most 1
	$\Longrightarrow$ d-stable	
$G = G^* > 0$ on $\operatorname{Im} P_l$	$\Longleftarrow$ d-stable index at most 1	$\Longleftarrow$ d-stable index at most 1
	$\Longrightarrow$ d-stable	$\Longrightarrow$ d-stable

Table 4.2: Generalized discrete-time Lyapunov equations with different right-hand sides

4.3.3 Projected discrete-time Lyapunov equations

Consider now the generalized discrete-time Lyapunov equation

$$A^*XA - E^*XE = -P_r^*GP_r + \xi(I - P_r)^*G(I - P_r), \tag{4.58}$$

where $\xi = -1$, 0 or 1. Note that, unlike the GCALE (4.29), equation (4.58) has two terms on the right-hand side. Which sign ξ is used, as we will see later, depends on the applications. We will study all three cases simultaneously. The following theorem gives necessary and sufficient condition for the existence of solutions of the GDALE (4.58).

Theorem 4.39. *Let $\lambda E - A$ be a regular pencil with finite eigenvalues $\{\lambda_1, \dots \lambda_{n_f}\}$ counted according to their multiplicities and let P_r and P_l be the spectral projections onto the right and left finite deflating subspaces of $\lambda E - A$. The GDALE* (4.58) *has a solution for every matrix G if and only if $\lambda_j \overline{\lambda}_k \neq 1$ for all $j, k = 1, \dots, n_f$. Moreover, if a solution of* (4.58) *satisfies $P_l^* X = XP_l$, then it is unique.*

Proof. Let the pencil $\lambda E - A$ be in Weierstrass canonical form (2.2), where J has eigenvalues $\{\lambda_1, \dots \lambda_{n_f}\}$. Substituting the matrices G and X be as in (4.30) in the GDALE (4.58), we obtain the system of matrix equations

$$J^*Y_{11}J - Y_{11} = -T_{11}, \tag{4.59}$$
$$J^*Y_{12} - Y_{12}N = 0, \tag{4.60}$$
$$Y_{21}J - N^*Y_{21} = 0, \tag{4.61}$$
$$Y_{22} - N^*Y_{22}N = \xi T_{22}. \tag{4.62}$$

The Lyapunov equation (4.59) has a solution for every T_{11} if and only if $\lambda_j \overline{\lambda}_k \neq 1$ for any two eigenvalues λ_j and λ_k of J, see [99]. Since N is nilpotent, equation (4.62) has a unique solution for every T_{22} [99]. Equations (4.60) and (4.61) are solvable and have, for example, trivial solutions. It follows from $P_l^* X = XP_l$ that

$$P_l^* X = W^{-*} \begin{bmatrix} Y_{11} & Y_{12} \\ 0 & 0 \end{bmatrix} W^{-1} = XP_l = W^{-*} \begin{bmatrix} Y_{11} & 0 \\ Y_{21} & 0 \end{bmatrix} W^{-1},$$

i.e., $Y_{12} = Y_{21} = 0$. Thus, the matrix

$$X = W^{-*} \begin{bmatrix} Y_{11} & 0 \\ 0 & Y_{22} \end{bmatrix} W^{-1} \tag{4.63}$$

is the unique Hermitian solution of the GDALE (4.58) together with $P_l^* X = XP_l$. □

If the GDALE (4.58) is solvable and if A is nonsingular, then the solution of (4.58) is unique. If both the matrices E and A are singular, then the nonuniqueness of the solution of (4.58) is resolved by requiring the extra condition for the nonunique part Y_{12} to be zero.

In terms of the original data this requirement is written as $P_l^* X = X P_l$. In the following a system of matrix equations of the form

$$\begin{aligned} &A^* X A - E^* X E = -P_r^* G P_r + \xi (I - P_r)^* G (I - P_r), \\ &P_l^* X = X P_l, \end{aligned} \tag{4.64}$$

is called *projected generalized discrete-time algebraic Lyapunov equation.*

Analogous to Corollaries 4.14 and 4.15 we can prove the following stability result for the projected GDALE (4.64).

Corollary 4.40. *Let $\lambda E - A$ be a regular pencil and let P_r and P_l be as in* (2.3)*. For every Hermitian, positive definite matrix G, the projected GDALE* (4.64) *has a unique Hermitian solution X which is positive definite on* $\operatorname{Im} P_l$ *if and only if the pencil $\lambda E - A$ is d-stable. This solution is given by*

$$X = \frac{1}{2\pi} \int_0^{2\pi} (e^{i\varphi} E - A)^{-*} \Big(P_r^* G P_r + \xi (I - P_r)^* G (I - P_r) \Big) (e^{i\varphi} E - A)^{-1} d\varphi. \tag{4.65}$$

Proof. Let the pencil $\lambda E - A$ be in Weierstrass canonical form (2.2) and let the matrices G and X as in (4.30) satisfy the projected GDALE (4.64). Since the matrix X is positive definite on the subspace $\operatorname{Im} P_l$ and $P_r^* G P_r - \xi (I - P_r)^* G (I - P_r)$ for $\xi = -1, 0, 1$ is positive definite on $\operatorname{Im} P_r$, by Theorem 4.32 the pencil $\lambda E - A$ is d-stable.

Assume now that $\lambda E - A$ is d-stable. Then by Theorem 4.39 the projected GDALE (4.64) has a unique solution for every G. This solution X is given by (4.63), where Y_{11} and Y_{22} satisfy equations (4.59) and (4.62), respectively. The solutions of (4.59) and (4.62) are represented as

$$Y_{11} = \frac{1}{2\pi} \int_0^{2\pi} (e^{i\varphi} I - J)^{-*} T_{11} (e^{i\varphi} I - J)^{-1} d\varphi$$

and

$$Y_{22} = \frac{\xi}{2\pi} \int_0^{2\pi} (e^{i\varphi} N - I)^{-*} T_{22} (e^{i\varphi} N - I)^{-1} d\varphi,$$

see [62]. Thus, (4.65) holds. Clearly, if G is Hermitian, positive definite, then Y_{11} and Y_{22} are Hermitian, and Y_{11} is positive definite. In this case the solution of (4.64) is positive definite on $\operatorname{Im} P_l$. □

Remark 4.41. Note that Corollary 4.40 remains valid if the matrix G is positive definite only on the subspace $\operatorname{Im} P_r$.

Remark 4.42. Assume that the pencil $\lambda E - A$ is d-stable and G is positive definite. Then the solution X of the projected GDALE (4.64) with $\xi = -1$ is positive definite on $\operatorname{Im} P_l$ and negative definite on $\operatorname{Ker} P_l$. For $\xi = 0$, the solution of (4.64) is positive definite on $\operatorname{Im} P_l$ and positive semidefinite on $\mathbb{F}^n$. If $\xi = 1$, then the solution of the projected GDALE (4.64) is positive definite on $\mathbb{F}^n$.

In Table 4.3 we review the projected GDALEs discussed above.

right-hand side $-P_r^*GP_r$	$A^*XA - E^*XE = -P_r^*GP_r, \quad P_l^*X = XP_l$	
	$X = X^* > 0$ on $\operatorname{Im} P_l$ unique	$X = X^* \geq 0$ on $\mathbb{F}^n$ unique
$G = G^* > 0$ on $\mathbb{F}^n$	$\Longleftrightarrow$ d-stable	$\Longleftrightarrow$ d-stable
$G = G^* > 0$ on $\operatorname{Im} P_r$	$\Longleftrightarrow$ d-stable	$\Longleftrightarrow$ d-stable
$G = G^* \geq 0$ on $\mathbb{F}^n$		$\Longleftarrow$ d-stable
right-hand side $-P_r^*GP_r - (I-P_r)^*G(I-P_r)$	$A^*XA - E^*XE = -P_r^*GP_r - (I-P_r)^*G(I-P_r)$ $P_l^*X = XP_l$	
	$X = X^* > 0$ on $\operatorname{Im} P_l$ unique	$X = X^* < 0$ on $\operatorname{Ker} P_l$ unique
$G = G^* > 0$ on $\mathbb{F}^n$	$\Longleftrightarrow$ d-stable	$\Longleftarrow$ d-stable
$G = G^* > 0$ on $\operatorname{Im} P_r$	$\Longleftrightarrow$ d-stable	
right-hand side $-P_r^*GP_r + (I-P_r)^*G(I-P_r)$	$A^*XA - E^*XE = -P_r^*GP_r + (I-P_r)^*G(I-P_r)$ $P_l^*X = XP_l$	
	$X = X^* > 0$ on $\operatorname{Im} P_l$ unique	$X = X^* > 0$ on $\mathbb{F}^n$ unique
$G = G^* > 0$ on $\mathbb{F}^n$	$\Longleftrightarrow$ d-stable	$\Longleftrightarrow$ d-stable
$G = G^* > 0$ on $\operatorname{Im} P_r$	$\Longleftrightarrow$ d-stable	$\Longrightarrow$ d-stable

Table 4.3: Projected generalized discrete-time Lyapunov equations with different right-hand sides

4.3.4 Inertia with respect to the unit circle

We recall that the *inertia of a matrix A with respect to the unit circle* (*d-inertia*) is defined by the triplet of integers

$$\operatorname{In}_d(A) = \{\ \pi_{<1}(A),\ \pi_{>1}(A),\ \pi_1(A)\ \},$$

where $\pi_{<1}(E,A)$, $\pi_{>1}(E,A)$ and $\pi_1(E,A)$ denote the numbers of the eigenvalues of A counted with their algebraic multiplicities inside, outside and on the unit circle, respectively.

Before extending the d-inertia for matrix pencils, it should be noted that in some problems it is necessary to distinguish the finite eigenvalues of a matrix pencil of modulus larger than 1 and the infinite eigenvalues although the latter also lie outside the unit circle. As we have seen in Section 3.2.2, the presence of infinite eigenvalues of $\lambda E - A$, in contrast to the finite eigenvalues outside the unit circle, does not affect the behavior at infinity of solutions of the discrete-time descriptor system (3.29).

Definition 4.43. The *d-inertia* of a regular pencil $\lambda E - A$ is defined by the quadruple of integers

$$\mathrm{In}_d(E, A) = \{\ \pi_{<1}(E, A),\ \pi_{>1}(E, A),\ \pi_1(E, A),\ \pi_\infty(E, A)\ \},$$

where $\pi_{<1}(E, A)$, $\pi_{>1}(E, A)$ and $\pi_1(E, A)$ denote the numbers of the finite eigenvalues of $\lambda E - A$ counted with their algebraic multiplicities inside, outside and on the unit circle, respectively, and $\pi_\infty(E, A)$ denotes the number of infinite eigenvalues of $\lambda E - A$.

If E is nonsingular, then $\pi_\infty(E, A) = 0$. A d-stable pencil $\lambda E - A$ has the d-inertia $\mathrm{In}_d(E, A) = \{ n_f, 0, 0, n_\infty \}$.

Although there is a difference between the discrete-time and continuous-time generalized Lyapunov equations, inertia theorems in the discrete-time case in many aspects resemble the continuous-time case. Thus, to avoid repetition, results for the d-inertia are only listed without proof unless necessary.

The following theorem gives a connection between the d-inertia of the pencil $\lambda E - A$ and the c-inertia of the Hermitian solution of the projected GDALE

$$\begin{aligned} &A^*XA - E^*XE = -P_r^*GP_r + (I - P_r)^*G(I - P_r),\\ &P_l^*X = XP_l. \end{aligned} \tag{4.66}$$

Theorem 4.44. *Let $\lambda E - A$ be a regular pencil. If there exists an Hermitian matrix X that satisfies the projected GDALE* (4.66) *with Hermitian, positive definite G, then*

$$\pi_{<1}(E, A) + \pi_\infty(E, A) = \pi_+(X), \quad \pi_{>1}(E, A) = \pi_-(X), \quad \pi_1(E, A) = \pi_0(X) = 0. \tag{4.67}$$

Conversely, if $\pi_1(E, A) = 0$, then there exist an Hermitian matrix X and an Hermitian, positive definite matrix G such that the GDALE in (4.66) *is satisfied and the inertia identities* (4.67) *hold.*

Proof. Every Hermitian solution X of (4.66) has the form (4.63), where the Hermitian matrix Y_{11} satisfies the Lyapunov equation (4.59) and the Hermitian matrix Y_{22} is a unique solution of the Lyapunov equation (4.62) with $\xi = 1$ that is given by

$$Y_{22} = \sum_{j=0}^{\nu-1} (N^*)^j T_{22} N^j.$$

If T_{22} is positive definite, then Y_{22} is also positive definite.

It follows from from the Sylvester law of inertia [29] and the matrix inertia theorem [172] that

$$\begin{aligned} &\pi_{<1}(E, A) = \pi_{<1}(J) = \pi_+(X_{11}) = \pi_+(X) - \pi_+(X_{22}) = \pi_+(X) - \pi_\infty(E, A),\\ &\pi_{>1}(E, A) = \pi_{>1}(J) = \pi_-(X_{11}) = \pi_-(X) - \pi_-(X_{22}) = \pi_-(X),\\ &\pi_1(E, A) = \pi_1(J) = \pi_0(X_{11}) = 0. \end{aligned}$$

Moreover, $\pi_0(X) = \pi_0(X_{11}) + \pi_0(X_{22}) = 0$.

Suppose that $\pi_1(E, A) = 0$. Then by the matrix inertia theorem [172] there exists Hermitian matrices X_{11}, X_{22} and Hermitian, positive definite matrices T_{11}, T_{22} such that (4.59) and (4.62) with $\xi = 1$ are satisfied and

$$\begin{array}{lll} \pi_{<1}(J) = \pi_+(X_{11}), & \pi_{>1}(J) = \pi_-(X_{11}), & \pi_1(J) = \pi_0(X_{11}) = 0, \\ \pi_\infty(E, A) = \pi_+(X_{22}), & \pi_-(X_{22}) = \pi_0(X_{22}) = 0. & \end{array}$$

Thus, the Hermitian matrices

$$X = W^{-*} \begin{bmatrix} X_{11} & 0 \\ 0 & X_{22} \end{bmatrix} W^{-1}, \qquad G = T^* \begin{bmatrix} T_{11} & 0 \\ 0 & T_{22} \end{bmatrix} T$$

satisfy the GDALE in (4.66), G is positive definite and the d-inertia identities (4.67) hold. □

There are also unit circle analogies of Theorem 4.22 and Corollary 4.23 that can be established in the same way.

Theorem 4.45. *Let $\lambda E - A$ be a regular pencil and let X be an Hermitian matrix that satisfy the projected GDALE* (4.66) *with Hermitian, positive semidefinite G.*

1. *If $\pi_1(E, A) = 0$, then $\pi_+(X) \le \pi_{<1}(E, A) + \pi_\infty(E, A)$ and $\pi_-(X) \le \pi_{>1}(E, A)$.*
2. *If $\pi_0(X) = 0$, then $\pi_+(X) \ge \pi_{<1}(E, A) + \pi_\infty(E, A)$ and $\pi_-(X) \ge \pi_{>1}(E, A)$.*

Corollary 4.46. *Let $\lambda E - A$ be regular and let G be an Hermitian, positive semidefinite. Assume that $\pi_1(E, A) = 0$. If there exists a nonsingular Hermitian matrix X that satisfies the projected GDALE* (4.66)*, then the inertia identities* (4.67) *hold.*

The inertia identities (4.67) can also be obtained from observability conditions for the discrete-time descriptor system (3.2). Consider the projected GDALE

$$\begin{aligned} &A^*XA - E^*XE = -P_r^*C^*CP_r + (I - P_r)^*C^*C(I - P_r), \\ &P_l^*X = XP_l. \end{aligned} \tag{4.68}$$

The presence of the second term in the right-hand side of the GDALE in (4.68) makes it possible to characterize not only R-observability but also S-observability and C-observability properties of the discrete-time descriptor system (3.2). We will show that the condition for the pencil $\lambda E - A$ to have no eigenvalues of modulus 1 and the condition for the solution of (4.68) to be nonsingular together are equivalent to the property for the triplet (E, A, C) to be C-observable.

Theorem 4.47. *Consider the discrete-time descriptor system* (3.2) *with a regular pencil $\lambda E - A$. Let X be an Hermitian solution of the projected GDALE* (4.68)*. The triplet (E, A, C) is C-observable if and only if $\pi_1(E, A) = 0$ and X is nonsingular.*

Proof. Let the pencil $\lambda E - A$ be in Weierstrass canonical form (2.2) and let the matrix $CT^{-1} = [\,C_1,\ C_2\,]$ be partitioned conformally to E and A. The solution of the projected GDALE (4.68) has the form (4.63), where Y_{11} is the solution of the Lyapunov equation

$$J^*Y_{11}J - Y_{11} = -C_1^*C_1 \tag{4.69}$$

and Y_{22} is the solution of the Lyapunov equation

$$Y_{22} - N^*Y_{22}N = C_2^*C_2. \tag{4.70}$$

Since (E, A, C) is C-observable, by Theorem 3.27 the matrices $\begin{bmatrix} \lambda I - J \\ C_1 \end{bmatrix}$ and $\begin{bmatrix} \lambda N - I \\ C_2 \end{bmatrix}$ have full column rank for all $\lambda \in \mathbb{C}$. In this case J has no eigenvalues on the unit circle and the solutions Y_{11} and Y_{22} of (4.69) and (4.69) are nonsingular [99, Theorem 13.2.4]. Thus, $\pi_1(E, A) = 0$ and the solution X of the projected GDALE (4.68) is nonsingular.

Conversely, let $v \in \operatorname{Im} P_r$ be a right eigenvector of $\lambda E - A$ corresponding to a finite eigenvalue λ with $|\lambda| \neq 1$. We have

$$-\|Cv\|^2 = -v^*C^*Cv = v^*(A^*XA - E^*XE)v = (|\lambda|^2 - 1)v^*E^*XEv.$$

Since X is nonsingular, $Ev \neq 0$ and $\pi_1(E, A) = 0$, we obtain that $Cv \neq 0$. Hence, the triplet (E, A, C) is R-observable. For $v \in \operatorname{Ker} E$, we have $\|Cv\|^2 = v^*C^*Cv = v^*A^*XAv \neq 0$, i.e, the triplet (E, A, C) is C-observable. □

It follows from Theorem 4.47 that if $\pi_1(E, A) = 0$ and an Hermitian solution X of (4.68) is nonsingular, then the triplet (E, A, C) is S-observable. However, S-observability of (E, A, C) does not imply that the solution of (4.68) is nonsingular.

Example 4.48. The projected GDALE (4.68) with

$$E = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \qquad C = [\,1,\ 0\,]$$

has the unique solution

$$X = \begin{bmatrix} -1/3 & 0 \\ 0 & 0 \end{bmatrix}$$

which is singular although $\operatorname{rank}\begin{bmatrix} \lambda E - A \\ C \end{bmatrix} = 2$ and $\operatorname{rank}\begin{bmatrix} E \\ K_{E^*}^*A \\ C \end{bmatrix} = 2$.

As an immediate consequence of Corollary 4.46 and Theorem 4.47 we obtain the following results.

Corollary 4.49. *Consider system* (3.2) *with a regular pencil $\lambda E - A$. Let the triplet (E, A, C) be C-observable. If an Hermitian matrix X satisfies the projected GDALE* (4.68)*, then the inertia identities* (4.67) *hold.*

Furthermore, from Theorem 4.47 and Corollary 4.49 we have the following connection between d-stability of the pencil $\lambda E - A$, the C-observability of the triplet (E, A, C) and the existence of an Hermitian solution of the projected GDALE (4.68).

Corollary 4.50. *Consider the statements:*

1. *the pencil $\lambda E - A$ is d-stable,*
2. *the triplet (E, A, C) is C-observable,*
3. *the projected GDALE (4.68) has a unique Hermitian, positive definite solution X.*

Any two of these statements together imply the third.

Remark 4.51. Note that Corollary 4.50 still holds if we replace the C-observability condition by the weaker condition for (E, A, C) to be R-observable, and if we require for solutions of (4.68) to be positive definite on $\operatorname{Im} P_l$ only.

If the triple (E, A, C) is not C-observable, then we can derive inertia inequalities similar to (4.41). Consider a proper observability matrix $\mathbf{O}_+$ and an improper observability matrix $\mathbf{O}_-$ as in as in (3.43). By Theorem 3.27 the triplet (E, A, C) is C-observable if and only if $\operatorname{rank}(\mathbf{O}_+) = n - \pi_\infty(E, A)$ and $\operatorname{rank}(\mathbf{O}_-) = \pi_\infty(E, A)$. Using the Weierstrass canonical form (2.2) and representation (4.63) for the solution X of the projected GDALE (4.68) we obtain the following inertia inequalities.

Theorem 4.52. *Let $\lambda E - A$ be a regular pencil and let X be an Hermitian solution of the projected GDALE (4.68). Then*

$$\begin{aligned} |\pi_{<1}(E, A) - \pi_+(X) + \operatorname{rank}(\mathbf{O}_-)| &\leq n - \pi_\infty(E, A) - \operatorname{rank}(\mathbf{O}_+), \\ |\pi_{>1}(E, A) - \pi_-(X)| &\leq n - \pi_\infty(E, A) - \operatorname{rank}(\mathbf{O}_+). \end{aligned}$$

Remark 4.53. All results of this subsection can also be reformulated for the projected generalized discrete-time Lyapunov equation

$$\begin{aligned} &A^* X A - E^* X E = -P_r^* G P_r + \xi (I - P_r)^* G (I - P_r), \\ &P_l^* X = X P_l \end{aligned}$$

with $\xi = 0$ or -1. For these equations we must consider instead of (4.67) the inertia identities

$$\pi_{<1}(E, A) = \pi_+(X), \quad \pi_{>1}(E, A) = \pi_-(X), \quad \pi_1(E, A) = 0, \quad \pi_\infty(E, A) = \pi_0(X)$$

for the case $\xi = 0$ and

$$\pi_{<1}(E, A) = \pi_+(X), \quad \pi_{>1}(E, A) + \pi_\infty(E, A) = \pi_-(X), \quad \pi_1(E, A) = \pi_0(X) = 0$$

for the case $\xi = -1$.

By duality of controllability and observability conditions, analogies of Theorems 4.47, 4.52 and Corollaries 4.49, 4.50 can be obtained for the dual projected GDALE

$$\begin{aligned} &A X A^* - E X E^* = -P_l B B^* P_l^* + \xi (I - P_l) B B^* (I - P_l)^*, \\ &P_r X = X P_r^*. \end{aligned} \tag{4.71}$$

4.4 Controllability and observability Gramians

In this section we establish relationships among solutions of projected generalized Lyapunov equations and the controllability and observability Gramians for descriptor systems introduced in [11]. Since the results for the continuous-time case are partly related to the discrete-time case, we begin our discussions with the discrete-time problem.

4.4.1 The discrete-time case

Consider the causal and noncausal controllability matrices $\mathbf{C}_+$ and $\mathbf{C}_-$ defined in (3.39) and the causal and noncausal observability matrices $\mathbf{O}_+$ and $\mathbf{O}_-$ defined in (3.43). Assume that the pencil $\lambda E - A$ is d-stable. Then the infinite sums

$$\mathcal{G}_{dcc} := \mathbf{C}_+\mathbf{C}_+^* = \sum_{k=0}^{\infty} F_k B B^* F_k^* \tag{4.72}$$

and

$$\mathcal{G}_{dco} := \mathbf{O}_+^*\mathbf{O}_+ = \sum_{k=0}^{\infty} F_k^* C^* C F_k, \tag{4.73}$$

where F_k are as in (2.7), converge. The matrix $\mathcal{G}_{dcc}$ is called the *causal controllability Gramian* of the discrete-time descriptor system (3.2) and the matrix $\mathcal{G}_{dco}$ is called the *causal observability Gramian* of (3.2). The matrices

$$\mathcal{G}_{dnc} := \mathbf{C}_-\mathbf{C}_-^* = \sum_{k=-\nu}^{-1} F_k B B^* F_k^* \tag{4.74}$$

and

$$\mathcal{G}_{dno} := \mathbf{O}_-^*\mathbf{O}_- = \sum_{k=-\nu}^{-1} F_k^* C^* C F_k \tag{4.75}$$

are called, respectively, the *noncausal controllability Gramian* and the *noncausal observability Gramian* of system (3.2). In summary, the *controllability Gramian* of the discrete-time descriptor system (3.2) is defined by

$$\mathcal{G}_{dc} = \mathcal{G}_{dcc} + \mathcal{G}_{dnc} \tag{4.76}$$

and the *observability Gramian* for the discrete-time descriptor system (3.2) is defined by

$$\mathcal{G}_{do} = \mathcal{G}_{dco} + \mathcal{G}_{dno}. \tag{4.77}$$

If $E = I$, then $\mathcal{G}_{dc} = \mathcal{G}_{dcc}$ and $\mathcal{G}_{do} = \mathcal{G}_{dco}$ are the classical controllability and observability Gramians for standard discrete-time state space systems [176].

The following lemma gives integral representations for the controllability and observability Gramians of the descriptor system (3.2).

Lemma 4.54. *Consider the discrete-time descriptor system* (3.2)*. Let* $\lambda E - A$ *be d-stable.*

1. *The controllability Gramian of system* (3.2) *can be represented as*

$$\mathcal{G}_{dc} = \frac{1}{2\pi}\int\limits_0^{2\pi}(e^{i\varphi}E - A)^{-1}\Big(P_l BB^* P_l^* + (I - P_l)BB^*(I - P_l)^*\Big)(e^{i\varphi}E - A)^{-*}d\varphi. \quad (4.78)$$

2. *The observability Gramian of system* (3.2) *can be represented as*

$$\mathcal{G}_{do} = \frac{1}{2\pi}\int\limits_0^{2\pi}(e^{i\varphi}E - A)^{-*}\Big(P_r^* C^* C P_r + (I - P_r)^* C^* C(I - P_r)\Big)(e^{i\varphi}E - A)^{-1}d\varphi. \quad (4.79)$$

Proof. Since all finite eigenvalues of the pencil $\lambda E - A$ lie inside the unit circle, the sequence $\|F_k\|$ is uniformly bounded for all integers k. Then the Fourier series

$$\sum_{k=-\infty}^{\infty} F_k e^{ik\varphi}$$

converges [135]. Using (2.11) we have

$$(E - e^{i\varphi}A)\sum_{k=-\infty}^{\infty} F_k e^{ik\varphi} = \sum_{k=-\infty}^{\infty}(EF_k - AF_{k-1})e^{ik\varphi} = I$$

and, hence,

$$(E - e^{i\varphi}A)^{-1} = \sum_{k=-\infty}^{\infty} F_k e^{ik\varphi} = \sum_{k=-\nu}^{\infty} F_k e^{ik\varphi} \quad (4.80)$$

is the Fourier expansion of the matrix-valued function $(E - e^{i\varphi}A)^{-1}$. It immediately follows from the Parseval identity [135] that

$$\begin{aligned}
\mathcal{G}_{dcc} &= \sum_{k=0}^{\infty} F_k BB^* F_k^* = \sum_{k=-\infty}^{\infty} F_k P_l BB^* P_l^* F_k^* \\
&= \frac{1}{2\pi}\int_0^{2\pi}(E - e^{i\varphi}A)^{-1}P_l BB^* P_l^*(E - e^{i\varphi}A)^{-*}d\varphi \\
&= \frac{1}{2\pi}\int_0^{2\pi}(e^{i\varphi}E - A)^{-1}P_l BB^* P_l^*(e^{i\varphi}E - A)^{-*}d\varphi, \quad (4.81)\\
\mathcal{G}_{dnc} &= \sum_{k=-\nu}^{-1} F_k BB^* F_k^* = \sum_{k=-\infty}^{\infty} F_k(I - P_l)BB^*(I - P_l)^* F_k^* \\
&= \frac{1}{2\pi}\int_0^{2\pi}(e^{i\varphi}E - A)^{-1}(I - P_l)BB^*(I - P_l)^*(e^{i\varphi}E - A)^{-*}d\varphi. \quad (4.82)
\end{aligned}$$

Thus, (4.76), (4.81) and (4.82) imply (4.78). The integral representation (4.79) for G_{do} can be obtained analogously. □

As a consequence of Corollaries 4.40, 4.50 and Lemma 4.54 we obtain the following result.

Corollary 4.55. *Consider the discrete-time descriptor system* (3.2)*. Let the pencil* $\lambda E - A$ *be d-stable.*

1. *The causal observability Gramian* $\mathcal{G}_{dco}$ *of* (3.2) *exists and is a unique Hermitian solution of the projected GDALE*

$$A^* X A - E^* X E = -P_r^* C^* C P_r, \qquad X = X P_l. \tag{4.83}$$

   Moreover, $\mathcal{G}_{dco}$ *is positive definite on the subspace* $\operatorname{Im} P_l$ *if and only if the triplet* (E, A, C) *is R-observable.*

2. *The noncausal observability Gramian* $\mathcal{G}_{dno}$ *of* (3.2) *is a unique Hermitian solution of the projected GDALE*

$$A^* X A - E^* X E = (I - P_r)^* C^* C (I - P_r), \qquad X = X(I - P_l). \tag{4.84}$$

   Moreover, $\mathcal{G}_{dno}$ *is positive definite on* $\operatorname{Ker} P_l$ *if and only if* $\operatorname{rank} \begin{bmatrix} E \\ C \end{bmatrix} = n$.

3. *The observability Gramian* $\mathcal{G}_{do}$ *of* (3.2) *exists and is a unique Hermitian solution of the projected GDALE* (4.68)*. Moreover,* $\mathcal{G}_{do}$ *is positive definite on* $\mathbb{F}^n$ *if and only if the triplet* (E, A, C) *is C-observable.*

An analogous result holds for the controllability Gramians.

Corollary 4.56. *Consider the discrete-time descriptor system* (3.2)*. Let the pencil* $\lambda E - A$ *be d-stable.*

1. *The causal controllability Gramian* $\mathcal{G}_{dcc}$ *of* (3.2) *exists and is a unique Hermitian solution of the projected GDALE*

$$A X A^* - E X E^* = -P_l B B^* P_l^*, \qquad X = P_r X.$$

   Moreover, $\mathcal{G}_{dcc}$ *is positive definite on the subspace* $\operatorname{Im} P_r^*$ *if and only if the triplet* (E, A, B) *is R-controllable.*

2. *The noncausal controllability Gramian* $\mathcal{G}_{dnc}$ *of* (3.2) *is a unique Hermitian solution of the projected GDALE*

$$A X A^* - E X E^* = (I - P_l) B B^* (I - P_l)^*, \qquad X = (I - P_r) X.$$

   Moreover, $\mathcal{G}_{dnc}$ *is positive definite on* $\operatorname{Ker} P_r^*$ *if and only if* $\operatorname{rank}[\,E,\ B\,] = n$.

3. *The controllability Gramian* $\mathcal{G}_{dc}$ *of* (3.2) *exists and is a unique Hermitian solution of the projected GDALE* (4.71) *with* $\xi = 1$*. Moreover,* $\mathcal{G}_{dc}$ *is positive definite on* $\mathbb{F}^n$ *if and only if the triplet* (E, A, B) *is C-controllable.*

4.4.2 The continuous-time case

Consider now the continuous-time descriptor system (3.1). Assume that the pencil $\lambda E - A$ is c-stable and the fundamental solution matrix $\mathcal{F}(t)$ is as in (3.12). Then the infinite integrals

$$\mathcal{G}_{cpc} = \int_0^\infty \mathcal{F}(t) B B^* \mathcal{F}^*(t)\, dt \tag{4.85}$$

and

$$\mathcal{G}_{cpo} = \int_0^\infty \mathcal{F}^*(t) C^* C \mathcal{F}(t)\, dt \tag{4.86}$$

exist. The matrix $\mathcal{G}_{cpc}$ is called the *proper controllability Gramian* and the matrix $\mathcal{G}_{cpo}$ is called the *proper observability Gramian* of the continuous-time descriptor system (3.1). The *improper controllability Gramian* and the *improper observability Gramian* of (3.1) are defined by

$$\mathcal{G}_{cic} = \sum_{k=-\nu}^{-1} F_k B B^* F_k^* \tag{4.87}$$

and

$$\mathcal{G}_{cio} = \sum_{k=-\nu}^{-1} F_k^* C^* C F_k, \tag{4.88}$$

respectively, were the matrices F_k are as in (2.7). The *controllability Gramian* of the descriptor system (3.1) is given by

$$\mathcal{G}_{cc} = \mathcal{G}_{cpc} + \mathcal{G}_{cic} \tag{4.89}$$

and the *observability Gramian* of the continuous-time descriptor system (3.1) has the form

$$\mathcal{G}_{co} = \mathcal{G}_{cpo} + \mathcal{G}_{cio}. \tag{4.90}$$

In the case $E = I$ the proper controllability and observability Gramians are classical controllability and observability Gramians of standard continuous-time state space systems [176].

It follows from (4.74), (4.75), (4.87) and (4.88) that the improper controllability and observability Gramians of the continuous-time descriptor system (3.1) coincide with the noncausal controllability and observability Gramians of the discrete-time descriptor system (3.2). Therefore, in the sequel we are discussing only the proper controllability and observability Gramians of (3.1).

The following lemma gives integral representations for the proper controllability and observability Gramians $\mathcal{G}_{cpc}$ and $\mathcal{G}_{cpo}$ in terms of the generalized resolvent $(\lambda E - A)^{-1}$.

Lemma 4.57. *Consider the continuous-time descriptor system* (3.1). *Let the pencil $\lambda E - A$ be c-stable.*

1. *The proper controllability Gramian of system* (3.1) *can be represented as*

$$\mathcal{G}_{cpc} = \frac{1}{2\pi}\int_{-\infty}^{\infty}(i\xi E - A)^{-1}P_l BB^* P_l^*(i\xi E - A)^{-*}d\xi. \tag{4.91}$$

2. *The proper observability Gramian of system* (3.1) *can be represented as*

$$\mathcal{G}_{cpo} = \frac{1}{2\pi}\int_{-\infty}^{\infty}(i\xi E - A)^{-*}P_r^* C^* C P_r(i\xi E - A)^{-1}d\xi. \tag{4.92}$$

Proof. From (3.12) we have that the entries of the matrices $P_r(i\xi E-A)^{-1}$ and $(i\xi E-A)^{-1}P_l$ are the Fourier transformations of the entries of $\mathcal{F}(t)$. Then the integrals (4.91) and (4.92) immediately follow from the Parseval identity [135]. □

If we compare the integrals (4.91) and (4.92) with the solutions of the projected GCALEs (4.42) and (4.39), respectively, then from Corollaries 4.15, 4.25 and Remark 4.27 we obtain the following result.

Corollary 4.58. *Consider the continuous-time descriptor system* (3.1). *Let the pencil* $\lambda E - A$ *be c-stable.*

1. *The proper controllability Gramian* $\mathcal{G}_{cpc}$ *of* (3.1) *exists and is a unique Hermitian solution of the projected GCALE* (4.42). *Moreover,* $\mathcal{G}_{cpc}$ *is positive definite on* $\operatorname{Im} P_r^*$ *if and only if the triplet* (E, A, B) *is R-controllable.*

2. *The proper observability Gramian* $\mathcal{G}_{cpo}$ *of* (3.1) *exists and is a unique Hermitian solution of the projected GCALE* (4.39). *Moreover,* $\mathcal{G}_{cpo}$ *is positive definite on* $\operatorname{Im} P_l$ *if and only if the triplet* (E, A, C) *is R-observable.*

Remark 4.59. Corollaries 4.55, 4.56 and 4.58 imply the following conditions.

1. The controllability Gramian $\mathcal{G}_{cc}$ of (3.1) is positive definite if and only if the pencil $\lambda E - A$ is c-stable and the triplet (E, A, B) is C-controllable.

2. The observability Gramian $\mathcal{G}_{co}$ of (3.1) is positive definite if and only if the pencil $\lambda E - A$ is c-stable and the triplet (E, A, C) is C-observable.

It should be noted that the proper controllability (observability) Gramian of (3.1) is defined via the projected generalized continuous-time Lyapunov equation and the improper controllability (observability) Gramian of (3.1) is defined via the projected generalized discrete-time Lyapunov equation. Unlike the discrete-time descriptor system (3.2), we do not know how to express the controllability and observability Gramians of the continuous-time descriptor system (3.1) via solutions of a single Lyapunov equation.

Chapter 5

Numerical solution of generalized Lyapunov equations

Due to the practical importance the numerical solution of Lyapunov equations has received a lot of attention, see [9, 17, 55, 64, 72, 80, 100, 109, 126, 127, 136, 146] and the references therein. The classical numerical methods for standard Lyapunov equations are the Bartels-Stewart method [9], the Hammarling method [72] and the Hessenberg-Schur method [65]. An extension of these methods to regular generalized Lyapunov equations is given in [34, 55, 56, 65, 117, 125]. These methods are based on the preliminary reduction of the matrix (matrix pencil) to the (generalized) Schur form [64] or the Hessenberg-Schur form [65], calculation of the solution of a reduced system and back transformation.

In this chapter we extend the Bartels-Stewart and Hammarling methods for projected Lyapunov equations. A review of iterative methods for (generalized) Lyapunov equations is also presented.

5.1 Generalized Schur-Bartels-Stewart method

Consider the projected GCALE

$$\begin{aligned} &E^T XA + A^T XE = -P_r^T G P_r, \\ &X = XP_l, \end{aligned} \tag{5.1}$$

where E, A, $G \in \mathbb{R}^{n,n}$ (the complex case is similar). Let the pencil $\lambda E - A$ be in the GUPTRI form (2.4). To compute the right and left deflating subspaces of $\lambda E - A$ corresponding to the finite eigenvalues we need to compute matrices Y and Z such that

$$\begin{bmatrix} I & -Z \\ 0 & I \end{bmatrix} \begin{bmatrix} \lambda E_f - A_f & \lambda E_u - A_u \\ 0 & \lambda E_\infty - A_\infty \end{bmatrix} \begin{bmatrix} I & Y \\ 0 & I \end{bmatrix} = \begin{bmatrix} \lambda E_f - A_f & 0 \\ 0 & \lambda E_\infty - A_\infty \end{bmatrix}.$$

This leads to the generalized Sylvester equation

$$\begin{aligned} E_f Y - Z E_\infty &= -E_u, \\ A_f Y - Z A_\infty &= -A_u. \end{aligned} \tag{5.2}$$

Since the pencils $\lambda E_f - A_f$ and $\lambda E_\infty - A_\infty$ have no common eigenvalues, equation (5.2) has a unique solution (Y, Z) [34]. Then the pencil $\lambda E - A$ can be reduced by an equivalence transformation to the Weierstrass-like canonical form

$$\begin{aligned} \lambda E - A &= V \begin{bmatrix} I & Z \\ 0 & I \end{bmatrix} \begin{bmatrix} \lambda E_f - A_f & 0 \\ 0 & \lambda E_\infty - A_\infty \end{bmatrix} \begin{bmatrix} I & -Y \\ 0 & I \end{bmatrix} U^T \\ &= W_1 \begin{bmatrix} \lambda E_f - A_f & 0 \\ 0 & \lambda E_\infty - A_\infty \end{bmatrix} T_1, \end{aligned}$$

where the matrices

$$W_1 = V \begin{bmatrix} I & Z \\ 0 & I \end{bmatrix} \qquad \text{and} \qquad T_1 = \begin{bmatrix} I & -Y \\ 0 & I \end{bmatrix} U^T$$

are nonsingular. In this case the spectral projections P_r and P_l onto the right and left finite deflating subspaces of $\lambda E - A$ have the form

$$P_r = T_1^{-1} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} T_1 = U \begin{bmatrix} I & -Y \\ 0 & 0 \end{bmatrix} U^T, \tag{5.3}$$

$$P_l = W_1 \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} W_1^{-1} = V \begin{bmatrix} I & -Z \\ 0 & 0 \end{bmatrix} V^T. \tag{5.4}$$

Assume that the pencil $\lambda E - A$ is c-stable. Setting

$$V^T X V = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \qquad \text{and} \qquad U^T G U = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix}, \tag{5.5}$$

we obtain from the GCALE in (5.1) the decoupled system of matrix equations

$$E_f^T X_{11} A_f + A_f^T X_{11} E_f = -G_{11}, \tag{5.6}$$

$$E_f^T X_{12} A_\infty + A_f^T X_{12} E_\infty = G_{11} Y - E_f^T X_{11} A_u - A_f^T X_{11} E_u, \tag{5.7}$$

$$E_\infty^T X_{21} A_f + A_\infty^T X_{21} E_f = Y^T G_{11} - E_u^T X_{11} A_f - A_u^T X_{11} E_f, \tag{5.8}$$

$$\begin{aligned} E_\infty^T X_{22} A_\infty + A_\infty^T X_{22} E_\infty = &- Y^T G_{11} Y - E_u^T X_{11} A_u - A_u^T X_{11} E_u - E_\infty^T X_{21} A_u \\ &- A_\infty^T X_{21} E_u - E_u^T X_{12} A_\infty - A_u^T X_{12} E_\infty. \end{aligned} \tag{5.9}$$

Since all eigenvalues of $\lambda E_f - A_f$ are finite and lie in the open left half-plane, by Theorem 4.4 the GCALE (5.6) has a unique solution X_{11}. The pencils $\lambda E_f - A_f$ and $\lambda E_\infty - A_\infty$ have no eigenvalues in common and, hence, by Theorem 4.2 the generalized Sylvester equations (5.7) and (5.8) are uniquely solvable. To show that the matrix $X_{12} = -X_{11} Z$ satisfies equation (5.7), we substitute this matrix in (5.7). Taking into account equations (5.2) and (5.6), we obtain

$$\begin{aligned} E_f^T X_{12} A_\infty + A_f^T X_{12} E_\infty &= - E_f^T X_{11} (A_f Y + A_u) - A_f^T X_{11} (E_f Y + E_u) \\ &= - (E_f^T X_{11} A_f + A_f^T X_{11} E_f) Y - E_f^T X_{11} A_u - A_f^T X_{11} E_u \\ &= G_{11} Y - E_f^T X_{11} A_u - A_f^T X_{11} E_u. \end{aligned}$$

Similarly, it can be verified that the matrix $X_{21} = -Z^T X_{11}$ is the solution of (5.8).

Consider now equation (5.9). Substitute the matrices $X_{12} = -X_{11}Z$, $X_{21} = -Z^T X_{11}$ in (5.9). Using (5.2) and (5.6) we obtain

$$\begin{aligned} E_\infty^T X_{22} A_\infty + A_\infty^T X_{22} E_\infty &= Y^T E_f^T X_{11}(ZA_\infty - A_f Y) + Y^T A_f^T X_{11}(ZE_\infty - E_f Y) \\ &\quad + E_u^T X_{11} Z A_\infty + A_u^T X_{11} Z E_\infty - Y^T G_{11} Y \\ &= (E_f Y + E_u)^T X_{11} Z A_\infty + (A_f Y + A_u)^T X_{11} Z E_\infty \\ &= E_\infty^T Z^T X_{11} Z A_\infty + A_\infty^T Z^T X_{11} Z E_\infty. \end{aligned}$$

Then

$$E_\infty^T (X_{22} - Z^T X_{11} Z) A_\infty + A_\infty^T (X_{22} - Z^T X_{11} Z) E_\infty = 0. \tag{5.10}$$

Clearly, $X_{22} = Z^T X_{11} Z$ satisfies (5.9). Moreover, we have

$$\begin{aligned} X &= V \begin{bmatrix} X_{11} & -X_{11}Z \\ -Z^T X_{11} & Z^T X_{11} Z \end{bmatrix} V^T \\ &= V \begin{bmatrix} X_{11} & -X_{11}Z \\ -Z^T X_{11} & Z^T X_{11} Z \end{bmatrix} \begin{bmatrix} I & -Z \\ 0 & 0 \end{bmatrix} V^T = X P_l. \end{aligned}$$

Thus, the matrix

$$X = V \begin{bmatrix} X_{11} & -X_{11}Z \\ -Z^T X_{11} & Z^T X_{11} Z \end{bmatrix} V^T \tag{5.11}$$

is the unique solution of the projected GCALE (5.1).

In some applications we need the matrix $E^T X E$ rather that the solution X itself [147]. Using (2.4), (5.2) and (5.11) we obtain that

$$E^T X E = U \begin{bmatrix} E_f^T X_{11} E_f & -E_f^T X_{11} E_f Y \\ -Y^T E_f^T X_{11} E_f & Y^T E_f^T X_{11} E_f Y \end{bmatrix} U^T.$$

Remark 5.1. It follows from (5.10) that the general solution of the GCALE in (5.1) has the form

$$X = V \begin{bmatrix} X_{11} & -X_{11}Z \\ -Z^T X_{11} & X_\infty + Z^T X_{11} Z \end{bmatrix} V^T,$$

where X_∞ is the general solution of the homogeneous GCALE $E_\infty^T X_\infty A_\infty + A_\infty^T X_\infty E_\infty = 0$. If we require for this solution to satisfy $X = XP_l$, then we obtain that $X_\infty = 0$.

In summary, we have the following algorithm for computing the solution X of the projected GCALE (5.1).

Algorithm 5.1.1. *Generalized Schur-Bartels-Stewart method for the projected GCALE.*
Input: *A real symmetric matrix G and a real regular pencil $\lambda E - A$ such that $\lambda_j + \lambda_k \neq 0$ for any two finite eigenvalues λ_j and λ_k of $\lambda E - A$.*
Output: *The symmetric solution X of the projected GCALE* (5.1).

Step 1. *Use the GUPTRI algorithm* [41, 42] *to compute* (2.4).
Step 2. *Use the generalized Schur method* [86, 87] *or the recursive blocked algorithm* [81] *to solve the generalized Sylvester equation* (5.2).
Step 3. *Compute the matrix*

$$U^T G U = \begin{bmatrix} G_{11} & G_{12} \\ G_{12}^T & G_{22} \end{bmatrix}. \tag{5.12}$$

Step 4. *Use the generalized Bartels-Stewart method* [9, 125] *or the recursive blocked algorithm* [82] *to solve the regular GCALE*

$$E_f^T X_{11} A_f + A_f^T X_{11} E_f = -G_{11}. \tag{5.13}$$

Step 5. *Compute the matrix*

$$X = V \begin{bmatrix} X_{11} & -X_{11}Z \\ -Z^T X_{11} & Z^T X_{11} Z \end{bmatrix} V^T. \tag{5.14}$$

Consider now the projected GDALE

$$\begin{aligned} & A^T X A - E^T X E = -P_r^T G P_r + \xi (I - P_r)^T G (I - P_r), \\ & P_l^T X = X P_l, \end{aligned} \tag{5.15}$$

where $\xi = -1$, 0 or 1. Assume that the pencil $\lambda E - A$ is d-stable. Using (2.4), (5.3) and (5.5) we obtain from the GDALE in (5.15) the following system of matrix equations

$$A_f^T X_{11} A_f - E_f^T X_{11} E_f = -G_{11}, \tag{5.16}$$

$$A_f^T X_{12} A_\infty - E_f^T X_{12} E_\infty = G_{11} Y - A_f^T X_{11} A_u + E_f^T X_{11} E_u, \tag{5.17}$$

$$A_\infty^T X_{21} A_f - E_\infty^T X_{21} E_f = Y^T G_{11} - A_u^T X_{11} A_f + E_u^T X_{11} E_f, \tag{5.18}$$

$$\begin{aligned} A_\infty^T X_{22} A_\infty - E_\infty^T X_{22} E_\infty = & -Y^T G_{11} Y + \xi \left(Y^T G_{11} Y + Y^T G_{12} + G_{21} Y + G_{22} \right) \\ & - A_u^T X_{11} A_u - A_u^T X_{12} A_\infty - A_\infty^T X_{21} A_u \\ & + E_u^T X_{11} E_u + E_u^T X_{12} E_\infty + E_\infty^T X_{21} E_u. \end{aligned} \tag{5.19}$$

Since all eigenvalues of the pencil $\lambda E_f - A_f$ lie inside the unit circle, by Theorem 4.30 the regular GDALE (5.16) has a unique solution X_{11}. It follows from $P_l^T X = X P_l$ that $X_{12} = -X_{11} Z$ and $X_{21} = -Z^T X_{11}$. Moreover, we can verify that these matrices satisfy equations (5.17) and (5.18), respectively. Substituting these matrices in (5.19) and taking into account equations (5.2) and (5.16), we obtain that

$$\begin{aligned} A_\infty^T X_{22} A_\infty - E_\infty^T X_{22} E_\infty = & \ \xi \left(Y^T G_{11} Y + Y^T G_{12} + G_{21} Y + G_{22} \right) \\ + & \ A_\infty^T Z^T X_{11} Z A_\infty - E_\infty^T Z^T X_{11} Z E_\infty. \end{aligned}$$

Thus, the solution of the projected GDALE (5.15) has the form

$$X = V \begin{bmatrix} X_{11} & -X_{11}Z \\ -Z^T X_{11} & X_\infty + Z^T X_{11} Z \end{bmatrix} V^T,$$

where X_{11} satisfies the regular GDALE (5.16) and X_∞ is a solution of the regular GDALE

$$A_\infty^T X_\infty A_\infty - E_\infty^T X_\infty E_\infty = \xi(Y^T G_{11} Y + Y^T G_{12} + G_{21} Y + G_{22}).$$

Analogous to the continuous-time case, we have the following algorithm for computing the solution X of the projected GDALE (5.15).

Algorithm 5.1.2. *Generalized Schur-Bartels-Stewart method for the projected GDALE.*
Input: *A real symmetric matrix G and a real regular pencil $\lambda E - A$ such that $\lambda_j \lambda_k \neq 1$ for any two finite eigenvalues λ_j and λ_k of $\lambda E - A$.*
Output: *The symmetric solution X of the projected GDALE* (4.64).
Step 1. *Use the GUPTRI algorithm* [41, 42] *to compute* (2.4).
Step 2. *Use the generalized Schur method* [86, 87] *or the recursive blocked algorithm* [81] *to solve the generalized Sylvester equation* (5.2).
Step 3. *Compute the matrix $U^T G U$ as in* (5.12).
Step 4a. *Use the generalized Bartels-Stewart method* [9, 125] *or the recursive blocked algorithm* [82] *to solve the regular GDALE*

$$A_f^T X_{11} A_f - E_f^T X_{11} E_f = -G_{11}. \tag{5.20}$$

Step 4b. *If $\xi = 0$, then $X_\infty = 0$. Otherwise, use the generalized Bartels-Stewart method* [9, 125] *or the recursive blocked algorithm* [82] *to solve the regular GDALE*

$$A_\infty^T X_\infty A_\infty - E_\infty^T X_\infty E_\infty = \xi(Y^T G_{11} Y + Y^T G_{12} + G_{12}^T Y + G_{22}). \tag{5.21}$$

Step 5. *Compute the matrix*

$$X = V \begin{bmatrix} X_{11} & -X_{11} Z \\ -Z^T X_{11} & X_\infty + Z^T X_{11} Z \end{bmatrix} V^T. \tag{5.22}$$

5.2 Generalized Schur-Hammarling method

In many applications it is necessary to have the Cholesky factor of the solution of the Lyapunov equation rather than the solution itself, e.g., [102]. An attractive algorithm for computing the Cholesky factor of solutions of regular Lyapunov equations with a positive semidefinite right-hand side is the generalized Hammarling method [72, 125] We will show that the Hammarling method can also be used to solve the projected GCALE

$$E^T X A + A^T X E = -P_r^T C^T C P_r, \qquad X = X P_l, \tag{5.23}$$

where $E,\ A \in \mathbb{R}^{n,n}$, $C \in \mathbb{R}^{p,n}$. In fact, we can compute the full rank factorization [99] of the solution $X = L^T L$ without constructing X and the matrix product $C^T C$ explicitly.

Let $\lambda E - A$ be in the GUPTRI form (2.4) and let $CU = [\, C_1,\ C_2 \,]$ be partitioned in blocks conformally to E and A. Then the solution of the projected GCALE (5.23) has the form (5.11), where the symmetric, positive semidefinite matrix X_{11} satisfies the GCALE

$$E_f^T X_{11} A_f + A_f^T X_{11} E_f = -C_1^T C_1.$$

Let $U_{X_{11}}$ be a Cholesky factor of the solution $X_{11} = U_{X_{11}}^T U_{X_{11}}$. Compute the QR decomposition

$$U_{X_{11}} = Q \begin{bmatrix} L_1 \\ 0 \end{bmatrix},$$

where Q is orthogonal and L_1 has full row rank [64]. Then

$$\begin{aligned} X &= V \begin{bmatrix} U_{X_{11}}^T \\ -Z^T U_{X_{11}}^T \end{bmatrix} [\, U_{X_{11}}, \;\; -U_{X_{11}} Z \,] V^T \\ &= V \begin{bmatrix} L_1^T \\ -Z^T L_1^T \end{bmatrix} [\, L_1, \;\; -L_1 Z \,] V^T = L^T L \end{aligned}$$

is the full rank factorization of X, where $L = [\, L_1, \;\; -L_1 Z \,] V^T$ has full row rank.

Thus, we have the following algorithm for computing the full row rank factor of the solution of the projected GCALE (5.23).

Algorithm 5.2.1. *Generalized Schur-Hammarling method for the projected GCALE* (5.23)
Input: *A real c-stable pencil $\lambda E - A$ and a real matrix C.*
Output: *A full row rank factor L of the solution $X = L^T L$ of* (5.23).
Step 1. *Use the GUPTRI algorithm* [41, 42] *to compute* (2.4).
Step 2. *Use the generalized Schur method* [86, 87] *or the recursive blocked algorithm* [81] *to compute the solution of the generalized Sylvester equation* (5.2).
Step 3. *Compute the matrix*

$$CU = [\, C_1, \quad C_2 \,]. \tag{5.24}$$

Step 4. *Use the generalized Hammarling method* [72, 125] *to compute the Cholesky factor $U_{X_{11}}$ of the solution $X_{11} = U_{X_{11}}^T U_{X_{11}}$ of the GCALE*

$$E_f^T X_{11} A_f + A_f^T X_{11} E_f = -C_1^T C_1. \tag{5.25}$$

Step 5a. *If* $\mathrm{rank}(U_{X_{11}}) < n_f$, *then use Householder or Givens transformations* [64] *to compute the full row rank matrix L_1 from the QR decomposition $U_{X_{11}} = Q_{L_1} \begin{bmatrix} L_1 \\ 0 \end{bmatrix}$. Otherwise, $L_1 := U_{X_{11}}$.*
Step 5b. *Compute the full row rank factor*

$$L = [\, L_1, \;\; -L_1 Z \,] V^T. \tag{5.26}$$

In some applications we need to compute the full column rank factor R of the solution $X = RR^T$ of the dual projected GCALE

$$EXA^T + AXE^T = -P_l B B^T P_l^T, \qquad X = P_r X, \tag{5.27}$$

where $E,\ A \in \mathbb{R}^{n,n}$, $B \in \mathbb{R}^{n,m}$. Algorithm 5.2.1 can be rewritten for this equation as follows.

Algorithm 5.2.2. *Generalized Schur-Hammarling method for the projected GCALE* (5.27)
Input: *A real c-stable pencil* $\lambda E - A$ *and a real matrix* B.
Output: *A full column rank factor* R *of the solution* $X = RR^T$ *of* (5.27).
Step 1. *Use the GUPTRI algorithm* [41, 42] *to compute* (2.4).
Step 2. *Use the generalized Schur method* [86, 87] *or the recursive blocked algorithm* [81] *to compute the solution of the generalized Sylvester equation* (5.2).
Step 3. *Compute the matrix*

$$V^T B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}.$$

Step 4. *Use the generalized Hammarling method* [72, 125] *to compute the Cholesky factor* $U_{X_{11}}$ *of the solution* $X_{11} = U_{X_{11}}^T U_{X_{11}}$ *of the regular GCALE*

$$E_f X_{11} A_f^T + A_f X_{11} E_f^T = -(B_1 - ZB_2)(B_1 - ZB_2)^T. \tag{5.28}$$

Step 5a. *If* $\operatorname{rank}(U_{X_{11}}) < n_f$, *then use Householder or Givens transformations* [64] *to compute the full column rank matrix* R_1 *from the QR decomposition* $U_{X_{11}} = Q_{R_1} \begin{bmatrix} R_1^T \\ 0 \end{bmatrix}$. *Otherwise,* $R_1 := U_{X_{11}}^T$.
Step 5b. *Compute the full column rank factor*

$$R = U \begin{bmatrix} R_1 \\ 0 \end{bmatrix}. \tag{5.29}$$

Analogous to the continuous-time case, we obtain the following algorithm for computing the full row rank factor L of the solution $X = L^T L$ of the projected GDALE

$$\begin{aligned} &A^T X A - E^T X E = -P_r^T C^T C P_r + \xi (I - P_r)^T C^T C (I - P_r), \\ &P_l^T X = X P_l \end{aligned} \tag{5.30}$$

where ξ is 0 or 1. Note that for $\xi = -1$, the solution of (5.30) is indefinite and, hence, the full rank factorization for this solution does not exist.

Algorithm 5.2.3. *Generalized Schur-Hammarling method for the projected GDALE*
Input: *A real d-stable pencil* $\lambda E - A$ *and a real matrix* C.
Output: *A full row rank factor* L *of the solution* $X = L^T L$ *of the projected GDALE* (5.30) *with* $\xi = 0$ *or* 1.
Step 1. *Use the GUPTRI algorithm* [41, 42] *to compute* (2.4).
Step 2. *Use the generalized Schur method* [86, 87] *or the recursive blocked algorithm* [81] *to solve the generalized Sylvester equation* (5.2).
Step 3. *Compute the matrix* CU *as in* (5.24).
Step 4a. *Use the generalized Hammarling method* [72, 125] *to compute the Cholesky factor* $U_{X_{11}}$ *of the solution* $X_{11} = U_{X_{11}}^T U_{X_{11}}$ *of the regular GDALE*

$$A_f^T X_{11} A_f - E_f^T X_{11} E_f = -C_1^T C_1. \tag{5.31}$$

Step 4b. *If $\xi = 0$, then $U_{X_\infty} = 0$. Otherwise, use the generalized Hammarling method* [72, 125] *to compute the Cholesky factor U_{X_∞} of the solution $X_\infty = U_{X_\infty}^T U_{X_\infty}$ of the regular GDALE*

$$A_\infty^T X_\infty A_\infty - E_\infty^T X_\infty E_\infty = (C_1 Y + C_2)^T (C_1 Y + C_2). \tag{5.32}$$

Step 5. *Use the Householder or Givens transformations* [64] *to compute the full row rank matrix L from the QR decomposition*

$$\begin{pmatrix} U_{X_{11}} & -U_{X_{11}} Z \\ 0 & U_{X_\infty} \end{pmatrix} V^T = Q \begin{bmatrix} L \\ 0 \end{bmatrix}. \tag{5.33}$$

An algorithm for computing the full column rank factor of the solution of the dual projected GDALE can be obtained in the same way.

5.3 Numerical aspects and complexity

We will now discuss numerical aspects and computational cost for the algorithms described in the previous subsections in detail. We focus on Algorithm 5.1.1 and give some notes about the differences to the other algorithms.

Step 1. To deflate the infinite eigenvalues of the pencil $\lambda E - A$ and to reduce this pencil to the quasi-triangular form (2.4) we use the GUPTRI algorithm [41, 42]. This algorithm is based on the computation of the infinity-staircase form [161] of $\lambda E - A$ which exposes the Jordan structure of the infinite eigenvalues, and the QZ algorithm [64] for a subpencil which gives quasi-triangular blocks with the finite eigenvalues. The GUPTRI algorithm is numerically backwards stable and requires $O(n^3)$ operations [41].

Step 2. To solve the generalized Sylvester equation (5.2) we can use the generalized Schur method [86, 87]. Note that the pencils $\lambda E_f - A_f$ and $\lambda E_\infty - A_\infty$ are already in the generalized real Schur form [64], that is, the matrices E_f and E_∞ are upper triangular, whereas the matrices A_f and A_∞ are upper quasi-triangular. Since the infinite eigenvalues of $\lambda E_\infty - A_\infty$ correspond to the zero eigenvalues of the reciprocal pencil $E_\infty - \mu A_\infty$, we obtain that A_∞ is upper triangular. Let $A_f = [A_{ij}^f]_{i,j=1}^k$ and $A_\infty = [A_{ij}^\infty]_{i,j=1}^l$ be partitioned such that the diagonal blocks A_{jj}^f are of size 1×1 or 2×2 and A_{jj}^∞ are of size 1×1. Let $E_f = [E_{ij}^f]_{i,j=1}^k$, $E_\infty = [E_{ij}^\infty]_{i,j=1}^l$, $E_u = [E_{ij}^u]_{i,j=1}^{k,l}$, $A_u = [A_{ij}^u]_{i,j=1}^{k,l}$, $Y = [Y_{ij}]_{i,j=1}^{k,l}$ and $Z = [Z_{ij}]_{i,j=1}^{k,l}$ be partitioned in blocks conformally to A_f and A_∞. Then (5.2) is equivalent to the kl equations

$$E_{tt}^f Y_{tq} - Z_{tq} E_{qq}^\infty = -E_{tq} - \sum_{j=t+1}^{k} E_{tj}^f Y_{jq} + \sum_{j=1}^{q-1} Z_{tj} E_{jq}^\infty =: -\check{E}_{tq}, \tag{5.34}$$

$$A_{tt}^f Y_{tq} - Z_{tq} A_{qq}^\infty = -A_{tq} - \sum_{j=t+1}^{k} A_{tj}^f Y_{jq} + \sum_{j=1}^{q-1} Z_{tj} A_{jq}^\infty =: -\check{A}_{tq} \tag{5.35}$$

for $t = 1, \ldots, k$ and $q = 1, \ldots, l$. The matrices Y_{tq} and Z_{tq} can be computed successively in a row-wise order beginning with $t = k$ and $q = l$ from these equations. Since $E^{\infty}_{qq} = 0$, the 1×1 or 2×1 matrix Y_{tq} can be computed from the linear equation (5.34) of size 1×1 or 2×2 using Gaussian elimination with partial pivoting [64]. Then from (5.35) we obtain

$$Z_{tq} = (A^f_{tt} Y_{pq} + \check{A}_{tq})(A^{\infty}_{qq})^{-1}.$$

The algorithm for solving the generalized Sylvester equation (5.2) via the generalized Schur method is available as the LAPACK subroutine `_TGSYL` [1] and costs $2n_f^2 n_\infty + 2n_f n_\infty^2$ flops [87].

To compute the solution of the quasi-triangular generalized Sylvester equation (5.2) we can also use the recursive blocked algorithm [81, Algorithm 3]. This algorithm consists in the recursive splitting equation (5.2) in smaller subproblems that can be solved using high-performance kernel solvers. For comparison of the recursive blocked algorithm and the LAPACK subroutine, see [81].

Step 3 is a matrix multiplication. In fact, in Algorithm 5.1.1 only the $n_f \times n_f$ block G_{11} in (5.12) is needed. Let $U = [\,U_1,\, U_2\,]$, where the columns of the $(n \times n_f)$-matrix U_1 form the basis of the right finite deflating subspace of $\lambda E - A$. Exploiting the symmetry of G, the computation of $G_{11} = U_1^T G U_1$ requires $n^2 n_f + n n_f^2/2$ flops. In Algorithm 5.2.1 we only need the $p \times n_f$ block C_1 in (5.24) which can be computed as $C_1 = CU_1$ in npn_f flops. The computation of $U^T G U$ in Algorithm 5.1.2, $V^T B$ in Algorithm 5.2.2 and CU in Algorithm 5.2.3 requires $3n^3/2$, mn^2 and pn^2 flops, respectively.

Step 4. To solve the regular GCALE (5.13) in Algorithm 5.1.1 and the regular GDALEs (5.20), (5.21) in Algorithm 5.1.2 we can use the generalized Bartels-Stewart method [9, 125]. Here we briefly describe the generalized Bartels-Stewart method for the GCALE (5.13). Let the matrices $X_{11} = [X'_{ij}]^k_{i,j=1}$ and $G_{11} = [G'_{ij}]^k_{i,j=1}$ be partitioned in blocks conformally to E_f and A_f. Then equation (5.13) is equivalent to k^2 equations

$$(E^f_{tt})^T X'_{tq} A^f_{qq} + (A^f_{tt})^T X'_{tq} E^f_{qq} = -\check{G}_{tq}, \qquad t, q = 1, \ldots k, \tag{5.36}$$

where

$$\begin{aligned} \check{G}_{tq} &= G'_{tq} + \sum_{\substack{i=1,j=1 \\ (i,j)\neq(t,q)}}^{t,q} \Big((E^f_{it})^T X'_{ij} A^f_{jq} + (A^f_{it})^T X'_{ij} E^f_{jq} \Big) \\ &= G'_{tq} + \sum_{i=1}^{t} \left[(E^f_{it})^T \left(\sum_{j=1}^{q-1} X'_{ij} A^f_{jq} \right) + (A^f_{it})^T \left(\sum_{j=1}^{q-1} X'_{ij} E^f_{jq} \right) \right] \\ &\quad + \sum_{i=1}^{t-1} \Big[(E^f_{it})^T X'_{iq} A^f_{qq} + (A^f_{it})^T X'_{iq} E^f_{qq} \Big]. \end{aligned}$$

We compute the blocks X'_{tq} in a row-wise order beginning with $t = q = 1$. Using the column-wise vector representation of the matrices X'_{tq} and $\check{G}_{tq}$ we can rewrite the generalized Sylvester equation (5.36) as a linear system

$$\Big((A^f_{qq})^T \otimes (E^f_{tt})^T + (E^f_{qq})^T \otimes (A^f_{tt})^T \Big) \operatorname{vec}(X'_{tq}) = -\operatorname{vec}(\check{G}_{tq}) \tag{5.37}$$

of size 2×2, 4×4 or 8×8. The solution $\mathrm{vec}(X'_{tq})$ can be computed by solving (5.37) via Gaussian elimination with partial pivoting [64].

To compute the Cholesky factors of solutions of the GCALE (5.25) in Algorithm 5.2.1 and the regular GDALEs (5.31), (5.32) in Algorithm 5.2.3 we can use the generalized Hammarling method, see [72, 125] for details.

The solutions of the regular Lyapunov equations (5.13) and (5.20) using the generalized Bartels-Stewart method requires $O(n_f^3)$ flops, while computing the Cholesky factors of solutions of equations (5.25) and (5.31) via the generalized Hammarling method requires $O(n_f^3 + pn_f^2 + p^2 n_f)$ flops [125]. The computation of the right-hand side in the regular GDALE (5.21) and the solution of this equation requires $O(n_\infty^3 + n_f^2 n_\infty + n_f n_\infty^2)$. Calculation of the right-hand sides in the regular GCALE (5.28) and the regular GDALE (5.32) and the Cholesky factors of the solutions of these equations costs $O(n_f^3 + m^2 n_f + mn_f^2 + mn_f n_\infty)$ and $O(n_\infty^3 + pn_\infty^2 + p^2 n_\infty + pn_f n_\infty)$ flops, respectively.

The generalized Bartels-Stewart method and the generalized Hammarling method are implemented in LAPACK-style subroutines `SG03AD` and `SG03BD`, respectively, that are available in the SLICOT Library [16].

The quasi-triangular generalized Lyapunov equations (5.13), (5.20) and (5.21) can also be solved using the recursive blocked algorithm [82, Algorithm 3]. Comparison of this algorithm with the SLICOT subroutines can be found in [82].

Step 5. The matrices X in (5.14) and (5.22) are computed in $O(n^3 + n_f^2 n_\infty + n_f n_\infty^2)$ flops. The computation of the full rank factor L in (5.26) and R in (5.29) requires, respectively, $O(n_f^3 + n_f n_\infty r_1 + n^2 r_1)$ and $O(n_f^3 + nn_f r_2)$ flops, where $r_1 = \mathrm{rank}(L)$ and $r_2 = \mathrm{rank}(R)$. The full row rank factor L in (5.33) is computed in $O(n^3 + n_f^2 n_\infty)$ flops.

Thus, the total computational cost of the generalized Schur-Bartels-Stewart method as well as the generalized Schur-Hammarling method is estimated as $O(n^3)$ and they require $O(n^2)$ memory location. These methods can be used, unfortunately, only for projected Lyapunov equations of small or medium size ($n \leq 1000$). Moreover, they do not take into account the sparsity and any structure of the coefficient matrices and are difficult to be parallelize.

5.4 Iterative methods

Iterative methods are very useful for large scale sparse problems because they are more suitable for parallelization than direct methods and often do not destroy sparsity. In this section we briefly review some iterative methods for (generalized) Lyapunov equations.

The matrix sign function method

One of the most popular approaches to solve large scale dense Lyapunov equations is the *matrix sign function method*. This method was proposed for standard Lyapunov equations in [133], see also [25, 92, 100], and extended to generalized Lyapunov equations in [12, 17, 54, 101].

Consider the GCALE (4.9) with real matrices E, A and G. The matrix sign function method for (4.9) is given by

$$\begin{aligned} A_0 &= A, \qquad G_0 = G, \\ A_{k+1} &= \frac{1}{2}\left(A_k + EA_k^{-1}E\right), \\ G_{k+1} &= \frac{1}{2}\left(G_k + E^T A_k^{-T} G_k A_k^{-1} E\right). \end{aligned} \tag{5.38}$$

If the matrix E is nonsingular and the pencil $\lambda E - A$ is c-stable, then iteration (5.38) is convergent globally quadratic and $X = \frac{1}{2}E^{-T}\Big(\lim\limits_{k\to\infty} G_k\Big)E^{-1}$ satisfy the GCALE (4.9), see [17].

The solution of the GCALE (4.9) with symmetric, positive semidefinite $G = C^T C$ can be computed directly in factored form $X = L^T L$ via

$$\begin{aligned} & A_0 = A, \qquad C_0 = C, \\ & A_{k+1} = \frac{1}{2}\left(A_k + EA_k^{-1}E\right), \\ & \begin{bmatrix} C_k \\ C_k A_k^{-1} E \end{bmatrix} = Q_{k+1} \begin{bmatrix} R_{k+1} \\ 0 \end{bmatrix}, \qquad \text{(QR decomposition)} \\ & C_{k+1} = \frac{1}{\sqrt{2}} R_{k+1}. \end{aligned} \tag{5.39}$$

In this case $L = \frac{1}{\sqrt{2}}\Big(\lim\limits_{k\to\infty} C_k\Big)E^{-1}$, see [18, 101] for details. The stopping criterion in (5.38) and (5.39) can be chosen as $\|A_k + E\| \le tol\|E\|$ for some matrix norm $\|\cdot\|$ and a user-defined tolerance tol. Scaling strategies to accelerate the convergence of the sign function iterations have been presented in [8, 25, 54, 133].

The matrix sign function method can also be used to solve the GDALE (4.43) with nonsingular E by applying to the Cayley-transformed equation (4.51).

Comparison of the matrix sign function method to the generalized Bartels-Stewart and Hammarling methods with respect to the accuracy and computational cost can be found in [17]. There it has been observed that the matrix sign function method is about as expensive as the Bartels-Stewart method and both methods require approximately the same amount of work space. However, the matrix sign function method is more appropriate for parallelization [15] than the generalized Bartels-Stewart method and is currently the only practicable approach to solve regular generalized Lyapunov equations with large scale dense coefficient matrices.

A disadvantage of the matrix sign function method is that a matrix inversion is required in every iteration step which may lead to significant roundoff errors for ill-conditioned A_k. Such difficulties may arise when eigenvalues of the pencil $\lambda E - A$ lie close to the imaginary axis or $\lambda E - A$ is nearly singular. Note that if the matrix E is singular, then A_k diverges for the pencil $\lambda E - A$ of index greater than two and converges to a singular matrix, otherwise, see [152]. Thus, the matrix sign function method cannot be directly utilized for projected generalized Lyapunov equations.

The Malyshev algorithm

A different approach to compute approximate solutions of generalized Lyapunov equations is the Malyshev algorithm proposed in [112, 113], see also [7, 12, 62], to compute deflating subspaces of a pencil corresponding to eigenvalues inside and outside the unit circle.

Consider the projected GDALE (4.64) with real matrices E, A and $G = I$. Assume that the pencil $\lambda E - A$ is d-stable. Note that E is not necessarily nonsingular. The Malyshev algorithm is described by the following schema

$$\begin{aligned} & E_0 = E^T, \quad A_0 = A^T, \\ & \begin{bmatrix} E_k \\ -A_k \end{bmatrix} = \begin{bmatrix} Q_{1k} & Q_{2k} \\ Q_{3k} & Q_{4k} \end{bmatrix} \begin{bmatrix} R_k \\ 0 \end{bmatrix}, \qquad \text{(QR decomposition)}, \\ & E_{k+1} = Q_{4k}^T E_k, \quad A_{k+1} = Q_{2k}^T A_k. \end{aligned} \tag{5.40}$$

Then the solution of the projected GDALE (4.64) is given by

$$\begin{aligned} X = \lim_{k\to\infty} \Big(& (E_k + A_k)^{-1} E_k (E_k + A_k)^{-1} (E_k + A_k)^{-T} E_k^T (E_k + A_k)^{-T} \\ & + \xi\, (E_k + A_k)^{-1} A_k (E_k + A_k)^{-1} (E_k + A_k)^{-T} A_k^T (E_k + A_k)^{-T} \Big), \end{aligned} \tag{5.41}$$

see [62, 112, 113] for details. For the case $\xi = 0$ or 1, the solution X of (4.64) is symmetric, positive (semi)definite and can be computed in factored form $X = L^T L$ with

$$L = \lim_{k\to\infty} \begin{bmatrix} (E_k + A_k)^{-T} E_k^T (E_k + A_k)^{-T} \\ \xi\, (E_k + A_k)^{-T} A_k^T (E_k + A_k)^{-T} \end{bmatrix}.$$

The Malyshev algorithm can also be used to solve the projected GDALE (4.64) with symmetric, positive definite $G = C^T C$ by applying to the pencil $\lambda E C^{-1} - A C^{-1}$. However, there is no straightforward way to utilize this algorithm for the projected GDALE (4.64), where G is singular.

Iteration (5.40) converges globally quadratically. However, as mentioned in [7, 112], some convergence difficulties may arise if eigenvalues of the pencil $\lambda E - A$ lie close to the unit circle or $\lambda E - A$ is nearly singular.

As a stopping criterion it has been proposed in [112] to use $\|R_k - R_{k-1}\| \le tol \|R_k\|$ with some matrix norm $\|\cdot\|$ and a tolerance tol. Note that for nonsingular E, the pencil $\lambda E - A$ is d-stable if and only if A_k converges to zero. This observation can be used to verify numerically whether the pencil $\lambda E - A$ with nonsingular E is d-stable. We are not aware of a similar d-stability criterion for the case when E is singular.

It should be noted that the Malyshev algorithm converges even if the pencil $\lambda E - A$ is not d-stable but it has no eigenvalues on the unit circle. In this case the matrix X as in (5.41) is a solution of the generalized Lyapunov equation

$$\begin{aligned} & A^T X A - E^T X E = -P_{r,0}^T P_{r,0} + \xi (I - P_{r,0})^T (I - P_{r,0}), \\ & P_{l,0}^T X = X P_{l,0}, \end{aligned}$$

where $P_{l,0}$ and $P_{r,0}$ are the spectral projections onto the left and right deflating subspaces of $\lambda E - A$ corresponding to the eigenvalues inside the unit circle. The projection $P_{l,0}$ is computed as $P_{l,0} = \lim\limits_{k\to\infty} E_k^T(E_k+A_k)^{-T}$, see [7, 112]. The projection $P_{r,0}$ can be determined in the same way via iteration (5.40) with the starting matrices $E_0 = E$ and $A_0 = A$. Note that the left and right deflating subspaces of $\lambda E - A$ corresponding to the eigenvalues inside the unit circle can be computed without inverting the matrix $E_k + A_k$ explicitly, see [7] for details.

Also, one can use the Malyshev algorithm to solve the GCALE (4.9) with nonsingular E by applying to the Cayley-transformed pencil $\lambda(A - E) - (E + A)$. However, if the matrix E is singular, then by Proposition 4.34 the infinite eigenvalues of the pencil $\lambda E - A$ are mapped by the Cayley transformation to eigenvalues on the unit circle. In this case the Malyshev algorithm cannot be applied.

Perturbation theory, error analysis and parallelization issues for the Malyshev algorithm can be found in [7, 12, 62, 112]. A connection between this algorithm and the matrix sign function method is discussed in [12, 112].

The ADI and Smith methods

The *alternating direction implicit* (ADI) method was originally proposed for linear systems [124] and then applied in [109, 141, 167] to the continuous-time Lyapunov equation

$$A^T X + XA = -C^T C. \tag{5.42}$$

The ADI iteration can be written as

$$\begin{aligned}(A^T + p_k I)X_{k-1/2} &= -C^T C - X_{k-1}(A - p_k I),\\ (A^T + p_k I)X_k^T &= -C^T C - X_{k-1/2}^T(A - p_k I)\end{aligned}$$

with $X_0 = 0$ and the shift parameters $p_1, \ldots, p_k \in \mathbb{C}^-$. If all eigenvalues of the matrix A lie in the open left half-plane, then X_k converges to the solution of equation (5.42). The rate of convergence is determined by the spectral radius of the error transfer operator given by

$$\mathcal{T}_k(X) = \left(r_k(A)r_k(-A)^{-1}\right)^T X \left(r_k(A)r_k(-A)^{-1}\right),$$

where r_k is the polynomial $r_k(t) = (t - p_1) \cdot \ldots \cdot (t - p_k)$. The minimization of this spectral radius with respect to the parameters $p_1, \ldots, p_k$ leads to the ADI minimax problem

$$\{p_1, \ldots, p_k\} = \operatorname*{arg\,min}_{\{p_1,\ldots,p_k\}\in\mathbb{C}^-} \max_{t\in\mathrm{Sp}(A)} \frac{|r_k(t)|}{|r_k(-t)|}. \tag{5.43}$$

This problem is solved for equations with symmetric A, e.g. [168], while the case of complex eigenvalues is still under development, see [109, 125, 141, 142, 168] for some contributions.

The computational cost of the ADI method is, in general, $O(n^3)$. However, computations can be reduced by previously transformation of A to tridiagonal form [109]. The

ADI method is efficient for structured matrices and sparse matrices with small bandwidth [167].

For any real $p < 0$, equation (5.42) is equivalent to the discrete-time Lyapunov equation

$$\mathcal{A}^T X \mathcal{A} - X = -\mathcal{C}^T \mathcal{C}, \tag{5.44}$$

where $\mathcal{A} = (A - pI)(A + pI)^{-1}$ and $\mathcal{C} = \sqrt{-2p}\, C(A + pI)^{-1}$. It can be shown that if all the eigenvalues of the matrix $\mathcal{A}$ lie inside the unit circle or, equivalently, all the eigenvalues of A are in the open left half-plane, then the *Smith iteration*

$$X_0 = \mathcal{C}^T \mathcal{C}, \qquad X_{k+1} = \mathcal{C}^T \mathcal{C} + \mathcal{A}^T X_k \mathcal{A}$$

converges linearly to the solution X, see [139]. The quadratic convergence can be achieved by using the *squared Smith method* [139] based on the iteration

$$\begin{aligned} X_0 &= \mathcal{C}^T \mathcal{C}, & \mathcal{A}_0 &= \mathcal{A}, \\ X_{k+1} &= \mathcal{A}_k^T X_k \mathcal{A}_k, & \mathcal{A}_{k+1} &= \mathcal{A}_k^2. \end{aligned}$$

The number of iterations required for a desired accuracy in the approximate solution X_k of equation (5.42) depends on the parameter p. It should be noted that the Smith method is, in fact, the ADI iteration with a single parameter. Therefore, an optimal value $p = p_1 = \ldots = p_k$ from (5.43) can be used to increase the convergence.

The Smith method costs $O(n^3)$ flops and has just as the ADI method the memory complexity $O(n^2)$, since the solution X is computed explicitly and it is dense even if the coefficient matrix A is sparse. Note that in many cases the storage requirement rather than the computational cost is a limiting factor for feasibility of numerical methods for large scale problems.

Recently, efficient modifications of the ADI and Smith methods have been proposed to compute low-rank approximations for solutions of standard Lyapunov equations [106, 125, 127]. These are the *low-rank ADI iterate* and the *cyclic low-rank Smith method*. It was observed that the eigenvalues of the symmetric solutions of large scale Lyapunov equations with low-rank right-hand side generally decay very rapidly, see [5, 128]. This makes it possible to approximate such solutions by low-rank matrices.

The cyclic low-rank Smith method consists of two stages. First one computes

$$\begin{aligned} Z_1 &= \sqrt{-2p_1}(A^T + p_1 I)^{-1} C^T, \\ Z_k &= \left[\, (A^T - p_k I)(A^T + p_k I)^{-1} Z_{k-1},\ \sqrt{-2p_k}(A^T + p_k I)^{-1} C^T \,\right], \quad k = 1, \ldots, l, \end{aligned} \tag{5.45}$$

with the shift parameters $p_1, \ldots, p_k$ and then one iterates

$$\begin{aligned} Z^{(l)} &= Z_l, \\ Z^{((k+1)l)} &= \left(\prod_{j=1}^{l} (A^T - p_j I)(A^T + p_j I)^{-1} \right) Z^{(kl)}, \quad k = 1, 2, \ldots, \\ Z_{(k+1)l} &= \left[\, Z_{kl},\ Z^{(k+1)l} \,\right]. \end{aligned}$$

In this case a low-rank approximate solution of equation (5.42) is computed as $X = Z_{kl}Z_{kl}^T$. Note that the cyclic low-rank Smith method is equivalent to the low-rank ADI iterate with the cyclically repeated shift parameters $p_1, \ldots, p_l$, see [127].

Numerical aspects, area of application and computing of the shift parameters for the low-rank ADI and Smith methods for sparse problems are discussed in detail in [126, 127]. Some convergence results and improvements on the memory requirements for these methods can be found in [3, 106].

Krylov subspace methods

An alternative technique to compute low-rank approximate solutions of large scale sparse Lyapunov equations is the *full orthogonalization method* (FOM) and the *generalized minimum residual* (GMRES) method [44, 80, 136]. These methods are based on the calculation of an orthonormal basis $V_k \in \mathbb{R}^{n,k}$ of the Krylov subspace

$$\mathcal{K}_k(A^T, C^T) = \mathrm{Im}\,\left[\, C^T,\ A^TC^T,\ \ldots,\ (A^T)^{k-1}C^T \,\right]$$

via the block Arnoldi or Lanczos process [64, 80, 171] together with solving reduced order linear matrix equations.

In the FOM a low-rank approximate solution of the Lyapunov equation (5.42) is computed as $\hat{X} = V_kX_kV_k^T$, where $X_k \in \mathbb{R}^{k,k}$ satisfies the Galerkin condition

$$V_k^T\left(A^TV_kX_kV_k^T + V_kX_kV_k^TA + C^TC\right)V_k = 0.$$

To provide this condition we have to solve the reduced order Lyapunov equation

$$(V_k^TAV_k)^TX_k + X_k(V_k^TAV_k) = -V_k^TC^TCV_k.$$

This equation can be solved by using any direct method.

In the GMRES method one constructs an approximate solution $\hat{X} = V_kX_kV_k^T$, where $X_k \in \mathbb{R}^{k,k}$ satisfies the minimization problem

$$\|A^TV_kX_kV_k^T + V_kX_kV_k^TA + C^TC\|_F \to \min!.$$

This problem leads to a low order generalized Sylvester equation, see [80, 136] for details.

Note that the FOM and the GMRES can be used in a similar way to solve the discrete-time Lyapunov equation $A^TXA - X = -C^TC$ [80].

A drawback of the Krylov subspace methods is that they often converge slowly and relatively many iterates should be performed to determine the approximate solution with high accuracy. However, for increasing k the storage requirements to save the dense matrix V_k become excessive and the computing X_k gets expensive.

All iterative methods presented above for standard Lyapunov equations can also be used to solve generalized Lyapunov equations (4.9) and (4.43) with nonsingular E by applying to equations (4.15) and (4.47), respectively. However, if the matrix E is ill-conditioned this is not a numerically feasible approach to solve generalized Lyapunov equations. Moreover, in many applications E is sparse, whereas inverse of E may be dense. An extension of these methods to projected generalized Lyapunov equations is an open problem.

Chapter 6

Perturbation theory for generalized Lyapunov equations

There are several papers concerned with the perturbation theory and the backward error bounds for standard continuous-time Lyapunov equations, see [50, 57, 61, 74, 75] and the references therein. The sensitivity analysis for regular generalized Lyapunov equations has been presented in [96]. In this chapter we discuss the perturbation theory for projected generalized Lyapunov equations.

A condition number for a problem is an important characteristic to measure the sensitivity of the solution of this problem to perturbations in the original data and to bound errors in the approximate solution. If the condition number is large, then the problem is ill-conditioned in the sense that small perturbations in the data may lead to large variations in the solution.

The solution of the projected generalized Lyapunov equations is determined essentially in two steps that include first a computation of the deflating subspaces of a pencil corresponding to the finite and infinite eigenvalues due reduction to the GUPTRI form and solving the generalized Sylvester equation and then a calculation of the solution of the regular generalized Lyapunov equation. In such situation it may happen that although the projected generalized Lyapunov equation is well-conditioned, one of the intermediate problems may be ill-conditioned. This may lead to large inaccuracy in the numerical solution of the original problem. In this case we may conclude that either the combined numerical method is unstable or the solution is ill-conditioned, since it is a composition of two mappings one of which is ill-conditioned. Therefore, along with the conditioning of the projected GCALE (4.36) and the projected GDALE (4.64) we consider the perturbation theory for deflating subspaces, the generalized Sylvester equation (5.2), the regular Lyapunov equations (4.9) and (4.43).

6.1 Conditioning of deflating subspaces

The perturbation analysis for deflating subspaces of a regular pencil corresponding to the specified eigenvalues and error bounds have been presented in [40, 85, 86, 143, 145]. Here we briefly review the main results.

To compute the right and left deflating subspaces of the pencil $\lambda E - A$ corresponding to the finite eigenvalues we have to solve the generalized Sylvester equation (5.2). Consider a *Sylvester operator* $\mathcal{S} : \mathbb{F}^{n_f,2n_\infty} \to \mathbb{F}^{n_f,2n_\infty}$ given by

$$\mathcal{S}(Y,Z) := (E_f Y - Z E_\infty,\; A_f Y - Z A_\infty). \tag{6.1}$$

Then equation (5.2) can be written in the operator form $\mathcal{S}(Y,Z) = (E_u, A_u)$. Using the column-wise vector representation for the matrices Y and Z we rewrite (5.2) as a linear system

$$\mathbf{S}\begin{bmatrix} \mathrm{vec}(Y) \\ \mathrm{vec}(Z) \end{bmatrix} = -\begin{bmatrix} \mathrm{vec}(E_u) \\ \mathrm{vec}(A_u) \end{bmatrix}, \tag{6.2}$$

where the $(2n_f n_\infty \times 2n_f n_\infty)$-matrix

$$\mathbf{S} = \begin{bmatrix} I_{n_\infty} \otimes E_f & -E_\infty^T \otimes I_{n_f} \\ I_{n_\infty} \otimes A_f & -A_\infty^T \otimes I_{n_f} \end{bmatrix}$$

is the matrix representation of the Sylvester operator $\mathcal{S}$. The norm of $\mathcal{S}$ induced by the Frobenius matrix norm is given by

$$\|\mathcal{S}\|_F := \sup_{\|(Y,Z)\|_F=1} \|(E_f Y - Z E_\infty, A_f Y - Z A_\infty)\|_F = \|\mathbf{S}\|_2.$$

We define the *separation* of two regular pencils $\lambda E_f - A_f$ and $\lambda E_\infty - A_\infty$ as

$$\mathrm{Dif}_u \equiv \mathrm{Dif}_u(E_f, A_f; E_\infty, A_\infty) := \inf_{\|(Y,Z)\|_F=1} \|(E_f Y - Z E_\infty, A_f Y - Z A_\infty)\|_F = \sigma_{\min}(\mathbf{S}),$$

where $\sigma_{\min}(\mathbf{S})$ is the smallest singular value of $\mathbf{S}$ [143]. Note that $\mathrm{Dif}_u(E_\infty, A_\infty; E_f, A_f)$ does not, in general, equal $\mathrm{Dif}_u(E_f, A_f; E_\infty, A_\infty)$. Therefore, we set

$$\mathrm{Dif}_l \equiv \mathrm{Dif}_l(E_f, A_f; E_\infty, A_\infty) := \mathrm{Dif}_u(E_\infty, A_\infty; E_f, A_f).$$

The values Dif_u and Dif_l measure how close the spectra of $\lambda E_f - A_f$ and $\lambda E_\infty - A_\infty$ are. In other words, if there is a small perturbation of $\lambda E_f - A_f$ and $\lambda E_\infty - A_\infty$ such that the perturbed pencils have a common eigenvalue, then either Dif_u or Dif_l is small. However, small separations do not imply that the corresponding deflating subspaces are ill-conditioned [145].

Important quantities that measure the sensitivity of the right and left finite deflating subspaces of the pencil $\lambda E - A$ to perturbations in E and A are the norms of the spectral projections P_r and P_l. If $\|P_r\|_2$ (or $\|P_l\|_2$) is large, then the right (left) deflating subspace of $\lambda E - A$ corresponding to the finite eigenvalues is close to the right (left) deflating subspace corresponding to the infinite eigenvalues.

Let the pencil $\lambda E - A$ be in the GUPTRI form (2.4) and let the transformation matrices $U = [\,U_1,\ U_2\,]$ and $V = [\,V_1,\ V_2\,]$ be partitioned conformally to the blocks associated with the finite and infinite eigenvalues. In this case $\mathcal{U} = \operatorname{Im} U_1$ and $\mathcal{V} = \operatorname{Im} V_1$ are the right and left finite deflating subspaces of $\lambda E - A$, respectively, and they have dimension n_f. Consider a perturbed pencil $\lambda\widetilde{E} - \widetilde{A} = \lambda(E + \Delta E) - (A + \Delta A)$. Let $\widetilde{\mathcal{U}}$ and $\widetilde{\mathcal{V}}$ be, respectively, the right and left finite deflating subspaces of $\lambda\widetilde{E} - \widetilde{A}$ and suppose that they have the same dimensions as $\mathcal{U}$ and $\mathcal{V}$. The *distance between two subspaces* $\mathcal{U}$ and $\widetilde{\mathcal{U}}$ is defined as

$$\theta_{\max}(\mathcal{U}, \widetilde{\mathcal{U}}) = \max_{u \in \mathcal{U}} \min_{\widetilde{u} \in \widetilde{\mathcal{U}}} \theta(u, \widetilde{u}),$$

where $\theta(u, \widetilde{u})$ is the acute angle between the vectors u and $\widetilde{u}$. Then one has the following perturbation bounds for the deflating subspaces of the regular pencil $\lambda E - A$.

Theorem 6.1. [40] *Suppose that the right and left finite deflating subspaces of a regular pencil $\lambda E - A$ and a perturbed pencil $\lambda\widetilde{E} - \widetilde{A} = \lambda(E + \Delta E) - (A + \Delta A)$ corresponding to the finite eigenvalues have the same dimensions. If*

$$\|(\Delta E, \Delta A)\|_F < \frac{\min(\mathrm{Dif}_u, \mathrm{Dif}_l)}{\sqrt{\|P_l\|_2^2 + \|P_r\|_2^2 + 2\max(\|P_l\|_2, \|P_r\|_2)}} =: \wp,$$

then

$$\begin{aligned} \tan\theta_{\max}(\mathcal{U}, \widetilde{\mathcal{U}}) &\le \frac{\|(\Delta E, \Delta A)\|_F}{\wp\|P_r\|_2 - \|(\Delta E, \Delta A)\|_F \sqrt{\|P_r\|_2^2 - 1}} \\ &\le \|(\Delta E, \Delta A)\|_F \frac{\|P_r\|_2 + \sqrt{\|P_r\|_2^2 - 1}}{\wp} \end{aligned} \tag{6.3}$$

and

$$\begin{aligned} \tan\theta_{\max}(\mathcal{V}, \widetilde{\mathcal{V}}) &\le \frac{\|(\Delta E, \Delta A)\|_F}{\wp\|P_l\|_2 - \|(\Delta E, \Delta A)\|_F \sqrt{\|P_l\|_2^2 - 1}} \\ &\le \|(\Delta E, \Delta A)\|_F \frac{\|P_l\|_2 + \sqrt{\|P_l\|_2^2 - 1}}{\wp}. \end{aligned} \tag{6.4}$$

Bounds (6.3) and (6.4) imply that for small enough $\|(\Delta E, \Delta A)\|_F$, the right and left finite deflating subspaces of the perturbed pencil $\lambda\widetilde{E} - \widetilde{A}$ are small perturbations of the corresponding right and left deflating subspaces of $\lambda E - A$. Perturbation $\|(\Delta E, \Delta A)\|_F$ is bounded by $\wp$ which is small if the separations Dif_u and Dif_l are small or the norms $\|P_l\|_2$ and $\|P_r\|_2$ are large.

Thus, the quantities Dif_u, Dif_l, $\|P_l\|_2$ and $\|P_r\|_2$ can be used to characterize the conditioning of the right and left finite deflating subspaces of the pencil $\lambda E - A$.

From representations (5.3) and (5.4) for the spectral projections P_r and P_l we have

$$\|P_r\|_2 = \sqrt{1 + \|Y\|_2^2}, \qquad \|P_l\|_2 = \sqrt{1 + \|Z\|_2^2}, \tag{6.5}$$

where (Y, Z) is the solution of the generalized Sylvester equation (5.2). We see that the norms of Y and Z also characterize the sensitivity of the deflating subspaces. It follows from (6.2) that

$$\|(Y, Z)\|_F \le \mathrm{Dif}_u^{-1} \|(E_u, A_u)\|_F. \tag{6.6}$$

This estimate gives a connection between the separation Dif_u and the norm of the solution of the generalized Sylvester equation (5.2).

The perturbation analysis, condition numbers and error bounds for generalized Sylvester equations are presented in [84, 87]. Consider a perturbed generalized Sylvester equation

$$\begin{array}{rcl} (E_f + \Delta E_f)\widetilde{Y} - \widetilde{Z}(E_\infty + \Delta E_\infty) & = & -(E_u + \Delta E_u), \\ (A_f + \Delta A_f)\widetilde{Y} - \widetilde{Z}(A_\infty + \Delta A_\infty) & = & -(A_u + \Delta E_u), \end{array} \tag{6.7}$$

where the perturbations are measured norm-wise by

$$\epsilon = \max\left\{ \frac{\|(\Delta E_f, \Delta A_f)\|_F}{\alpha}, \ \frac{\|(\Delta E_\infty, \Delta A_\infty)\|_F}{\beta}, \ \frac{\|(\Delta E_u, \Delta A_u)\|_F}{\gamma} \right\} \tag{6.8}$$

with $\alpha = \|(E_f, A_f)\|_F$, $\beta = \|(E_\infty, A_\infty)\|_F$ and $\gamma = \|(E_u, A_u)\|_F$. Then one has the following first order relative perturbation bound for the solution of the generalized Sylvester equation (5.2).

Theorem 6.2. [84] *Let the perturbations in* (6.7) *satisfy* (6.8). *Assume that both the generalized Sylvester equations* (5.2) *and* (6.7) *are uniquely solvable. Then*

$$\frac{\|(\widetilde{Y}, \widetilde{Z}) - (Y, Z)\|_F}{\|(Y, Z)\|_F} \le \sqrt{3}\,\epsilon\, \frac{\|\mathbf{S}^{-1} M_S\|_2}{\|(Y, Z)\|_F}, \tag{6.9}$$

where the matrix M_S of size $2n_f n_\infty \times 2(nn_\infty + n_f^2)$ has the form $M_S = \mathrm{diag}(B_S, B_S)$ with $B_S = [\,\alpha(Y^T \otimes I_{n_f}), \ -\beta(I_{n_\infty} \otimes Z), \ \gamma I_{n_f n_\infty}\,]$.

The number

$$\kappa_{st} = \frac{\|\mathbf{S}^{-1} M_S\|_2}{\|(Y, Z)\|_F}$$

is called the *structured condition number* for the generalized Sylvester equation (5.2). Bound (6.9) shows that the relative error in the solution of the perturbed equation (6.7) is small if κ_{st} is not too large, i.e., if the problem is well-conditioned.

From (6.9) we obtain another relative error bound

$$\frac{\|(\widetilde{Y}, \widetilde{Z}) - (Y, Z)\|_F}{\|(Y, Z)\|_F} \le \sqrt{3}\,\epsilon\, \mathrm{Dif}_u^{-1} \frac{(\alpha + \beta)\,\|(Y, Z)\|_F + \gamma}{\|(Y, Z)\|_F}$$

that, in general, is worse than (6.9), since it does not take account of the special structure of perturbations in the generalized Sylvester equation (6.7).

We define the *condition number* for the generalized Sylvester equation (5.2) induced by the Frobenius norm as

$$\kappa := \Big(\|(E_f, A_f)\|_F^2 + \|(E_\infty, A_\infty)\|_F^2\Big)^{1/2} \mathrm{Dif}_u^{-1}.$$

Applying the standard linear system perturbation analysis [64] to (6.2) we obtain the following relative perturbation bounds.

Theorem 6.3. [87] *Suppose that the generalized Sylvester equation* (5.2) *has a unique solution* (Y, Z). *Let the perturbations in* (6.7) *satisfy* (6.8). *If* $\epsilon\kappa < 1$, *then the perturbed generalized Sylvester equation* (6.7) *has a unique solution* $(\widetilde{Y}, \widetilde{Z})$ *and*

$$\frac{\|(\widetilde{Y}, \widetilde{Z}) - (Y, Z)\|_F}{\|(Y, Z)\|_F} \leq \frac{\epsilon\,(\kappa\|(Y, Z)\|_F + \|(E_u, A_u)\|_F)}{(1 - \epsilon\,\kappa)\|(Y, Z)\|_F} \leq \frac{2\,\epsilon\,\kappa}{1 - \epsilon\,\kappa}. \tag{6.10}$$

Note that both the bounds in (6.10) may overestimate the true relative error in the solution, since they do not take into account the structured perturbations in the matrix $\mathbf{S}$. Nevertheless, quantities Dif_u^{-1} and κ are used in practice to characterize the conditioning of the generalized Sylvester equation (5.2).

The computation of $\mathrm{Dif}_u = \sigma_{\min}(\mathbf{S})$ is expensive even for modest n_f and n_∞, since the cost of computing the smallest singular value of the matrix $\mathbf{S}$ is $O(n_f^3 n_\infty^3)$ flops. It is more practical to compute lower bounds for Dif_u^{-1}, see [86, 87] for details. The Frobenius norm based Dif_u^{-1}-estimator can be computed by solving one generalized Sylvester equation in triangular form and costs $(2n_f^2 n_\infty + 2n_f n_\infty^2)$ flops. The one-norm based estimator is a factor 3 to 10 times more expensive and it does not differ more than a factor $\sqrt{2n_f n_\infty}$ from Dif_u^{-1} [86]. The computation of both these Dif_u^{-1}-estimators is implemented in the LAPACK subroutine `_TGSEN` [1].

6.2 Condition numbers for regular generalized Lyapunov equations

The perturbation theory and some useful condition numbers for the standard Lyapunov equations were presented in [50, 61, 74, 75], see also the references therein. The case of generalized Lyapunov equations with nonsingular E was considered in [95, 96, 113]. In this subsection we review some results from there.

Consider the regular GCALE (4.9). Let $\mathcal{L}_c$ be a continuous-time Lyapunov operator as in (4.10). The norm of $\mathcal{L}_c$ induced by the Frobenius matrix norm is computed via

$$\|\mathcal{L}_c\|_F := \sup_{\|X\|_F = 1} \|E^* X A + A^* X E\|_F = \|\mathbf{L}_c\|_2,$$

where $\mathbf{L}_c$ is as in (4.13). Analogously to the Sylvester equation, an important quantity in the sensitivity analysis for Lyapunov equations is the *separation* defined for the GCALE (4.9) by

$$\mathrm{Sep}_c(E, A) := \inf_{\|X\|_F=1} \|E^*XA + A^*XE\|_F = \sigma_{\min}(\mathbf{L}_c),$$

where $\sigma_{\min}(\mathbf{L}_c)$ is the smallest singular value of $\mathbf{L}_c$, see [55]. If the GCALE (4.9) is regular, then the Lyapunov operator $\mathcal{L}_c$ is invertible and the matrix $\mathbf{L}_c$ is nonsingular. The norm of the inverse $\mathcal{L}_c^{-1}$ induced by the Frobenius norm can be computed as

$$\|\mathcal{L}_c^{-1}\|_F = \|\mathbf{L}_c^{-1}\|_2 = \mathrm{Sep}_c^{-1}(E, A).$$

Consider a perturbed GCALE

$$(E + \Delta E)^*\widetilde{X}(A + \Delta A) + (A + \Delta A)^*\widetilde{X}(E + \Delta E) = -(G + \Delta G), \tag{6.11}$$

where

$$\|\Delta E\|_F \le \varepsilon_F, \qquad \|\Delta A\|_F \le \varepsilon_F, \qquad \|\Delta G\|_F \le \varepsilon_F, \qquad (\Delta G)^* = \Delta G. \tag{6.12}$$

Using the equivalent formulation (4.12) for the GCALE (4.9) we have the following perturbation estimate for the solution of (4.9) in the real case, see [96] for the complex case.

Theorem 6.4. [96] *Let E, A, $G \in \mathbb{R}^{n,n}$ and let G be symmetric. Assume that the GCALE* (4.9) *is regular. Let the absolute perturbations in the GCALE* (6.11) *satisfy* (6.12). *If*

$$\varepsilon_F \left(l_{c,E} + l_{c,A} + 2\varepsilon_F \mathrm{Sep}_c^{-1}(E, A) \right) < 1,$$

then the perturbed GCALE (6.11) *is regular and the norm-wise absolute perturbation bound*

$$\|\widetilde{X} - X\|_F \le \frac{\sqrt{3}\,\varepsilon_F \|\mathbf{L}_c^{-1} M_c\|_2 + 2\varepsilon_F^2 \mathrm{Sep}_c^{-1}(E, A)\|X\|_2}{1 - \varepsilon_F \left(l_{c,E} + l_{c,A} + 2\varepsilon_F \mathrm{Sep}_c^{-1}(E, A) \right)} \tag{6.13}$$

holds, where

$$\begin{aligned}
M_c &= \Big[(I_{n^2} + \Pi_{n^2}) \left(I_n \otimes (A^T X) \right), \quad (I_{n^2} + \Pi_{n^2}) \left(I_n \otimes (E^T X) \right), \quad I_{n^2} \Big], \\
l_{c,E} &= \left\| \mathbf{L}_c^{-1} (I_{n^2} + \Pi_{n^2}) \left(I_n \otimes A^T \right) \right\|_2, \\
l_{c,A} &= \left\| \mathbf{L}_c^{-1} (I_{n^2} + \Pi_{n^2}) \left(I_n \otimes E^T \right) \right\|_2.
\end{aligned}$$

The number

$$\kappa_{c,st}(E, A) = \frac{\|\mathbf{L}_c^{-1} M_c\|_2}{\|X\|_F}$$

is called the *structured condition number* for the GCALE (4.9). Bound (6.13) shows that if $\kappa_{c,st}(E, A)$, $\mathrm{Sep}_c^{-1}(E, A)$, $l_{c,E}$ and $l_{c,A}$ are not too large, then the solution of the perturbed GCALE (6.11) is a small perturbation of the solution of (4.9). Note that bound (6.13) is asymptotically sharp [96].

We define the *condition number* for the GCALE (4.9) induced by the Frobenius norm as

$$\kappa_{c,F}(E,A) := 2\|E\|_2\|A\|_2 \mathrm{Sep}_c^{-1}(E,A). \tag{6.14}$$

This condition number allows to obtain relative perturbation bounds for the solution of the GCALE (4.9).

Corollary 6.5. *Suppose that the GCALE* (4.9) *is regular. Let the perturbations in* (6.11) *satisfy* $\|\Delta E\|_2 \le \varepsilon\|E\|_2$, $\|\Delta A\|_2 \le \varepsilon\|A\|_2$ *and* $\|\Delta G\|_2 \le \varepsilon\|G\|_2$. *If* $\varepsilon\,(2+\varepsilon)\kappa_{c,F}(E,A) < 1$, *then the perturbed GCALE* (6.11) *is regular and*

$$\begin{aligned} \frac{\|\widetilde{X}-X\|_F}{\|X\|_F} &\le \frac{(2\varepsilon+\varepsilon^2)\kappa_{c,F}(E,A)\|X\|_F + \varepsilon\|G\|_2\mathrm{Sep}_c^{-1}(E,A)}{(1-\varepsilon(2+\varepsilon)\kappa_{c,F}(E,A)\,)\|X\|_F} \\ &\le \frac{\varepsilon(3+\varepsilon)\kappa_{c,F}(E,A)}{1-\varepsilon(2+\varepsilon)\kappa_{c,F}(E,A)}. \end{aligned} \tag{6.15}$$

Proof. The result immediately follows from Theorem 6.4. □

It should be noted that bounds (6.15) may overestimate the true relative error, since they do not take account of the specific structure of perturbations in (6.11). In the case of symmetric perturbations in G, sharp sensitivity estimates for general Lyapunov operators can be derived by using so-called Lyapunov singular values instead of standard singular values, see [95, 96] for details. Note that for the Lyapunov operator $\mathcal{L}_c$ as in (4.10), the Lyapunov singular values are equal to the standard singular values.

Let $\hat{X}$ be an approximate solution of the GCALE (4.9) and let

$$\mathcal{R}_c := E^*\hat{X}A + A^*\hat{X}E + G \tag{6.16}$$

be a *residual* of (4.9) corresponding to $\hat{X}$. Then from Corollary 6.5 we obtain the following forward error bound

$$\frac{\|\hat{X}-X\|_F}{\|X\|_F} \le \frac{\kappa_{c,F}(E,A)\,\|\mathcal{R}_c\|_F}{2\|E\|_2\|A\|_2\|X\|_F} =: Est_{c,F}. \tag{6.17}$$

This bound shows that for well-conditioned problems, a small relative residual implies a small error in the approximate solution $\hat{X}$. However, if the condition number $\kappa_{c,F}(E,A)$ is large, then $\hat{X}$ may be inaccurate even for a small residual.

It follows from bounds (6.15) and (6.17) that $\kappa_{c,F}(E,A)$ and $\mathrm{Sep}_c(E,A) = \sigma_{\min}(\mathbf{L}_c)$ can be used as a measure of the sensitivity of the solution of the regular GCALE (4.9). Since computing the smallest singular value of the $(n^2\times n^2)$-matrix $\mathbf{L}_c$ is not acceptable even for modest n, it is more practical to compute estimates for $\mathrm{Sep}_c^{-1}(E,A)$. A Sep_c^{-1}-estimator based on the one-norm differs from $\mathrm{Sep}_c^{-1}(E,A)$ at most by a factor n. Computing this estimator is implemented in the LAPACK subroutine `_LACON` [1] and costs $O(n^3)$ flops.

Unfortunately, if the matrix E is singular, then $\mathrm{Sep}_c(E,A) = 0$ and $\kappa_{c,F}(E,A) = \infty$. In this case we cannot use (6.14) as the condition number for the projected GCALE (4.36).

Consider now the regular GDALE (4.43). Let $\mathcal{L}_d$ be a discrete-time Lyapunov operator given in (4.46). Analogous to the continuous-time case, the *separation* for the GDALE (4.43) is defined by

$$\mathrm{Sep}_d(E, A) := \inf_{\|X\|_F=1} \|A^*XA - E^*XE\|_F = \sigma_{\min}(\mathbf{L}_d),$$

where the matrix $\mathbf{L}_d$ is as in (4.46). If the GDALE (4.43) is regular, then $\mathcal{L}_d$ is invertible and the matrix $\mathbf{L}_d$ is nonsingular. In this case we obtain that

$$\|\mathcal{L}_d^{-1}\|_F = \|\mathbf{L}_d^{-1}\|_2 = \mathrm{Sep}_d^{-1}(E, A).$$

There is a discrete-time analogue of Theorem 6.4 for the perturbed GDALE

$$(A + \Delta A)^*\widetilde{X}(A + \Delta A) - (E + \Delta E)^*\widetilde{X}(E + \Delta E) = -(G + \Delta G). \tag{6.18}$$

Theorem 6.6. [96] *Let E, A, $G \in \mathbb{R}^{n,n}$ and let G be symmetric. Suppose that the GDALE* (4.43) *is regular. Let the absolute perturbations in* (6.18) *satisfy* (6.12)*. If*

$$\varepsilon_F \left(l_{d,E} + l_{d,A} + 2\varepsilon_F \mathrm{Sep}_d^{-1}(E, A)\right) < 1,$$

then the perturbed GDALE (6.18) *is regular and*

$$\|\widetilde{X} - X\|_F \le \frac{\sqrt{3}\varepsilon_F\|L_d^{-1}M_d\|_2 + 2\varepsilon_F^2 \mathrm{Sep}_d^{-1}(E, A)\|X\|_2}{1 - \varepsilon_F\left(l_{d,E} + l_{d,A} + 2\varepsilon_F \mathrm{Sep}_d^{-1}(E, A)\right)}, \tag{6.19}$$

holds, where

$$\begin{aligned} M_d &= \Big[-(I_{n^2} + \Pi_{n^2})\left(I_n \otimes (E^TX)\right), \quad (I_{n^2} + \Pi_{n^2})\left(I_n \otimes (A^TX)\right), \quad I_{n^2} \Big], \\ l_{d,E} &= \left\|\mathbf{L}_d^{-1}(I_{n^2} + \Pi_{n^2})\left(I_n \otimes E^T\right)\right\|_2 \\ l_{d,A} &= \left\|\mathbf{L}_d^{-1}(I_{n^2} + \Pi_{n^2})\left(I_n \otimes A^T\right)\right\|_2. \end{aligned}$$

The number

$$\kappa_{d,st} = \frac{\|\mathbf{L}_d^{-1}M_d\|_2}{\|X\|_F}$$

is called the *structured condition number* for the regular GDALE (4.43).

Similar to the continuous-time case, we define the *condition number* for the GDALE (4.43) induced by the Frobenius norm as

$$\kappa_{d,F}(E, A) := \left(\|E\|_2^2 + \|A\|_2^2\right) \mathrm{Sep}_d^{-1}(E, A). \tag{6.20}$$

From Theorem 6.6 we obtain the following relative perturbation bounds for the solution of the GDALE (4.43).

Corollary 6.7. *Let the GDALE* (4.43) *be regular. Suppose that the perturbations in* (6.18) *satisfy* $\|\Delta E\|_2 \le \varepsilon \|E\|_2$, $\|\Delta A\|_2 \le \varepsilon \|A\|_2$ *and* $\|\Delta G\|_2 \le \varepsilon \|G\|_2$. *If* $\varepsilon(2+\varepsilon)\kappa_{d,F}(E,A) < 1$, *then the perturbed GDALE* (6.18) *is regular and*

$$\begin{aligned}\frac{\|\widetilde{X}-X\|_F}{\|X\|_F} &\le \frac{(2\varepsilon+\varepsilon^2)\kappa_{d,F}(E,A)+\varepsilon \mathrm{Sep}_d^{-1}(E,A)\|G\|_2}{(1-\varepsilon(2+\varepsilon)\kappa_{d,F}(E,A))\|X\|_F} \\ &\le \frac{\varepsilon(3+\varepsilon)\kappa_{d,F}(E,A)}{1-\varepsilon(2+\varepsilon)\kappa_{d,F}(E,A)}.\end{aligned} \tag{6.21}$$

Let $\hat{X}$ be an approximate solution of the GDALE (4.43). A *residual* of (4.43) corresponding to $\hat{X}$ is defined by

$$\mathcal{R}_d := A^*\hat{X}A - E^*\hat{X}E + G. \tag{6.22}$$

By Corollary 6.7 we have the following forward error estimate

$$\frac{\|\hat{X}-X\|_F}{\|X\|_F} \le \frac{\kappa_{d,F}(E,A)\,\|\mathcal{R}_d\|_F}{(\|E\|_2^2+\|A\|_2^2)\|X\|_F} =: Est_{d,F}.$$

This bound shows that if the GDALE (4.43) is well-conditioned and if the relative residual is small, then the error in the approximate solution $\hat{X}$ is also small. However, for ill-conditioned problems, $\hat{X}$ may be inaccurate even if the residual is small.

Thus, $\mathrm{Sep}_d(E,A)$ and $\kappa_{d,F}(E,A)$ can be used to measure the sensitivity of the solution of the regular GDALE (4.43) to perturbations in the data. However, in the case when both the matrices E and A are singular we obtain that $\mathrm{Sep}_d^{-1}(E,A)=\infty$. Thus, it is impossible to use $\kappa_{d,F}(E,A)$ as the condition number for the projected GDALE (4.64).

In [50, 61, 74] condition numbers based on the spectral norm have been used as a measure of sensitivity of the standard continuous-time and discrete-time Lyapunov equations. In the following subsections we extend this idea to the projected GCALE (4.36) and the projected GDALE (4.64).

6.3 Conditioning of the projected GCALE

Assume that the pencil $\lambda E - A$ is c-stable. Consider the matrix H_c as in (3.16). Using the Parseval identity [135], we obtain the integral representation

$$H_c = \frac{1}{2\pi}\int_{-\infty}^{\infty}(i\omega E - A)^{-*}P_r^*P_r(i\omega E-A)^{-1}d\omega. \tag{6.23}$$

Consider a linear operator $\mathcal{L}_c^- : \mathbb{F}^{n,n} \to \mathbb{F}^{n,n}$ defined as follows: for a matrix G, the image $X = -\mathcal{L}_c^-(G)$ is the unique solution of the projected GCALE (4.36). Note that the operator $\mathcal{L}_c^-$ is a (2)-pseudoinverse [32] of the Lyapunov operator $\mathcal{L}_c$, since it satisfies $\mathcal{L}_c^-\mathcal{L}_c\mathcal{L}_c^- = \mathcal{L}_c^-$.

Lemma 6.8. *Let $\lambda E - A$ be c-stable. Then $\|\mathcal{L}_c^-\|_2 = \|H_c\|_2$.*

Proof. Let u and v be the left and right singular vectors of unit length corresponding to the largest singular value of the solution X of the projected GCALE (4.36) with some matrix G. Then

$$\begin{aligned}\|\mathcal{L}_c^-(G)\|_2 &= \|X\|_2 = u^*Xv = \frac{1}{2\pi}\int_{-\infty}^{\infty} u^*(i\omega E - A)^{-*}P_r^*GP_r(i\omega E - A)^{-1}v\,d\omega \\ &\leq \frac{1}{2\pi}\|G\|_2\int_{-\infty}^{\infty} \left\|P_r(i\omega E - A)^{-1}u\right\|_2 \left\|P_r(i\omega E - A)^{-1}v\right\|_2\,d\omega.\end{aligned}$$

Using the Cauchy-Schwarz inequality [89] and (6.23) we obtain

$$\begin{aligned}\|\mathcal{L}_c^-(G)\|_2 &\leq \frac{1}{2\pi}\|G\|_2\left(\int_{-\infty}^{\infty} \left\|P_r(i\omega E - A)^{-1}u\right\|_2^2\,d\omega\right)^{\frac{1}{2}}\left(\int_{-\infty}^{\infty} \left\|P_r(i\omega E - A)^{-1}v\right\|_2^2\,d\omega\right)^{\frac{1}{2}} \\ &\leq \|G\|_2\left\|\frac{1}{2\pi}\int_{-\infty}^{\infty}(i\omega E - A)^{-*}P_r^*P_r(i\omega E - A)^{-1}\,d\omega\right\|_2 = \|G\|_2\|H_c\|_2.\end{aligned}$$

Hence, $\|\mathcal{L}_c^-\|_2 \leq \|H_c\|_2$.

On the other hand, we have

$$\|\mathcal{L}_c^-\|_2 = \sup_{\|G\|_2=1}\|\mathcal{L}_c^-(-G)\|_2 \geq \|\mathcal{L}_c^-(-I)\|_2 = \|H_c\|_2.$$

Thus, $\|\mathcal{L}_c^-\|_2 = \|H_c\|_2$. □

If E is nonsingular, then $\mathcal{L}_c^- = \mathcal{L}_c^{-1}$ is the inverse of the Lyapunov operator $\mathcal{L}_c$ and $\|\mathcal{L}_c^{-1}\|_2 = \|H_c\|_2$.

By Corollary 4.15 the matrix H_c is the unique Hermitian, positive semidefinite solution of the projected GCALE

$$E^*H_cA + A^*H_cE = -P_r^*P_r, \qquad H_c = H_cP_l. \tag{6.24}$$

We define the *spectral condition number* for the projected GCALE (4.36) as

$$\kappa_{c,2}(E,A) := 2\|E\|_2\|A\|_2\|H_c\|_2.$$

In Section 3.1.2 we have seen that the parameter $\kappa_{c,2}(E,A)$ is closely related to the analysis of the asymptotic behavior of solutions of the continuous-time singular system (3.13). Here we will show that $\kappa_{c,2}(E,A)$ can also be used to estimate the distance from the finite eigenvalues of a c-stable pencil $\lambda E - A$ to the imaginary axis as well as to measure the sensitivity of the solution of the projected GCALE (4.36).

Theorem 6.9. *Let $\lambda E - A$ be a c-stable pencil. Then all finite eigenvalues of $\lambda E - A$ lie in the closed half-plane*

$$\left\{ z \in \mathbb{C} \quad : \quad \Re e(z) \leq -\frac{\|A\|_2}{\|E\|_2\kappa_{c,2}(E,A)} \right\}.$$

Moreover, for all $\omega \in \mathbb{R}$, the estimate

$$\|P_r(i\omega E - A)^{-1}\|_2 \leq \frac{5\pi\kappa_{c,2}(E,A)}{2\|A\|_2} \tag{6.25}$$

holds.

Proof. Let λ_0 be a finite eigenvalues of the pencil $\lambda E - A$ and $v \in \operatorname{Im} P_r$ be an eigenvector corresponding to λ_0. Then from the projected GCALE (6.24) we have

$$-\|v\|^2 = -\|P_r v\|^2 = v^*(E^* H_c A + A^* H_c E)v = 2\Re e(\lambda_0) v^* E^* H_c E v.$$

Hence,

$$\Re e(\lambda_0) = -\frac{\|v\|^2}{2v^* E^* H_c E v} \leq -\frac{1}{2\|E\|_2^2 \|H_c\|_2} = -\frac{\|A\|_2}{\|E\|_2 \kappa_{c,2}(E,A)}.$$

To prove (6.25), consider the integral representation (6.23) for the matrix H_c. For any vector v of unit length we obtain that

$$\frac{\kappa_{c,2}(E,A)}{2\|E\|_2\|A\|_2} = \|H_c\|_2 \geq \frac{1}{2\pi}\int_{-\infty}^{\infty} \|P_r(i\omega E - A)^{-1} v\|^2 d\omega. \tag{6.26}$$

Let ω_0 be a point on the real line where the norm $\|P_r(i\omega E - A)^{-1}\|_2$ achieves its maximal value. Using the relation

$$P_r(i\omega E - A)^{-1} E = P_r(i\omega E - A)^{-1} E P_r = (i\omega E - A)^{-1} E P_r$$

we obtain

$$P_r(i\omega E - A)^{-1} = P_r(i\omega_0 E - A)^{-1} - i(\omega - \omega_0) P_r(i\omega E - A)^{-1} E P_r(i\omega_0 E - A)^{-1}.$$

Then we have the estimate

$$\|P_r(i\omega E - A)^{-1}\|_2 \leq \frac{\|P_r(i\omega_0 E - A)^{-1}\|_2}{1 - |\omega - \omega_0|\|E\|_2\|P_r(i\omega_0 E - A)^{-1}\|_2}$$

which is valid for all ω such that $|\omega - \omega_0|\|E\|_2\|P_r(i\omega_0 E - A)^{-1}\|_2 < 1$. Furthermore, choosing a vector v such that $\|P_r(i\omega_0 E - A)^{-1} v\| = \|P_r(i\omega_0 E - A)^{-1}\|_2$, we obtain that

$$\begin{aligned} \|P_r(i\omega E - A)^{-1} v\| &\geq \|P_r(i\omega_0 E - A)^{-1} v\| \left(1 - |\omega - \omega_0|\|E\|_2\|P_r(i\omega E - A)^{-1}\|_2\right) \\ &\geq \|P_r(i\omega_0 E - A)^{-1}\|_2 \frac{1 - 2|\omega - \omega_0|\|E\|_2\|P_r(i\omega_0 E - A)^{-1}\|_2}{1 - |\omega - \omega_0|\|E\|_2\|P_r(i\omega_0 E - A)^{-1}\|_2}. \end{aligned}$$

Setting $\tau = \|E\|_2\|P_r(i\omega_0 E - A)^{-1}\|_2$, we get from (6.26) that

$$\begin{aligned} \frac{\pi\|E\|_2\kappa_{c,2}(E,A)}{\|A\|_2} &\geq \int_{\omega_0 - \frac{1}{2\tau}}^{\omega_0 + \frac{1}{2\tau}} \tau^2 \left(\frac{1 - 2|\omega - \omega_0|\tau}{1 - |\omega - \omega_0|\tau}\right)^2 d\omega \\ &= 2\tau \int_0^{1/2} \left(\frac{1 - 2t}{1 - t}\right)^2 dt = 2\tau(3 - 4\ln 2) \geq \frac{2\tau}{5}. \end{aligned}$$

Therefore, (6.25) holds. □

Bound (6.25) implies that the finite eigenvalues of the c-stable pencil $\lambda E - A$ are separated from the imaginary axis by a distance not less than $2\|A\|_2/(5\pi\kappa_{c,2}(E,A))$. In other words, (6.25) yields a lower bound for perturbations which preserve the dimension of the finite deflating subspace of $\lambda E - A$ and cause the pencil to obtain a finite eigenvalue on the imaginary axis. Thus, the parameter $\kappa_{c,2}(E,A)$ characterizes the absence of eigenvalues of the pencil $\lambda E - A$ not only on the imaginary axis but in a neighbourhood of it. This result generalizes the matrix case ($E = I$), see [22, 23, 62], for matrix pencils.

To measure the smallest real (complex) perturbation to a stable matrix required to make the perturbed matrix unstable, the real (complex) stability radius can be used [77, 163]. For numerical methods for the computation of the stability radius see, e.g., [26, 73, 130] and the references therein. Unfortunately, these results are not immediately applicable to matrix pencils. The general problem to measure or estimate the distance to instability for the pencil, i.e., the distance from the given pencil to the "nearest" pencil that is singular or has an eigenvalue in the closed right half-plane, is more difficult. Only partial solutions are known. A lower bound for the stability radius for the pencil $\lambda E - A$, allowing perturbations in A only, is given in [131]. A computable expression for the stability radius for the regular pencil of index less than or equal to one is studied in [28]. Computationally attractive upper and lower bounds for smallest norm de-regularizing perturbation are discussed in [27].

Consider now a perturbed pencil $\lambda\widetilde{E} - \widetilde{A} = \lambda(E + \Delta E) - (A + \Delta A)$ with $\|\Delta E\|_2 \le \varepsilon\|E\|_2$ and $\|\Delta A\|_2 \le \varepsilon\|A\|_2$. Assume that the dimension of the right and left deflating subspaces of $\lambda E - A$ corresponding to the infinite eigenvalues is not changed under perturbations. In many practical applications this is justified [28]. Consider, for example, semi-explicit differential-algebraic equations

$$\begin{aligned} \dot{x}_1(t) &= A_{11}x_1(t) + A_{12}x_2(t) + B_1u(t), && (6.27)\\ 0 &= A_{21}x_1(t) + A_{22}x_2(t) + B_2u(t). && (6.28) \end{aligned}$$

Equation (6.27) describes the dynamic behavior of the system, while equation (6.28) gives algebraic constraints on the states. Obviously, it is unreasonable to consider perturbations which cause the algebraic constraints to become differential.

Note that in the study of asymptotic stability for the differential-algebraic equation (3.13) it is allowed for the index of the pencil $\lambda E - A$ to be changed under perturbations. It is important only that finite eigenvalues stay finite and infinite eigenvalues must stay infinite. However, the perturbation analysis in this case is very complicated. We will deal only with perturbations which preserve the nilpotency structure of the pencil $\lambda E - A$, i.e., the right and left infinite deflating subspaces of $\lambda E - A$ are not changed. This condition can be written as

$$\ker P_r = \ker \widetilde{P}_r, \qquad \ker P_l = \ker \widetilde{P}_l, \tag{6.29}$$

where $\widetilde{P}_r$ and $\widetilde{P}_l$ are the spectral projections onto the right and left finite deflating subspaces of the perturbed pencil $\lambda\widetilde{E} - \widetilde{A}$. Moreover, we will assume for such allowable perturbations that we have an error bound $\|\widetilde{P}_r - P_r\|_2 \le \varepsilon K$ with some constant K (for such estimate for the pencil $\lambda E - A$ of index one, see [147]). This estimate implies that the right finite deflating subspace of the perturbed pencil $\lambda\widetilde{E} - \widetilde{A}$ is close to the right finite deflating subspace of $\lambda E - A$.

Consider now the perturbed projected GCALE

$$\widetilde{E}^*\widetilde{X}\widetilde{A} + \widetilde{A}^*\widetilde{X}\widetilde{E} = -\widetilde{P}_r^*\widetilde{G}\widetilde{P}_r, \qquad \widetilde{X} = \widetilde{X}\widetilde{P}_l. \tag{6.30}$$

The following theorem gives a relative error bound for the solution of (4.36).

Theorem 6.10. *Let $\lambda E - A$ be a c-stable pencil and let X satisfy the projected GCALE* (4.36). *Consider a perturbed pencil $\lambda\widetilde{E} - \widetilde{A} = \lambda(E + \Delta E) - (A + \Delta A)$ with $\|\Delta E\|_2 \le \varepsilon\|E\|_2$ and $\|\Delta A\|_2 \le \varepsilon\|A\|_2$. Assume that for the spectral projections $\widetilde{P}_r$ and $\widetilde{P}_l$ onto the right and left deflating subspaces corresponding to the finite eigenvalues of $\lambda\widetilde{E} - \widetilde{A}$, relations* (6.29) *are satisfied and a bound $\|\widetilde{P}_r - P_r\|_2 \le \varepsilon K < 1$ holds with some constant K. Let $\widetilde{G}$ be a perturbation of G such that $\|\Delta G\|_2 \le \varepsilon\|G\|_2$. If $\varepsilon(2 + \varepsilon)\kappa_{c,2}(E, A) < 1$, then the perturbed projected GCALE* (6.30) *has a unique solution $\widetilde{X}$ and*

$$\frac{\|\widetilde{X} - X\|_2}{\|X\|_2} \le \frac{\varepsilon\Big((\varepsilon K + \|P_r\|_2)(K + \|P_r\|_2)\|G\|_2 + 3\|E\|_2\|A\|_2\|X\|_2\Big)\kappa_{c,2}(E, A)}{\|E\|_2\|A\|_2\|X\|_2(1 - \varepsilon(2 + \varepsilon)\kappa_{c,2}(E, A))}. \tag{6.31}$$

Proof. It follows from (6.29) that

$$\widetilde{P}_r P_r = \widetilde{P}_r, \qquad P_r\widetilde{P}_r = P_r, \qquad \widetilde{P}_l P_l = \widetilde{P}_l, \qquad P_l\widetilde{P}_l = P_l. \tag{6.32}$$

The perturbed GCALE in (6.30) can be rewritten as

$$E^*\widetilde{X}A + A^*\widetilde{X}E = -\left(\widetilde{P}_r^*\widetilde{G}\widetilde{P}_r + \Delta\mathcal{L}_c(\widetilde{X})\right),$$

where $\Delta\mathcal{L}_c(\widetilde{X}) = (\Delta E)^*\widetilde{X}\widetilde{A} + E^*\widetilde{X}\Delta A + (\Delta A)^*\widetilde{X}E + \widetilde{A}^*\widetilde{X}\Delta E$. Using (2.4) and (5.2) we can verify that $P_l E = P_l E P_r = E P_r$ and $P_l A = P_l A P_r = A P_r$. Analogous relations hold for the perturbed pencil $\lambda\widetilde{E} - \widetilde{A}$. Then by (6.32) we obtain that $\widetilde{X} = \widetilde{X}P_l = \widetilde{X}P_l\widetilde{P}_l = \widetilde{X}\widetilde{P}_l$ and

$$\begin{aligned}
\widetilde{X}E &= \widetilde{X}P_l E = \widetilde{X}EP_r = \widetilde{X}P_l E P_r\widetilde{P}_r = \widetilde{X}E\widetilde{P}_r,\\
\widetilde{X}\widetilde{E} &= \widetilde{X}\widetilde{P}_l\widetilde{E} = \widetilde{X}\widetilde{E}\widetilde{P}_r = \widetilde{X}\widetilde{P}_l\widetilde{E}\widetilde{P}_r P_r = \widetilde{X}\widetilde{E}P_r.
\end{aligned}$$

These relationships remain valid if we replace E by A and $\widetilde{E}$ by $\widetilde{A}$. Combining these relations we obtain

$$\widetilde{P}_r^*\widetilde{G}\widetilde{P}_r + \Delta\mathcal{L}_c(\widetilde{X}) = P_r^*\left(\widetilde{P}_r^*\widetilde{G}\widetilde{P}_r + \Delta\mathcal{L}_c(\widetilde{X})\right)P_r = \widetilde{P}_r^*\left(\widetilde{P}_r^*\widetilde{G}\widetilde{P}_r + \Delta\mathcal{L}_c(\widetilde{X})\right)\widetilde{P}_r. \tag{6.33}$$

Then the perturbed projected GCALE (6.30) is equivalent to the projected GCALE

$$E^*\widetilde{X}A + A^*\widetilde{X}E = -P_r^*\left(\widetilde{P}_r^*\widetilde{G}\widetilde{P}_r + \Delta\mathcal{L}_c(\widetilde{X})\right)P_r, \qquad \widetilde{X} = \widetilde{X}P_l. \tag{6.34}$$

Since the pencil $\lambda E - A$ is c-stable, this equation has a unique solution given by

$$\widetilde{X} = \frac{1}{2\pi}\int_{-\infty}^{\infty}(i\omega E - A)^{-*}P_r^*\left(\widetilde{P}_r^*\widetilde{G}\widetilde{P}_r + \Delta\mathcal{L}_c(\widetilde{X})\right)P_r(i\omega E - A)^{-1}d\omega. \tag{6.35}$$

Thus, we have an integral equation $\widetilde{X} = \mathcal{I}(\widetilde{X})$ for the unknown matrix $\widetilde{X}$, where

$$\mathcal{I}(\widetilde{X}) = \frac{1}{2\pi}\int_{-\infty}^{\infty}(i\omega E - A)^{-*}P_r^*\left(\widetilde{P}_r^*\widetilde{G}\widetilde{P}_r + \Delta\mathcal{L}_c(\widetilde{X})\right)P_r(i\omega E - A)^{-1}d\omega.$$

From $\|\Delta\mathcal{L}_c(\widetilde{X})\|_2 \le 2(\|\Delta E\|_2\|\widetilde{A}\|_2 + \|\Delta A\|_2\|E\|_2)\|\widetilde{X}\|_2 \le 2\varepsilon(2+\varepsilon)\|E\|_2\|A\|_2\|\widetilde{X}\|_2$, we obtain for any matrices X_1 and X_2, that

$$\begin{aligned}\|\mathcal{I}(X_1) - \mathcal{I}(X_2)\|_2 &= \left\|\frac{1}{2\pi}\int_{-\infty}^{\infty}(i\omega E - A)^{-*}P_r^*\Delta\mathcal{L}_c(X_1 - X_2)P_r(i\omega E - A)^{-1}\,d\omega\right\|_2 \\ &\le \|\Delta\mathcal{L}_c(X_1 - X_2)\|_2\|H_c\|_2 \le \varepsilon(2+\varepsilon)\kappa_{c,2}(E,A)\|X_1 - X_2\|_2.\end{aligned}$$

Since $\varepsilon(2+\varepsilon)\kappa_{c,2}(E,A) < 1$, the operator $\mathcal{I}$ is contractive. Then by the fixed point theorem [89] the equation $\widetilde{X} = \mathcal{I}(\widetilde{X})$ has a unique solution $\widetilde{X}$ and we can estimate the error

$$\begin{aligned}\|\widetilde{X} - X\|_2 &= \left\|\frac{1}{2\pi}\int_{-\infty}^{\infty}(i\omega E - A)^{-*}P_r^*\Big(\widetilde{P}_r^*\widetilde{G}\widetilde{P}_r + \Delta\mathcal{L}_c(\widetilde{X}) - P_r^*GP_r\Big)P_r(i\omega E - A)^{-1}d\omega\right\|_2 \\ &\le \left(\|\widetilde{P}_r^*\widetilde{G}\widetilde{P}_r - P_r^*GP_r\|_2 + \|\Delta\mathcal{L}_c(\widetilde{X})\|_2\right)\|H_c\|_2.\end{aligned}$$

Taking into account that

$$\begin{aligned}\|\widetilde{P}_r^*\widetilde{G}\widetilde{P}_r - P_r^*GP_r\|_2 &\le \|\widetilde{P}_r - P_r\|_2(\|\widetilde{G}\|_2\|\widetilde{P}_r\|_2 + \|P_r\|_2\|G\|_2) + \|P_r\|_2\|\widetilde{G} - G\|_2\|\widetilde{P}_r\|_2 \\ &\le \varepsilon\left((\varepsilon K + \|P_r\|_2)((1+\varepsilon)K + \|P_r\|_2) + \varepsilon K\|P_r\|_2\right)\|G\|_2 \\ &\le 2\varepsilon\left(\varepsilon K + \|P_r\|_2\right)(K + \|P_r\|_2)\|G\|_2\end{aligned}$$

and $\|\Delta\mathcal{L}_c(\widetilde{X})\|_2 \le 2\varepsilon(2+\varepsilon)\|E\|_2\|A\|_2(\|X\|_2 + \|\widetilde{X} - X\|_2)$ we obtain the relative perturbation bound (6.31). □

Bound (6.31) implies that if perturbations in (6.30) satisfy (6.29) and if $\kappa_{c,2}(E,A)$, K and $\|P_r\|_2$ are not too large, then the solution of the perturbed projected GCALE (6.30) is a small perturbation of the solution of the projected GCALE (4.36).

Thus, $\kappa_{c,2}(E,A)$ can be used to characterize the sensitivity of the solution of the projected GCALE (4.36) to perturbations in the input data. To compute $\kappa_{c,2}(E,A)$ we need to solve the projected GCALE (6.24). The solution H_c of this equation can be calculated via the generalized Schur-Bartels-Stewart method or the generalized Schur-Hammarling methods presented in Sections 5.1 and 5.2.

From Theorem 6.10 we can obtain some useful consequences.

Corollary 6.11. *Under the assumptions of Theorem* 6.10 *we have that if the matrix G is Hermitian, positive definite and if*

$$2\varepsilon\left(2(1+2\varepsilon)(\varepsilon K+\|P_r\|_2)^2+1\right)\kappa_{c,2}(E,A)\|G\|_2<\lambda_{\min}(G), \tag{6.36}$$

where $\lambda_{\min}(G)$ is the smallest eigenvalue of G, then the perturbed pencil $\lambda\widetilde{E}-\widetilde{A}$ is c-stable and the following relative perturbation bound

$$\frac{|\kappa_{c,2}(\widetilde{E},\widetilde{A})-\kappa_{c,2}(E,A)|}{\kappa_{c,2}(E,A)}\le\frac{3\varepsilon\left(K(K+2\|P_r\|_2)+\kappa_{c,2}(E,A)+1\right)}{1-\varepsilon(2+\varepsilon)\kappa_{c,2}(E,A)} \tag{6.37}$$

holds.

Proof. First we will show that the matrix $\widetilde{P}_r^*\widetilde{G}\widetilde{P}_r+\Delta\mathcal{L}_c(\widetilde{X})$ is positive definite on the subspace $\operatorname{Im}P_r$. For all nonzero $v\in\operatorname{Im}P_r$, we have

$$\begin{array}{rcl}((\widetilde{P}_r^*\widetilde{G}\widetilde{P}_r+\Delta\mathcal{L}_c(\widetilde{X}))v,v) & = & ((\widetilde{P}_r^*(G+\Delta G)\widetilde{P}_r+\widetilde{P}_r^*\Delta\mathcal{L}_c(\widetilde{X})\widetilde{P}_r)v,v)\\ & \ge & \left(\lambda_{\min}(G)-\|\Delta\mathcal{L}_c(\widetilde{X})\|_2-\|\Delta G\|_2\right)\|\widetilde{P}_rv\|^2.\end{array} \tag{6.38}$$

It follows from (6.35) that

$$\|\widetilde{X}\|_2\le\frac{\|\widetilde{P}_r\|_2^2\|\widetilde{G}\|_2\|H_c\|_2}{1-\varepsilon(2+\varepsilon)\kappa_{c,2}(E,A)}\le\frac{(1+\varepsilon)(\varepsilon K+\|P_r\|_2)^2\|G\|_2\|H_c\|_2}{1-\varepsilon(2+\varepsilon)\kappa_{c,2}(E,A)}. \tag{6.39}$$

Then taking into account estimate (6.36) we get

$$\|\Delta\mathcal{L}_c(\widetilde{X})\|_2+\|\Delta G\|_2\le\frac{\varepsilon\left(2(1+2\varepsilon)(\varepsilon K+\|P_r\|_2)^2+1\right)\kappa_{c,2}(E,A)\|G\|_2}{1-\varepsilon(2+\varepsilon)\kappa_{c,2}(E,A)}<\lambda_{\min}(G).$$

Since $\widetilde{P}_rv\neq 0$, we have from (6.38) that $((\widetilde{P}_r^*\widetilde{G}\widetilde{P}_r+\Delta\mathcal{L}_c(\widetilde{X}))v,v)>0$ for all nonzero $v\in\operatorname{Im}P_r$, i.e., the matrix $\widetilde{P}_r^*\widetilde{G}\widetilde{P}_r+\Delta\mathcal{L}_c(\widetilde{X})$ is positive definite on the subspace $\operatorname{im}P_r$. Hence, by Corollary 4.15 the solution $\widetilde{X}$ of the projected GCALE (6.34) is positive semidefinite. Moreover, (6.36) yields that the matrix $\widetilde{G}$ is positive definite. Applying now Corollary 4.14 to the perturbed projected GCALE (6.30) we obtain that the pencil $\lambda\widetilde{E}-\widetilde{A}$ is c-stable.

From the proof of Theorem 6.10 with $\widetilde{G}=G=I$ it follows that

$$\|\widetilde{H}_c-H_c\|_2\le\frac{\varepsilon\left(K(\varepsilon K+2\|P_r\|_2)+(2+\varepsilon)\kappa_{c,2}(E,A)\right)\|H_c\|_2}{1-\varepsilon(2+\varepsilon)\kappa_{c,2}(E,A)},$$

where $\widetilde{H}_c$ is the solution of the perturbed projected GCALE (6.30) with $\widetilde{G}=I$. Then

$$\begin{array}{rcl}|\kappa_{c,2}(\widetilde{E},\widetilde{A}) & - & \kappa_{c,2}(E,A)|=2\left|\,\|\widetilde{E}\|_2\|\widetilde{A}\|_2\|\widetilde{H}_c\|_2-\|E\|_2\|A\|_2\|H_c\|_2\,\right|\\ & \le & 2\left(\|\widetilde{E}\|_2\|\widetilde{A}\|_2\|\widetilde{H}_c-H_c\|_2+\|\widetilde{E}-E\|\|\widetilde{A}\|_2\|H_c\|_2+\|E\|_2\|\widetilde{A}-A\|_2\|H_c\|_2\right)\\ & \le & \dfrac{3\varepsilon\kappa_{c,2}(E,A)\ \left(K(K+2\|P_r\|_2)+\kappa_{c,2}(E,A)+1\right)}{1-\varepsilon(2+\varepsilon)\kappa_{c,2}(E,A)}.\end{array}$$

□

Furthermore, from the proof of Theorem 6.10 for $\widetilde{P}_r = P_r = I$ we obtain the following perturbation bound for the solution of the regular GCALE (4.9).

Corollary 6.12. *Consider the GCALE* (4.9)*, where the pencil $\lambda E - A$ is c-stable and the matrix E is nonsingular. Assume that perturbations in* (6.11) *satisfy $\|\Delta E\|_2 \le \varepsilon\|E\|_2$, $\|\Delta A\|_2 \le \varepsilon\|A\|_2$ and $\|\Delta G\|_2 \le \varepsilon\|G\|_2$. If $\varepsilon\,(2+\varepsilon)\,\kappa_{c,2}(E,A) < 1$, then the perturbed GCALE* (6.11) *has a solution $\widetilde{X}$ and the relative error bound*

$$\frac{\|\widetilde{X} - X\|_2}{\|X\|_2} \le \frac{\varepsilon(3+\varepsilon)\kappa_{c,2}(E,A)}{1-\varepsilon(2+\varepsilon)\kappa_{c,2}(E,A)} \tag{6.40}$$

holds.

Note that bound (6.40) can also be obtained by applying the linear operator perturbation theory [90] to the regular GCALE (4.9) in the operator form $\mathcal{L}_c(X) = -G$.

If $\hat{X}$ is an approximate solution of the GCALE (4.9) and if $\mathcal{R}_c$ is a residual given by (6.16), then from Corollary 6.12 with $\Delta E = 0$, $\Delta A = 0$ and $\Delta G = \mathcal{R}_c$ we obtain the following forward error bound

$$\frac{\|\hat{X} - X\|_2}{\|X\|_2} \le \frac{\kappa_{c,2}(E,A)\,\|\mathcal{R}_c\|_2}{2\|E\|_2\|A\|_2\|X\|_2} =: Est_{c,2}. \tag{6.41}$$

Bounds (6.40) and (6.41) show that $\kappa_{c,2}(E,A)$ just as $\kappa_{c,F}(E,A)$ may also be used to measure the sensitivity of the solution of the regular GCALE (4.9). From the relationship

$$\frac{1}{\sqrt{n}}\|\mathcal{L}_c^{-1}\|_2 \le \|\mathcal{L}_c^{-1}\|_F \le \sqrt{n}\|\mathcal{L}_c^{-1}\|_2$$

we obtain that the Frobenius norm based condition number $\kappa_{c,F}(E,A)$ does not differ more than a factor $\sqrt{n}$ from the spectral condition number $\kappa_{c,2}(E,A)$. Thus, $\kappa_{c,2}(E,A)$ may be used as an estimator of $\kappa_{c,F}(E,A)$. Note that to compute Sep_c^{-1}-estimators we need to solve approximately five generalized Lyapunov equations of the form $E^*XA + A^*XE = -G$ and $EXA^* + AXE^* = -G$, see [1, 75], whereas the computation of $\|H_c\|_2$ requires solving only one additional generalized Lyapunov equation $E^*XA + A^*XE = -I$.

6.4 Conditioning of the projected GDALE

In this subsection we present the perturbation theory for the projected GDALE

$$\begin{aligned} &A^*XA - E^*XE = -P_r^*GP_r + (I-P_r)^*G(I-P_r), \\ &P_l^*X = XP_l. \end{aligned} \tag{6.42}$$

All results are based on the approach developed in [112, 113].

Assume that the pencil $\lambda E - A$ is d-stable. We define a *spectral condition number* for the projected GDALE (6.42) as

$$\kappa_{d,2}(E,A) := (\|E\|_2^2 + \|A\|_2^2)\|H_d\|_2, \tag{6.43}$$

where H_d is the positive definite matrix as in (3.32). Using (4.80), we obtain from the Parseval identity [135] that

$$H_d = \frac{1}{2\pi}\int_0^{2\pi} (e^{i\varphi}E - A)^{-*}\Big(P_r^*P_r + (I-P_r)^*(I-P_r)\Big)(e^{i\varphi}E - A)^{-1}d\varphi. \tag{6.44}$$

Then by Theorem 4.39 we have that H_d is the unique Hermitian, positive definite solution of the projected GDALE

$$\begin{aligned} &A^*H_dA - E^*H_dE = -P_r^*P_r + (I-P_r)^*(I-P_r),\\ &P_l^*H_d = H_dP_l. \end{aligned} \tag{6.45}$$

As we have seen in Section 3.2.2, the parameter $\kappa_{d,2}(E,A)$ characterizes the asymptotic stability of the singular difference equation (3.29). Here we will show that $\kappa_{d,2}(E,A)$ can also be used to estimate the distance from the finite eigenvalues of a d-stable pencil $\lambda E - A$ to the unite circle and to measure the sensitivity of the solution of the projected GDALE (6.42).

Set

$$\Theta := \max_{0\le\varphi\le 2\pi} \|(e^{i\varphi}E - A)^{-1}\|_2.$$

Clearly, if the pencil $\lambda E - A$ is d-stable, then $\Theta < \infty$. However, the boundedness of Θ does not imply that $\lambda E - A$ is d-stable. The following lemma gives lower and upper bounds for Θ by means of $\kappa_{d,2}(E,A)$.

Theorem 6.13. *Assume that the pencil $\lambda E - A$ is d-stable. Then all finite eigenvalues of $\lambda E - A$ lie in the closed disk*

$$\left\{ z\in\mathbb{C} \quad : \quad |z| \le \frac{\kappa_{d,2}(E,A) - 1}{\kappa_{d,2}(E,A)} \right\}.$$

Moreover,

$$\frac{\sqrt{\kappa_{d,2}(E,A)}}{\|P_r\|_2\sqrt{2(\|E\|_2^2 + \|A\|_2^2)}} \le \Theta \le \frac{10\pi\|E\|_2\kappa_{d,2}(E,A)}{\|E\|_2^2 + \|A\|_2^2}. \tag{6.46}$$

Proof. Let λ_0 be a finite eigenvalue of the pencil $\lambda E - A$ and let $v \in \operatorname{Im} P_r$ be an eigenvector corresponding to λ_0. Then from the projected GDALE (6.45) we obtain that

$$-\|v\|^2 = -\|P_rv\|^2 = v^*(A^*H_dA - E^*H_dE)v = (|\lambda_0| - 1)v^*E^*H_dEv,$$

and, hence,

$$|\lambda_0| = 1 - \frac{\|v\|^2}{v^*E^*H_dEv} \le 1 - \frac{1}{\kappa_{d,2}(E,A)} = \frac{\kappa_{d,2}(E,A) - 1}{\kappa_{d,2}(E,A)}.$$

The first estimate in (6.46) immediately follows from the inequalities

$$\begin{aligned} \kappa_{d,2}(E,A) &\le \frac{1}{2\pi}(\|E\|_2^2 + \|A\|_2^2)(\|P_r\|_2^2 + \|I-P_r\|_2^2)\int_0^{2\pi} \|(e^{i\varphi}E - A)^{-1}\|_2^2 d\varphi \\ &\le 2(\|E\|_2^2 + \|A\|_2^2)\|P_r\|_2^2\Theta^2. \end{aligned}$$

On the other hand, we have

$$\begin{aligned}\frac{\kappa_{d,2}(E,A)}{\|E\|_2^2+\|A\|_2^2} &= \|H_d\|_2 = \max_{\|v\|=1}(H_d v, v)\\ &\geq \frac{1}{2\pi}\int_0^{2\pi} v^*(e^{i\varphi}E-A)^{-*}\Big(P_r^*P_r+(I-P_r)^*(I-P_r)\Big)(e^{i\varphi}E-A)^{-1}v\,d\varphi\\ &= \frac{1}{2\pi}\int_0^{2\pi}\big(\|P_r(e^{i\varphi}E-A)^{-1}v\|^2+\|(I-P_r)(e^{i\varphi}E-A)^{-1}v\|^2\big)\,d\varphi\\ &\geq \frac{1}{4\pi}\int_0^{2\pi}\|(e^{i\varphi}E-A)^{-1}v\|^2 d\varphi.\end{aligned}$$

Let $0\leq\varphi_0\leq 2\pi$ be a point where the norm $\|(e^{i\varphi}E-A)^{-1}\|_2$ achieves its maximal value. It follows from the generalized resolvent equation (2.5) with $\lambda=e^{i\varphi}$ and $\mu=e^{i\varphi_0}$ that the estimate

$$\|(e^{i\varphi}E-A)^{-1}\|_2 \leq \frac{\|(e^{i\varphi_0}E-A)^{-1}\|_2}{1-|\varphi-\varphi_0|\|E\|_2\|(e^{i\varphi_0}E-A)^{-1}\|_2}$$

holds for all φ_0 such that $|\varphi-\varphi_0|\|E\|_2\|(e^{i\varphi_0}E-A)^{-1}\|_2<1$. Let v be the right singular vector of unit length corresponding to the largest singular value of the matrix $(e^{i\varphi_0}E-A)^{-1}$. Then $\|(e^{i\varphi_0}E-A)^{-1}v\|=\|(e^{i\varphi_0}E-A)^{-1}\|_2$ and

$$\begin{aligned}\|(e^{i\varphi}E-A)^{-1}v\| &\geq \|(e^{i\varphi_0}E-A)^{-1}v\|(1-|\varphi-\varphi_0|\|E\|_2\|(e^{i\varphi}E-A)^{-1}\|_2)\\ &\geq \|(e^{i\varphi_0}E-A)^{-1}\|_2\frac{1-2|\varphi-\varphi_0|\|E\|_2\|(e^{i\varphi_0}E-A)^{-1}\|_2}{1-|\varphi-\varphi_0|\|E\|_2\|(e^{i\varphi_0}E-A)^{-1}\|_2}.\end{aligned}$$

Hence, for $\psi=\|E\|_2\|(e^{i\varphi_0}E-A)^{-1}\|_2$, we have

$$\frac{4\pi\|E\|_2^2\kappa_{d,2}(E,A)}{\|E\|_2^2+\|A\|_2^2}\geq\int_{\varphi_0-\frac{1}{2\psi}}^{\varphi_0+\frac{1}{2\psi}}\psi^2\left(\frac{1-2|\varphi-\varphi_0|\psi}{1-|\varphi-\varphi_0|\psi}\right)^2 d\varphi = 2\psi(3-4\ln 2)\geq\frac{2\psi}{5}.$$

Thus, the upper bound in (6.46) holds. □

The second estimate in (6.46) shows that the eigenvalues of the d-stable pencil $\lambda E-A$ are separated from the unit circle by a distance not less than $(\|E\|_2+\|A\|_2)/(10\pi\kappa_{d,2}(E,A))$. Thus, the parameter $\kappa_{d,2}(E,A)$ characterizes the absence of eigenvalues of $\lambda E-A$ on the unite circle as well as in a neighbourhood of it.

Consider now a perturbed projected GDALE

$$\begin{aligned}&\widetilde{A}^*\widetilde{X}\widetilde{A}-\widetilde{E}^*\widetilde{X}\widetilde{E} = -\widetilde{P}_r^*\widetilde{G}\widetilde{P}_r+(I-\widetilde{P}_r)^*\widetilde{G}(I-\widetilde{P}_r),\\ &\widetilde{P}_l^*\widetilde{X}=\widetilde{X}\widetilde{P}_l,\end{aligned}\qquad(6.47)$$

where $\widetilde{P}_r$ and $\widetilde{P}_l$ are the spectral projections onto the right and left finite deflating subspaces of the perturbed pencil $\lambda\widetilde{E}-\widetilde{A}=\lambda(E+\Delta E)-(A+\Delta A)$. Note that small perturbations in E and A can make the infinite eigenvalues of the pencil $\lambda E-A$ to be finite.

In this case the perturbation analysis for the projected GDALE (6.42) becomes difficult. In the sequel we will consider only perturbations that do not change the dimension of the deflating subspaces of $\lambda E - A$ corresponding to the finite eigenvalues. The following lemma gives an error bound for the spectral projection P_r.

Lemma 6.14. *Let $\lambda E - A$ be a d-stable pencil and let $\lambda\widetilde{E} - \widetilde{A}$ be a perturbation of $\lambda E - A$ such that $\|\widetilde{E} - E\|_2 \le \varepsilon\|E\|_2$ and $\|\widetilde{A} - A\|_2 \le \varepsilon\|A\|_2$. Assume that the finite deflating subspaces of $\lambda E - A$ and $\lambda\widetilde{E} - \widetilde{A}$ have the same dimension. If*

$$20\pi\varepsilon\kappa_{d,2}(E,A)\Big(30\pi\kappa_{d,2}(E,A) + 1\Big) < 1, \tag{6.48}$$

then the pencil $\lambda\widetilde{E} - \widetilde{A}$ is d-stable and we have the following estimate

$$\|\widetilde{P}_r - P_r\|_2 \le 20\pi\varepsilon\kappa_{d,2}(E,A)\Big(30\pi\kappa_{d,2}(E,A) + 1\Big). \tag{6.49}$$

Proof. It follows from the generalized resolvent equation (2.5) that

$$(e^{i\varphi}\widetilde{E}-\widetilde{A})^{-1} = (e^{i\varphi}E-A)^{-1}-(e^{i\varphi}E-A)^{-1}\Big(e^{i\varphi}(\widetilde{E}-E)-(\widetilde{A}-A)\Big)(e^{i\varphi}\widetilde{E}-\widetilde{A})^{-1}. \tag{6.50}$$

Using (6.46) we get

$$\begin{aligned} \|(e^{i\varphi}\widetilde{E} - \widetilde{A})^{-1}\|_2 &\le \frac{\|(e^{i\varphi}E - A)^{-1}\|_2}{1 - \varepsilon(\|E\|_2 + \|A\|_2)\|(e^{i\varphi}E - A)^{-1}\|_2} \\ &\le \frac{10\pi\|E\|_2\kappa_{d,2}(E,A)}{(1 - 20\pi\varepsilon\kappa_{d,2}(E,A))(\|E\|_2^2 + \|A\|_2^2)}. \end{aligned} \tag{6.51}$$

Thus, if $\lambda E - A$ is d-stable and $20\pi\varepsilon\kappa_{d,2}(E,A) < 1$, then the pencil $\lambda\widetilde{E} - \widetilde{A}$ has no eigenvalues on the unit circle.

By Lemma 2.6 the spectral projection P_r onto the right finite deflating subspace of the d-stable pencil has the form

$$P_r = \frac{1}{2\pi}\int_0^{2\pi} e^{i\varphi}(e^{i\varphi}E - A)^{-1}E\,d\varphi,$$

and the spectral projection onto the right deflating subspace of the pencil $\lambda\widetilde{E} - \widetilde{A}$ corresponding to the eigenvalues inside the unit circle is given by

$$\widetilde{P} = \frac{1}{2\pi}\int_0^{2\pi} e^{i\varphi}(e^{i\varphi}\widetilde{E} - \widetilde{A})^{-1}\widetilde{E}\,d\varphi.$$

From (6.46) and (6.50) we have

$$\begin{aligned} \|(e^{i\varphi}\widetilde{E} - \widetilde{A})^{-1} - (e^{i\varphi}E - A)^{-1}\|_2 &\le \frac{\varepsilon\|(e^{i\varphi}E - A)^{-1}\|_2^2(\|E\|_2 + \|A\|_2)}{1 - \varepsilon(\|E\|_2 + \|A\|_2)\|(e^{i\varphi}E - A)^{-1}\|_2} \\ &\le \frac{2\varepsilon(10\pi\kappa_{d,2}(E,A))^2}{(1 - 20\pi\varepsilon\kappa_{d,2}(E,A))\|E\|_2}. \end{aligned} \tag{6.52}$$

Therefore,

$$\begin{aligned}\|\widetilde{P}-P_r\|_2 &= \left\|\frac{1}{2\pi}\int_0^{2\pi} e^{i\varphi}\Big((e^{i\varphi}\widetilde{E}-\widetilde{A})^{-1}\widetilde{E}-(e^{i\varphi}E-A)^{-1}E\Big)d\varphi\right\|_2\\ &\le \frac{1}{2\pi}\int_0^{2\pi}\Big(\|(e^{i\varphi}\widetilde{E}-\widetilde{A})^{-1}-(e^{i\varphi}E-A)^{-1}\|_2\|\widetilde{E}\|_2+\varepsilon\|(e^{i\varphi}E-A)^{-1}\|_2\|E\|_2\Big)d\varphi\\ &\le \frac{10\pi\varepsilon\kappa_{d,2}(E,A)\,(30\pi\kappa_{d,2}(E,A)+1)}{1-20\pi\varepsilon\kappa_{d,2}(E,A)}.\end{aligned}$$

It follows from estimate (6.48) that $\|\widetilde{P}-P_r\|_2<1$, and, hence, $\operatorname{Im}P_r$ and $\operatorname{Im}\widetilde{P}$ have the same dimension. In this case $\widetilde{P}=\widetilde{P}_r$ is the spectral projection onto the right finite deflating subspace of $\lambda\widetilde{E}-\widetilde{A}$. Thus, $\lambda\widetilde{E}-\widetilde{A}$ is d-stable and bound (6.49) holds. □

The following theorem gives a relative error bound for the solution of the projected GDALE (6.42).

Theorem 6.15. *Let $\lambda E-A$ be a d-stable pencil and let X be a solution of the projected GDALE (6.42). Let perturbations in (6.47) satisfy $\|\widetilde{E}-E\|_2\le\varepsilon\|E\|_2$, $\|\widetilde{A}-A\|_2\le\varepsilon\|A\|_2$ and $\|\widetilde{G}-G\|_2\le\varepsilon\|G\|_2$. Assume that the right and left finite deflating subspaces of $\lambda E-A$ and $\lambda\widetilde{E}-\widetilde{A}$ have the same dimension. If (6.48) is fulfilled, then then the perturbed projected GDALE (6.47) has a unique solution $\widetilde{X}$ and an error bound*

$$\|\widetilde{X}-X\|_2\le\frac{\varepsilon\kappa_{d,2}(E,A)(80\pi\kappa_{d,2}(E,A)+1)(60\pi\kappa_{d,2}(E,A)(1+2\|P_r\|_2)+\|P_r\|_2^2)\|G\|_2}{(1-20\pi\varepsilon\kappa_{d,2}(E,A))^2(\|E\|_2^2+\|A\|_2^2)} \tag{6.53}$$

holds.

Proof. It follows from Lemma 6.14 that the perturbed pencil $\lambda\widetilde{E}-\widetilde{A}$ is d-stable. Then by Theorem 4.39 the projected GDALE (6.47) has a unique solution $\widetilde{X}$ given by

$$\widetilde{X}=\frac{1}{2\pi}\int_0^{2\pi}(e^{i\varphi}\widetilde{E}-\widetilde{A})^{-*}\Big(\widetilde{P}_r^*\widetilde{G}\widetilde{P}_r+(I-\widetilde{P}_r)^*\widetilde{G}(I-\widetilde{P}_r)\Big)(e^{i\varphi}\widetilde{E}-\widetilde{A})^{-1}d\varphi.$$

The solution X of the projected GDALE (6.42) has the form (4.65) with $\xi=1$, and, hence,

$$\widetilde{X}-X=\frac{1}{2\pi}\int_0^{2\pi}\Big((e^{i\varphi}\widetilde{E}-\widetilde{A})^{-*}\widetilde{D}(e^{i\varphi}\widetilde{E}-\widetilde{A})^{-1}-(e^{i\varphi}E-A)^{-*}D(e^{i\varphi}E-A)^{-1}\Big)d\varphi,$$

where $\widetilde{D}=\widetilde{P}_r^*\widetilde{G}\widetilde{P}_r+(I-\widetilde{P}_r)^*\widetilde{G}(I-\widetilde{P}_r)$ and $D=P_r^*GP_r+(I-P_r)^*G(I-P_r)$. Taking into account estimates (6.51) and (6.52) we obtain

$$
\begin{aligned}
&\|(e^{i\varphi}\widetilde{E} - \widetilde{A})^{-*}\widetilde{D}(e^{i\varphi}\widetilde{E} - \widetilde{A})^{-1} - (e^{i\varphi}E - A)^{-*}D(e^{i\varphi}E - A)^{-1}\|_2 \\
&\quad \le \|(e^{i\varphi}\widetilde{E} - \widetilde{A})^{-1} - (e^{i\varphi}E - A)^{-1}\|_2\|D\|_2 \Big(\|(e^{i\varphi}\widetilde{E} - \widetilde{A})^{-1}\|_2 + \|(e^{i\varphi}E - A)^{-1}\|_2\Big) \\
&\qquad + \|(e^{i\varphi}\widetilde{E} - \widetilde{A})^{-1}\|_2^2\|\widetilde{D} - D\|_2 \\
&\quad \le \frac{\|\widetilde{D} - D\|_2 + 2\varepsilon(\|E\|_2 + \|A\|_2)\|(e^{i\varphi}E - A)^{-1}\|_2\|D\|_2}{(1 - \varepsilon(\|E\|_2 + \|A\|_2)\|(e^{i\varphi}E - A)^{-1}\|_2)^2}\|(e^{i\varphi}E - A)^{-1}\|_2^2 \\
&\quad \le \frac{\|\widetilde{D} - D\|_2 + 40\pi\varepsilon\|D\|_2\kappa_{d,2}(E, A)}{(1 - 20\pi\varepsilon\kappa_{d,2}(E, A))^2}\|(e^{i\varphi}E - A)^{-1}\|_2^2.
\end{aligned}
$$

Using (6.48) and (6.49) we have

$$
\begin{aligned}
\|\widetilde{D} - D\|_2 &\le 2(\|\widetilde{P}_r - P_r\|_2\|\widetilde{G}\|_2(\|\widetilde{P}_r\|_2 + \|P_r\|_2) + \|\widetilde{G} - G\|_2\|P_r\|_2^2 \\
&\le 60\pi\varepsilon\kappa_{d,2}(E, A)(30\pi\varepsilon\kappa_{d,2}(E, A) + 1)\|G\|(1 + 2\|P_r\|_2) + \varepsilon\|G\|_2\|P_r\|_2^2.
\end{aligned}
$$

Thus, bound (6.53) holds. □

Bound (6.53) shows that if $\|P_r\|_2$ and $\kappa_{d,2}(E, A)$ are not too large, then the solution of the perturbed projected GDALE (6.47) is a small perturbation of the solution of the projected GDALE (6.42). A large $\|P_r\|_2$ implies that the right finite deflating subspace of the pencil $\lambda E - A$ is ill-conditioned, whereas a large condition number $\kappa_{d,2}(E, A)$ implies that a finite eigenvalue of the d-stable pencil $\lambda E - A$ lies close to the unit circle.

Thus, $\kappa_{d,2}(E, A)$ can be used to estimate the sensitivity of the solution of the projected GDALE (6.42) to perturbations in the data. To compute $\kappa_{d,2}(E, A)$ we need to solve the projected GDALE (6.45). The solution H_d of this equation can be calculated via the generalized Schur-Bartels-Stewart method or the generalized Schur-Hammarling methods presented in Sections 5.1 and 5.2. Note that for the the projected GDALE (6.45) one can also use the Malyshev algorithm [112, 113].

From Theorem 6.15 we have the following perturbation bound for the spectral condition number $\kappa_{d,2}(E, A)$.

Corollary 6.16. *Under the assumptions of Theorem* 6.15 *we have the following relative perturbation bound*

$$
\frac{|\kappa_{d,2}(\widetilde{E}, \widetilde{A}) - \kappa_{d,2}(E, A)|}{\kappa_{d,2}(E, A)} \le \frac{60\pi\varepsilon\kappa_{d,2}(E, A)((30\pi\varepsilon\kappa_{d,2}(E, A) + 1)(1 + 2\|P_r\|_2) + 1) + 3\varepsilon}{(1 - 20\pi\varepsilon\kappa_{d,2}(E, A))^2}.
$$

Proof. From the proof of Theorem 6.15 with $\widetilde{G} = G = I$ we obtain that

$$
\|\widetilde{H}_d - H_d\|_2 \le \frac{40\pi\varepsilon\kappa_{d,2}(E, A)((30\pi\varepsilon\kappa_{d,2}(E, A) + 1)(1 + 2\|P_r\|_2) + 1)}{(1 - 20\pi\varepsilon\kappa_{d,2}(E, A))^2}\|H_d\|_2,
$$

where $\widetilde{H}_d$ is the solution of the perturbed GDALE (6.47) with $\widetilde{G} = I$. In this case

$$\begin{aligned} |\kappa_{d,2}(\widetilde{E},\widetilde{A}) - \kappa_{d,2}(E,A)| &= \left|(\|\widetilde{E}\|_2^2 + \|\widetilde{A}\|_2^2)\|\widetilde{H}_d\|_2 + (\|E\|_2^2 + \|A\|_2^2)\|H_d\|_2\right| \\ &\le \left|(\varepsilon+1)^2(\|E\|_2^2 + \|A\|_2^2)(\|\widetilde{H}_d - H_d\|_2 + \|H_d\|_2) - (\|E\|_2^2 + \|A\|_2^2)\|H_d\|_2\right| \\ &\le \left(\frac{60\pi\varepsilon\kappa_{d,2}(E,A)((30\pi\varepsilon\kappa_{d,2}(E,A)+1)(1+2\|P_r\|_2)+1)+3\varepsilon}{(1-20\pi\varepsilon\kappa_{d,2}(E,A))^2}\right)\kappa_{d,2}(E,A). \end{aligned}$$

□

The following corollary gives a perturbation bound for the regular GDALE (4.43). It can be obtained from the proof of Theorem 6.15 with $\widetilde{P}_r = P_r = I$ or by applying the linear operator perturbation theory [90] to (4.43) in the operator form $\mathcal{L}_d(X) = -G$.

Corollary 6.17. *Consider the GDALE* (4.43)*, where the pencil $\lambda E - A$ is d-stable and the matrix E is nonsingular. Assume that perturbations in* (6.18) *satisfy $\|\Delta E\|_2 \le \varepsilon\|E\|_2$, $\|\Delta A\|_2 \le \varepsilon\|A\|_2$ and $\|\Delta G\|_2 \le \varepsilon\|G\|_2$. If*

$$\varepsilon\,(2+\varepsilon)\,\kappa_{d,2}(E,A) < 1,$$

then the perturbed GDALE (6.18) *has a solution $\widetilde{X}$ and the relative error bound*

$$\frac{\|\widetilde{X} - X\|_2}{\|X\|_2} \le \frac{\varepsilon(3+\varepsilon)\kappa_{d,2}(E,A)}{1-\varepsilon(2+\varepsilon)\kappa_{d,2}(E,A)} \tag{6.54}$$

holds.

Let $\hat{X}$ be an approximate solution of the GDALE (4.43) and let $\mathcal{R}_d$ be a residual given in (6.22). Then from Corollary 6.17 with $\Delta E = 0$, $\Delta A = 0$ and $\Delta G = \mathcal{R}_d$ we have the following forward error bound

$$\frac{\|\hat{X} - X\|_2}{\|X\|_2} \le \frac{\kappa_{d,2}(E,A)\,\|\mathcal{R}_d\|_2}{(\|E\|_2^2 + \|A\|_2^2)\|X\|_2} =: Est_{d,2}.$$

This bound implies that if the regular GDALE (4.43) is well-conditioned and the residual is small, then the approximate solution $\hat{X}$ is a small perturbation of the exact solution. We see that the spectral condition number $\kappa_{d,2}(E,A)$ likewise the Frobenius norm condition number $\kappa_{d,F}(E,A)$ may be used to measure the sensitivity of the solution of the regular GDALE (4.43). Similar the continuous-time case, it can be shown that $\kappa_{d,2}(E,A)$ does not differ more than a factor $\sqrt{n}$ from $\kappa_{d,F}(E,A)$. However, to compute the one-norm estimators for $\kappa_{d,F}(E,A)$ we need to solve several generalized Lyapunov equations of the form $A^*XA - E^*XE = -G$ and $AXA^* - EXE^* = -G$, see [1, 75], whereas the computation of the spectral condition number $\kappa_{d,2}(E,A)$ requires solving only one additional generalized Lyapunov equation $A^*XA - E^*XE = -I$.

6.5 Numerical examples

In this section we present results of two sets of numerical experiments. The goal of the first set is to compare the spectral norm condition numbers and the Frobenius norm based condition numbers for regular generalized Lyapunov equations. In the second set we demonstrate the relevance of the spectral condition numbers proposed for projected generalized Lyapunov equations. Computations were carried out on IBM RS 6000 44P Modell 270 with relative machine precision $\texttt{EPS} \approx 2.22 \cdot 10^{-16}$.

Example 6.18. [125] The matrices E and A are defined as

$$\begin{aligned} E &= I_n + 2^{-t} U_n, \\ A &= (1 - 2^{-t}) I_n - \operatorname{diag}(1, 2, \ldots, n) - U_n^T \end{aligned}$$

in the continuous-time case and

$$\begin{aligned} E &= 2^{-t} I_n + \operatorname{diag}(1, 2, \ldots, n) + U_n^T, \\ A &= I_n + 2^{-t} U_n \end{aligned}$$

in the discrete-time case, where U_n is the $n \times n$ strictly lower triangular matrix with unit entries below the main diagonal. Note that E is nonsingular. The matrix G is defined so that a true solution X of the GCALE (4.9) or the GDALE (4.43) is a random matrix with entries uniformly distributed in $(0, 100)$.

We generated the generalized Lyapunov equations for a medium size $n = 100$ and different values of the parameter t. To compute the solutions of the GCALE (4.9) and the GDALE (4.43), the matrices H_c and H_d satisfying, respectively, (4.9) and (4.43) with $G = I$ as well as the Frobenius norm based estimators for $\operatorname{Sep}_c^{-1}(E, A)$ and $\operatorname{Sep}_d^{-1}(E, A)$, we use the SLICOT library subroutine `SG04AD` [16].

We compare the spectral condition numbers and the Frobenius norm based condition number in Figure 6.1 in the continuous-time case and in Figure 6.2 in the discrete-time case. One can see that $\kappa_{c,2}(E, A)$ is a factor 2-3 smaller than $\kappa_{c,F}(E, A)$ and $\kappa_{d,2}(E, A)$ is a factor 2-8 smaller than $\kappa_{d,F}(E, A)$. Both problems become ill-conditioned as the parameter t increases. Figures 6.3 and 6.4 show the relative errors in the spectral and Frobenius norms

$$\texttt{RERR2} = \frac{\|\hat{X} - X\|_2}{\|X\|_2}, \qquad \texttt{RERRF} = \frac{\|\hat{X} - X\|_F}{\|X\|_F},$$

where $\hat{X}$ is an approximate solution of (4.9) or (4.43) computed by the generalized Bartels-Stewart method. As expected from the perturbation theory, the accuracy of $\hat{X}$ may get worse as the condition numbers are large, while the relative residuals

$$\texttt{RRESC2} = \frac{\|E^* \hat{X} A + A^* \hat{X} E + G\|_2}{2\|E\|_2 \|A\|_2 \|X\|_2}, \qquad \texttt{RRESCF} = \frac{\|E^* \hat{X} A + A^* \hat{X} E + G\|_F}{2\|E\|_2 \|A\|_2 \|X\|_F}$$

in the continuous-time case (Figure 6.5) and

$$\texttt{RRESD2} = \frac{\|A^* \hat{X} A - E^* \hat{X} E + G\|_2}{(\|E\|_2^2 + \|A\|_2^2) \|X\|_2}, \qquad \texttt{RRESDF} = \frac{\|A^* \hat{X} A - E^* \hat{X} E + G\|_F}{(\|E\|_2^2 + \|A\|_2^2) \|X\|_F}$$

in the discrete-time case (Figure 6.6), remain small.

The continuous-time case

The discrete-time case

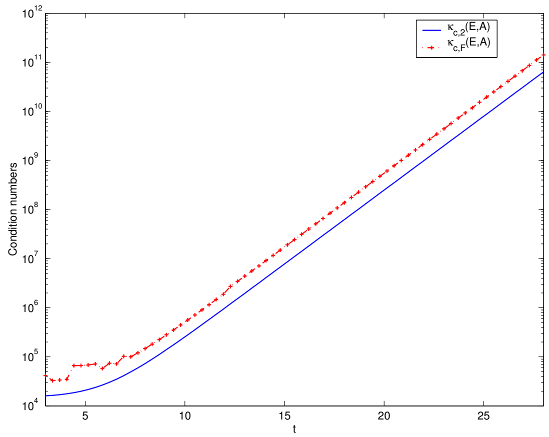


Figure 6.1: $\kappa_{c,2}(E, A)$ and $\kappa_{c,F}(E, A)$

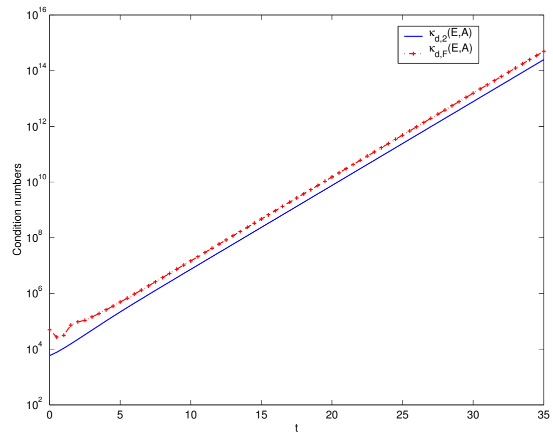


Figure 6.2: $\kappa_{d,2}(E, A)$ and $\kappa_{d,F}(E, A)$

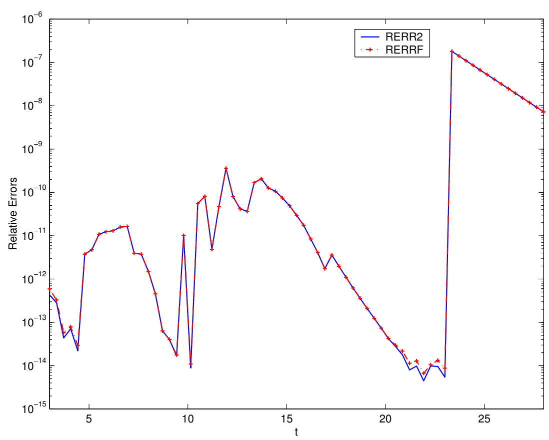


Figure 6.3: Relative errors in the solution

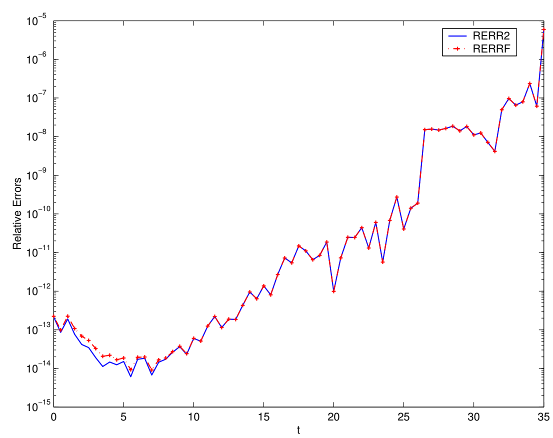


Figure 6.4: Relative errors in the solution

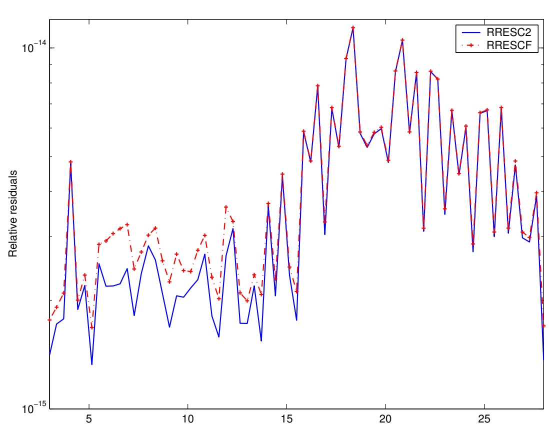


Figure 6.5: Relative residuals `RRESC2` and `RRESCF`

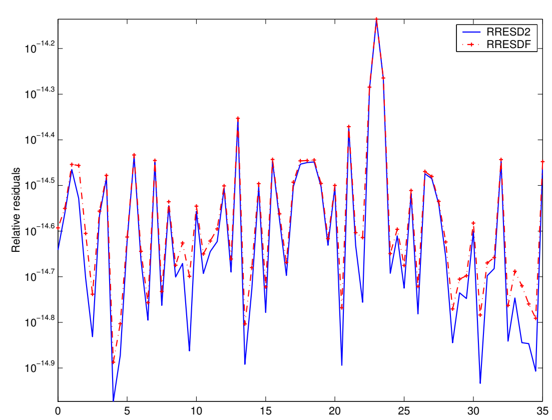


Figure 6.6: Relative residuals `RRESD2` and `RRESDF`

The continuous-time case

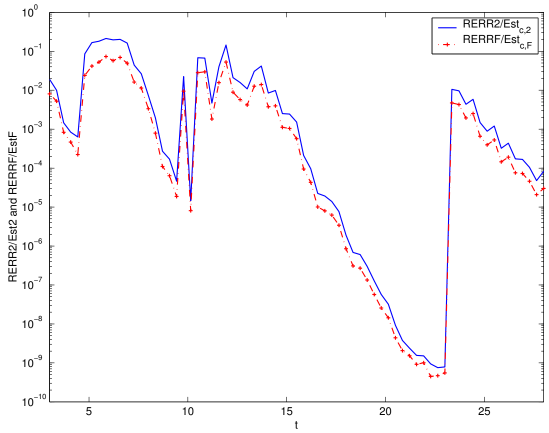


Figure 6.7: RERR/$Est_{c,2}$ and RERR/$Est_{c,F}$

The discrete-time case

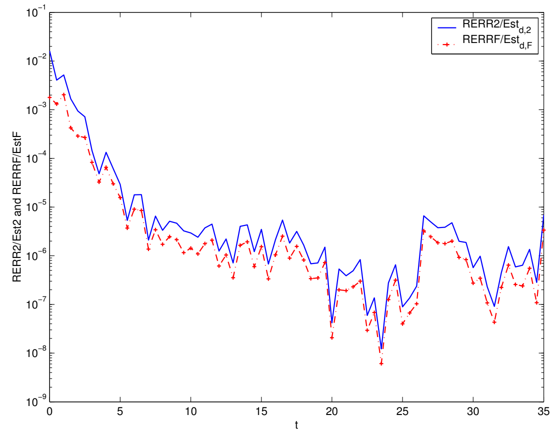


Figure 6.8: RERR/$Est_{d,2}$ and RERR/$Est_{d,F}$

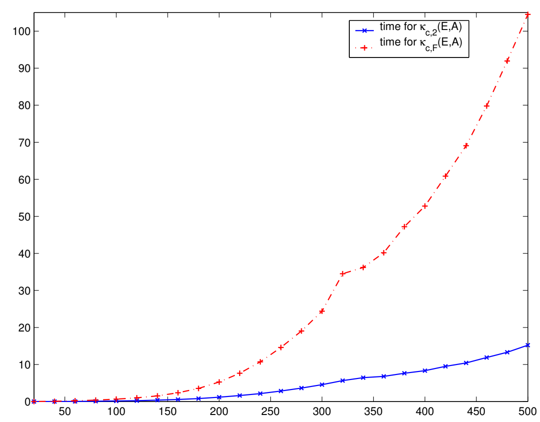


Figure 6.9: CPU-time in seconds required for computing $\kappa_{c,2}(E, A)$ and $\kappa_{c,F}(E, A)$

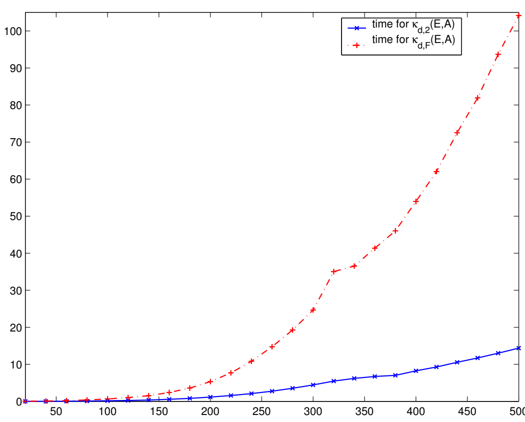


Figure 6.10: CPU-time in seconds required for computing $\kappa_{d,2}(E, A)$ and $\kappa_{d,F}(E, A)$

Figure 6.7 shows the ratios RERR2/$Est_{c,2}$ and RERRF/$Est_{c,F}$ between the relative errors and the computed residual based error estimates given by (6.17) and (6.41). An analogous result for the discrete-time case is presented in Figure 6.8. We see that the estimates in the spectral norm are sharper than the estimates in the Frobenius norm.

Finally, in Figures 6.9 and 6.10 we compare the CPU-time (in seconds) obtained via the LAPACK subroutine DSECND [1] that is required to compute the spectral norm and Frobenius norm condition numbers of the GCALE (4.9) and the GDALE (4.43) for the fixed parameter $t = 5$ and different sizes $n \in \{20, \ldots, 500\}$. We see that the computation of $\kappa_{c,2}(E, A)$ and $\kappa_{d,2}(E, A)$ is significantly faster especially for large problems than the estimators for $\kappa_{c,F}(E, A)$ and $\kappa_{d,F}(E, A)$. This is not surprising because to compute the spectral norm condition numbers we need to solve only one additional generalized Lyapunov equation, while computing the Frobenius norm based condition numbers requires solving approximately five generalized Lyapunov equations.

This numerical example shows that the spectral norm condition numbers and the Frobenius norm based condition numbers gives the similar information on the conditioning of regular generalized Lyapunov equations. However, from the point of view of computational costs the first are superior.

Example 6.19. Consider a family of projected GCALEs with

$$E = V \begin{bmatrix} I_3 & D(N_3 - I_3) \\ 0 & N_3 \end{bmatrix} U^T, \qquad A = V \begin{bmatrix} J & (I_3 - J)D \\ 0 & I_3 \end{bmatrix} U^T,$$
$$G = U \begin{bmatrix} G_{11} & -G_{11}D \\ -DG_{11} & DG_{11}D \end{bmatrix} U^T,$$

where N_3 is a nilpotent Jordan block of order 3,

$$\begin{aligned} J &= \operatorname{diag}(-10^{-k},\ -2,\ -3 \times 10^{k}), & k \geq 0, \\ D &= \operatorname{diag}(10^{-q},\ 1,\ 10^{q}), & q \geq 0, \\ G_{11} &= \operatorname{diag}(2,\ 4,\ 6). \end{aligned}$$

The transformation matrices V and U are elementary reflections chosen as

$$\begin{aligned} V &= I_6 - \frac{1}{3}ee^T, \qquad e = (1,\ 1,\ 1,\ 1,\ 1,\ 1)^T, \\ U &= I_6 - \frac{1}{3}ff^T, \qquad f = (1, -1,\ 1, -1,\ 1, -1)^T. \end{aligned} \tag{6.55}$$

The exact solution of the projected GCALE (4.36) is given by

$$X = V \begin{bmatrix} X_{11} & -X_{11}D \\ -DX_{11} & DX_{11}D \end{bmatrix} V^T \tag{6.56}$$

with $X_{11} = \operatorname{diag}(10^k,\ 1,\ 10^{-k})$. The problem becomes ill-conditioned when k and q increase.

To solve the projected GCALE (4.36) we use Algorithm 5.1.1. Computations were performed using MATLAB mex-functions based on the `GUPTRI` routine [41, 42] and the SLICOT routines `SG04OD` and `SG03AD` [16, 165].

In Figures 6.11 and 6.12 we show the values of Dif_u^{-1} and $\kappa_{c,2}(E_f, A_f)$ as functions of k and q. We see that the condition numbers of the generalized Sylvester equation (5.2) and the regular GCALE (5.13) are independent of q and increase with k.

In Figure 6.13 we show the values of $\|H_c\|_2$ and the condition number $\kappa_{c,2}(E, A)$ of the projected GCALE (4.36) for the same values of k and q. When k and q are increased, the condition number $\kappa_{c,2}(E, A)$ increases more quickly than $\|H_c\|_2$. Note that the projected GCALE (4.36) may be ill-conditioned even if both the intermediate problems are well-conditioned.

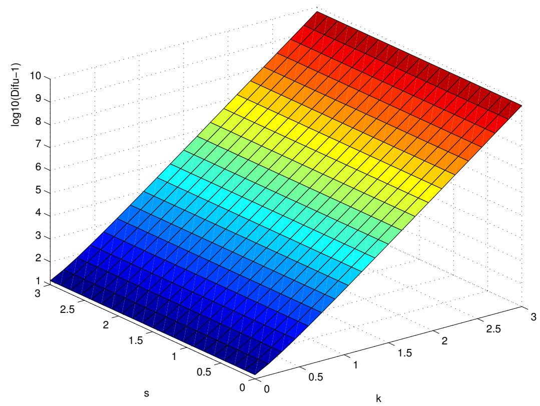


Figure 6.11: Conditioning of the generalized Sylvester equation in Example 6.19

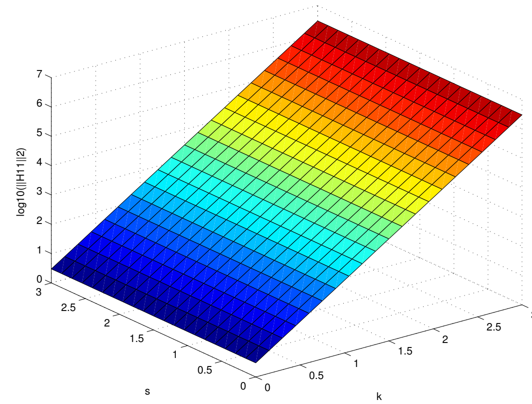


Figure 6.12: Conditioning of the regular GCALE

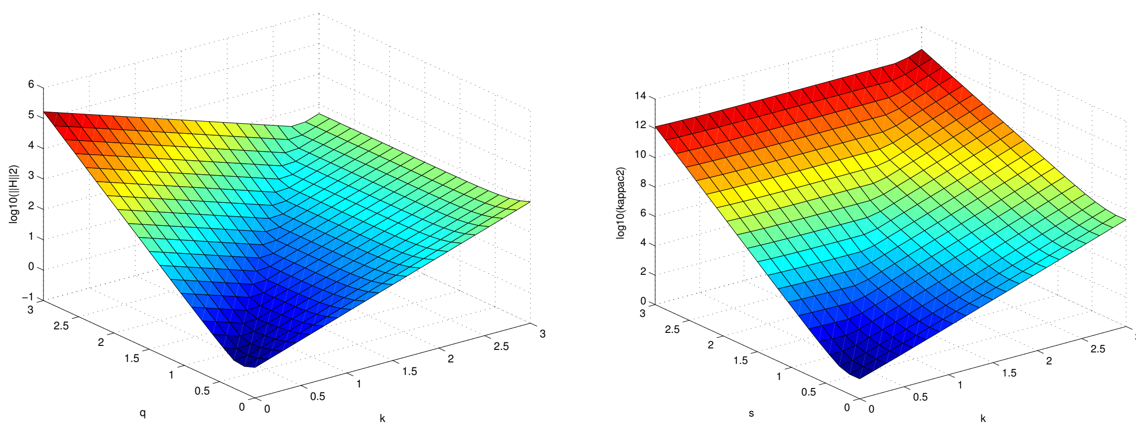


Figure 6.13: Conditioning of the projected GCALE

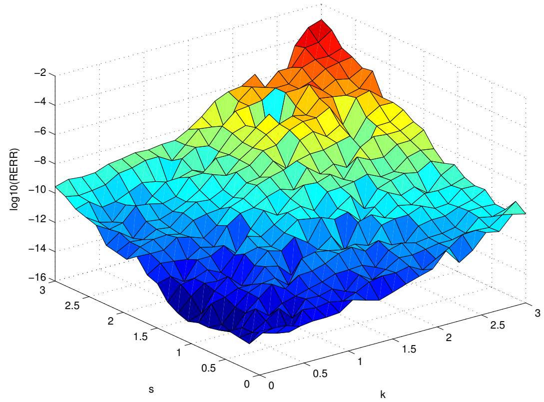


Figure 6.14: Relative error in the computed solution of the projected GCALE

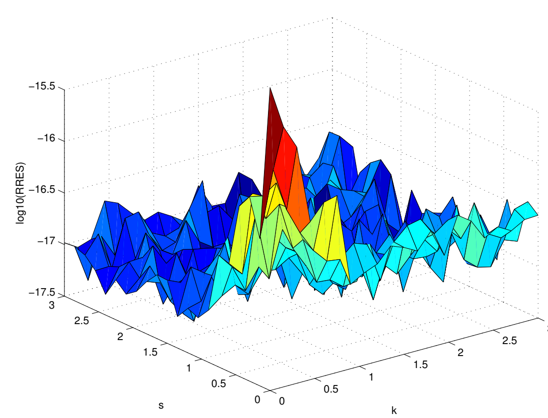


Figure 6.15: Relative residual `RRESC`

Finally, Figure 6.14 shows the relative error RERR $= \|\hat{X} - X\|_2/\|X\|_2$, where $\hat{X}$ is the computed solution, and Figure 6.15 shows the relative residual

$$\texttt{RRESC} = \frac{\|E^T\hat{X}A + A^T\hat{X}E + \hat{P}_r^T G\hat{P}_r\|_2}{2\|E\|_2\|A\|_2\|X\|_2},$$

where $\hat{P}_r$ is the computed projection onto the right deflating subspace of the pencil $\lambda E - A$ corresponding to the finite eigenvalues. We see that the relative residual is small even for the ill-conditioned problem. However, this does not imply that the relative error in the computed solution remains close to zero when the condition number $\kappa_{c,2}(E,A)$ is large. The relative error in $\hat{X}$ increases as $\kappa_{c,2}(E,A)$ grows.

Example 6.20. Consider the projected GDALE (6.42) with

$$E = V\begin{bmatrix} I_3 & D(N_3 - I_3) \\ 0 & N_3 \end{bmatrix} U^T, \qquad A = V\begin{bmatrix} J_1 & DJ_2 - J_1 D \\ 0 & J_2 \end{bmatrix} U^T,$$

$$G = U\begin{bmatrix} G_{11} & -G_{11}D \\ -DG_{11} & DG_{11}D \end{bmatrix} U^T,$$

where

$$\begin{aligned} J_1 &= \mathrm{diag}(1 - 10^{-k},\ 1/2,\ 0), & k \geq 0, \\ J_2 &= \mathrm{diag}(10^{2q/3},\ 1,\ 10^{-2q/3}), & q \geq 0, \\ D &= \mathrm{diag}(10^{-q},\ 1,\ 10^{q}), & \\ G_{11} &= \mathrm{diag}(2 - 10^{-k},\ 3/4,\ 10^{-k}), & \end{aligned}$$

and U, V are given by (6.55). The exact solution of the projected GDALE (6.42) has the form (6.56) with $X_{11} = \mathrm{diag}(10^k,\ 1,\ 10^{-k})$. An approximate solution $\hat{X}$ of (6.42) is computed using Algorithm 5.1.2.

In Figures 6.16 we show the values of Dif_u^{-1} as functions of k and q. One can see that the generalized Sylvester equation (5.2) is well-conditioned for all $k \in [0, 9]$ and $q \in [0,\ 2.7]$. Figures 6.17 and 6.18 show the spectral condition numbers $\kappa_{d,2}(E_f, A_f)$ and $\kappa_{d,2}(E_{inf}, A_{inf})$ of the regular GDALE (5.20) and the regular GDALE (5.21). The condition number of (5.20) does not depend on q and increases with k, while the condition number of (5.21) grows with q and is independent of k.

The spectral condition number $\kappa_{d,2}(E,A)$ of the projected GDALE (6.42) is depicted in Figure 6.19. We see that equation (6.42) becomes ill-conditioned when k and q increase.

The relative error RERR $= \|\hat{X} - X\|_2/\|X\|_2$ and the relative residual

$$\texttt{RRESD} = \frac{\|A^T\hat{X}A - E^T\hat{X}E + \hat{P}_r^T G\hat{P}_r - (I - \hat{P}_r)^T G(I - \hat{P}_r)\|_2}{(\|E\|_2^2 + \|A\|_2^2)\|X\|_2}$$

are shown in Figure 6.20 and Figure 6.21, respectively. Here $\hat{P}_r$ is the computed projection onto the right deflating subspace of $\lambda E - A$ corresponding to the finite eigenvalues. We see that even though the relative residual remains small, the accuracy in $\hat{X}$ may get worse for the large condition number $\kappa_{d,2}(E,A)$. Moreover, the computed solution may be inaccurate, if one of intermediate problems is ill-conditioned.

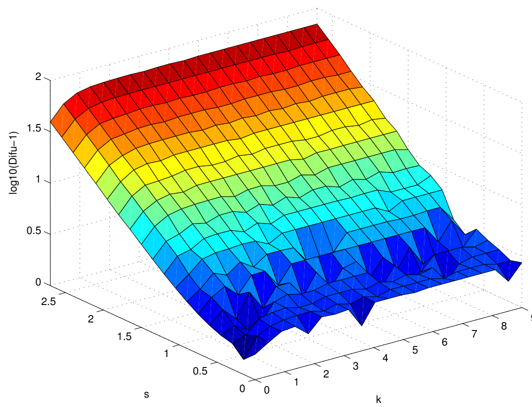


Figure 6.16: Conditioning of the generalized Sylvester equation in Example 6.20

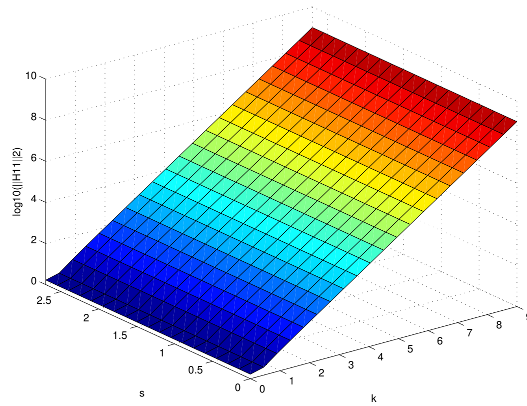


Figure 6.17: The spectral condition number $\kappa_{d,2}(E_f, A_f)$

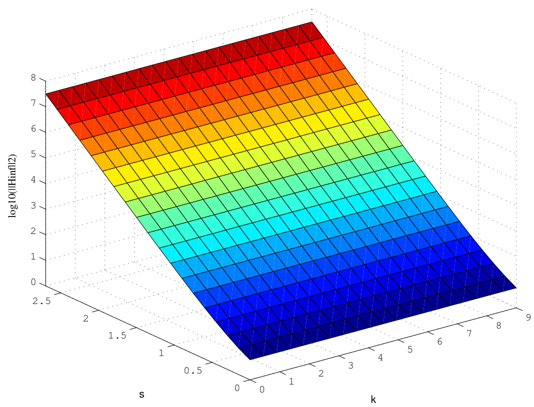


Figure 6.18: The spectral condition number $\kappa_{d,2}(E_{inf}, A_{inf})$

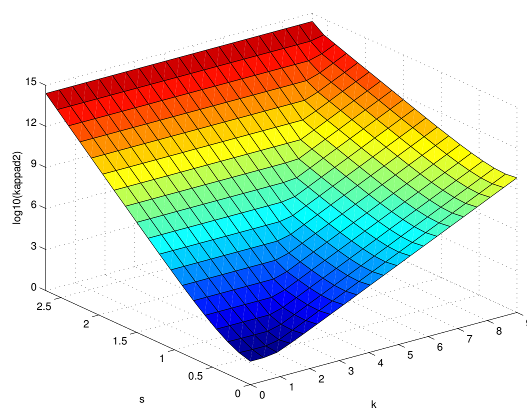


Figure 6.19: Conditioning of the projected GDALE

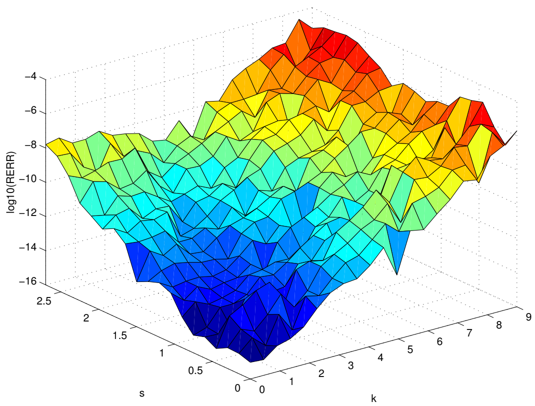


Figure 6.20: Relative error in the computed solution of the projected GDALE

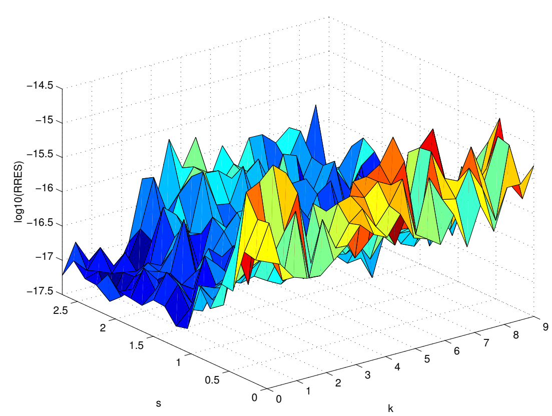


Figure 6.21: Relative residual `RRESD`

Chapter 7

Model reduction

An important field of applications for projected generalized Lyapunov equations is the model reduction of large scale descriptor systems that arise, for instance, from electrical circuit simulation and discretization of partial differential equations. The numerical methods for solving large systems or real time controller design cannot be applied to such systems due their computational complexity and storage requirements. This motivates the model order reduction that consists in the continuous-time case in an approximation of the descriptor system

$$\begin{aligned} E\,\dot{x}(t) &= A\,x(t) + B\,u(t), \quad x(0) = x^0, \\ y(t) &= C\,x(t) \end{aligned} \tag{7.1}$$

with E, $A \in \mathbb{R}^{n,n}$, $B \in \mathbb{R}^{n,m}$, $C \in \mathbb{R}^{p,n}$ by a reduced order system

$$\begin{aligned} \widetilde{E}\,\dot{\widetilde{x}}(t) &= \widetilde{A}\,\widetilde{x}(t) + \widetilde{B}\,u(t), \quad \widetilde{x}(0) = \widetilde{x}^0, \\ \widetilde{y}(t) &= \widetilde{C}\,\widetilde{x}(t), \end{aligned} \tag{7.2}$$

where $\widetilde{E}$, $\widetilde{A} \in \mathbb{R}^{\ell,\ell}$, $\widetilde{B} \in \mathbb{R}^{\ell,m}$, $\widetilde{C} \in \mathbb{R}^{p,\ell}$ and $\ell \ll n$. Note that systems (7.1) and (7.2) have the same input $u(t)$. One requires for the approximate system (7.2) to preserve properties of the original system (7.1) like regularity and stability. Since the descriptor system (7.1) consists of differential equations that describe the dynamic behavior of the system as well as algebraic equations characterizing a constraint manifold for the solution, it is natural to require for the reduced order system to have the same algebraic constraints as the original one. Clearly, it is also desirable that the approximation error is small. Moreover, the computation of the reduced order system should be numerically stable and efficient.

There exist various model reduction approaches for standard state space systems such as balanced truncation [102, 119, 129, 137, 156, 164], moment matching approximation [52, 68], singular perturbation approximation [94, 107] and optimal Hankel norm approximation [58]. Surveys on system approximation and model reduction can be found in [4, 48]. One of the most effective and well studied model reduction techniques is balanced truncation which is closely related to the controllability and observability Gramians. The balanced truncation method consists in transforming the state space system to a balanced form whose controllability and observability Gramians become diagonal and equal together with a truncation of states that are both difficult to reach and to observe. The diagonal elements of the transformed Gramians are known as the Hankel singular values of the dynamical system, and the truncated states correspond to the small Hankel singular values, see [119] for details. An important advantage of the balanced truncation approach is that if the original system is asymptotically stable then the reduced system is also asymptotically stable. Moreover, a priory bounds on the approximation error can be derived [46, 58].

In this chapter we generalize the Hankel singular values for descriptor systems and present an extension of known balanced truncation algorithms such as the square root method [102, 156] and the balancing free square root method [164] to descriptor systems.

7.1 Transfer function and realization

Consider the Laplace transform of a function $f(t)$, $t \in \mathbb{R}$, given by

$$\mathbf{f}(s) = \mathfrak{L}[f(t)] = \int_0^\infty e^{-st} f(t)\, dt,$$

where s is a complex variable called *frequency*. A discussion of the convergence region of $\mathbf{f}(s)$ in the complex plane and properties of the Laplace transform may be found in [43]. If we take in (7.1) the Laplace transform, then we obtain that

$$\mathbf{x}(s) = (sE - A)^{-1}B\mathbf{u}(s) + (sE - A)^{-1}Ex(0), \tag{7.3}$$

$$\mathbf{y}(s) = C(sE - A)^{-1}B\mathbf{u}(s) + C(sE - A)^{-1}Ex(0), \tag{7.4}$$

where $\mathbf{x}(s)$, $\mathbf{u}(s)$ and $\mathbf{y}(s)$ are the Laplace transforms of $x(t)$, $u(t)$ and $y(t)$, respectively. A rational matrix-valued function

$$\mathbf{G}(s) := C(sE - A)^{-1}B \tag{7.5}$$

is called the *transfer function* of the continuous-time descriptor system (7.1). We see in (7.4) that if $Ex(0) = 0$, then $\mathbf{G}(s)$ gives the relation between the Laplace transforms of the input $u(t)$ and the output $y(t)$. In other words, the transfer function $\mathbf{G}(s)$ describe the input-output behavior of system (7.1) in the frequency domain. The transfer function $\mathbf{G}(s)$ is said to be *c-stable* if the matrix pencil $\lambda E - A$ is c-stable, i.e., all finite eigenvalues of $\lambda E - A$ lie in the open left half-plane.

If for any rational matrix-valued function $\mathbf{G}(s)$ there exist matrices E, A, B and C such that $\mathbf{G}(s) = C(sE - A)^{-1}B$, then system (7.1) with these matrices is called a *realization* of $\mathbf{G}(s)$. We will also denote a realization of $\mathbf{G}(s)$ by $\mathbf{G} = [\,E,\ A,\ B,\ C\,]$ or by

$$\mathbf{G} = \left[\begin{array}{c|c} sE - A & B \\ \hline C & 0 \end{array}\right].$$

Note that the realization of $\mathbf{G}(s)$ is, in general, not unique [36].

Definition 7.1. Two realizations $[\,E,\,A,\,B,\,C\,]$ and $[\,\check{E},\,\check{A},\,\check{B},\,\check{C}\,]$ are *restricted system equivalent* if there exist nonsingular matrices $\check{W}$ and $\check{T}$ such that

$$E = \check{W}\check{E}\check{T}, \quad A = \check{W}\check{A}\check{T}, \quad B = \check{W}\check{B}, \quad C = \check{C}\check{T}.$$

A pair $(\check{W}, \check{T})$ is called *system equivalence transformation*.

The notion of the restricted system equivalence is in line with [134]. A characteristic quantity of system (7.1) is *input-output invariant* if it is preserved under a system equivalence transformation. The transfer function $\mathbf{G}(s)$ is input-output invariant, since

$$\mathbf{G}(s) = C(sE - A)^{-1}B = \check{C}\check{T}\check{T}^{-1}(s\check{E} - \check{A})^{-1}\check{W}^{-1}\check{W}\check{B} = \check{C}(s\check{E} - \check{A})^{-1}\check{B}.$$

Definition 7.2. A transfer function $\mathbf{G}(s)$ is called *proper* if $\lim\limits_{s\to\infty} \mathbf{G}(s) < \infty$. Otherwise, $\mathbf{G}(s)$ is *improper*. If $\lim\limits_{s\to\infty} \mathbf{G}(s) = 0$, then $\mathbf{G}(s)$ is said to be *strictly proper*.

Let the pencil $\lambda E - A$ be in Weierstrass canonical form (2.2) and let the matrices B and C be as in (3.3). Using the Laurent expansion (2.6) for the generalized resolvent $(\lambda E - A)^{-1}$, the transfer function $\mathbf{G}(s)$ can be written as

$$\mathbf{G}(s) = C_1(sI - J)^{-1}B_1 + C_2(sN - I)^{-1}B_2 = \mathbf{G}_{sp}(s) + \mathbf{P}(s),$$

where

$$\mathbf{G}_{sp}(s) = C_1(sI - J)^{-1}B_1 = \sum_{k=1}^{\infty} CF_{k-1}Bs^{-k}$$

is the *strictly proper part* of $\mathbf{G}(s)$ and

$$\mathbf{P}(s) = C_2(sN - I)^{-1}B_2 = \sum_{k=0}^{\nu-1} CF_{-k-1}Bs^{k}$$

is the *polynomial part* of $\mathbf{G}(s)$. The *proper part* of $\mathbf{G}(s)$ is given by

$$\mathbf{G}_p(s) = C_1(sI - J)^{-1}B_1 - C_2B_2 = \sum_{k=0}^{\infty} CF_{k-1}Bs^{-k}.$$

The matrices $M_k = CF_{k-1}B$ are called the *Markov parameters* of system (7.1). Clearly, they are input-output invariants. The transfer function $\mathbf{G}(s) = \mathbf{G}_p(s)$ is proper if and only if $M_k = CF_{k-1}B = 0$ for $k < 0$. For example, if the pencil $\lambda E - A$ is of index at most one, then $\mathbf{G}(s)$ is proper. The transfer function $\mathbf{G}(s) = \mathbf{G}_{sp}(s)$ is strictly proper if and only if $M_k = CF_{k-1}B = 0$ for $k \le 0$.

Other important results from the theory of rational functions and realization theory may be found in [36, 79, 166].

7.2 Hankel singular values

Assume that the pencil $\lambda E - A$ in the continuous-time descriptor system (7.1) is c-stable. Consider the controllability and observability Gramians of (7.1) introduced in Section 4.4.2. Similar to the state space systems [176], these Gramians can be used to define Hankel singular values for system (7.1) that will play a significant role in the model reduction via balanced truncation.

Note that under a system equivalence transformation $(\check{W}, \check{T})$ the proper and improper controllability Gramians $\mathcal{G}_{cpc}$ and $\mathcal{G}_{cic}$ of (7.1) are transformed to $\check{\mathcal{G}}_{cpc} = \check{T}^{-1}\mathcal{G}_{cpc}\check{T}^{-T}$ and $\check{\mathcal{G}}_{cic} = \check{T}^{-1}\mathcal{G}_{cic}\check{T}^{-T}$, respectively, whereas the proper and improper observability Gramians $\mathcal{G}_{cpo}$ and $\mathcal{G}_{cio}$ are transformed to $\check{\mathcal{G}}_{cpo} = \check{W}^{-T}\mathcal{G}_{cpo}\check{W}^{-1}$ and $\check{\mathcal{G}}_{cio} = \check{W}^{-T}\mathcal{G}_{cio}\check{W}^{-1}$, respectively. Thus, the Gramians are not input-output invariants. However, we know that for standard state space systems the spectrum of the product of the controllability and observability Gramians does not change under the system equivalence transformation [176]. For the descriptor system (7.1), an analogous result holds for the matrices

$$\begin{aligned} \Phi_{c,1} &:= \mathcal{G}_{cpc}E^T\mathcal{G}_{cpo}E, \\ \Phi_{c,2} &:= E\mathcal{G}_{cpc}E^T\mathcal{G}_{cpo}, \\ \Phi_{c,3} &:= \mathcal{G}_{cpo}E\mathcal{G}_{cpc}E^T, \\ \Phi_{c,4} &:= E^T\mathcal{G}_{cpo}E\mathcal{G}_{cpc}. \end{aligned} \tag{7.6}$$

Indeed, under a system equivalence transformation $(\check{W}, \check{T})$ these matrices are transformed to

$$\begin{aligned} \check{\mathcal{G}}_{cpc}\check{E}^T\check{\mathcal{G}}_{cpo}\check{E} &= \check{T}\mathcal{G}_{cpc}E^T\mathcal{G}_{cpo}E\check{T}^{-1}, \\ \check{E}\check{\mathcal{G}}_{cpc}\check{E}^T\check{\mathcal{G}}_{cpo} &= \check{W}^{-1}E\mathcal{G}_{cpc}E^T\mathcal{G}_{cpo}\check{W}, \\ \check{\mathcal{G}}_{cpo}\check{E}\check{\mathcal{G}}_{cpc}\check{E}^T &= \check{W}^T\mathcal{G}_{cpo}E\mathcal{G}_{cpc}E^T\check{W}^{-T}, \\ \check{E}^T\check{\mathcal{G}}_{cpo}\check{E}\check{\mathcal{G}}_{cpc} &= \check{T}^{-T}E^T\mathcal{G}_{cpo}E\mathcal{G}_{cpc}\check{T}^T, \end{aligned}$$

and, hence, the eigenvalues of $\Phi_{c,q}$ are input-output invariants. Moreover, we can prove that the matrices $\Phi_{c,q}$, $q = 1, \ldots, 4$, have the same spectrum.

Lemma 7.3. *Let $\lambda E - A$ be c-stable. Then the matrices $\Phi_{c,q}$, $q = 1, \ldots, 4$, given in* (7.6) *are diagonalizable and have the same eigenvalues that are real and non-negative.*

Proof. It follows from (4.85) and (4.86) that the matrices $\mathcal{G}_{cpc}$ and $E^T\mathcal{G}_{cpo}E$ are symmetric and positive semidefinite. In this case there exists a nonsingular matrix $\check{T}$ such that

$$\check{T}\mathcal{G}_{cpc}\check{T}^T = \begin{bmatrix} \Sigma_1 & & & 0 \\ & \Sigma_2 & & \\ & & 0 & \\ 0 & & & 0 \end{bmatrix}, \qquad \check{T}^{-T}E^T\mathcal{G}_{cpo}E\check{T}^{-1} = \begin{bmatrix} \Sigma_1 & & & 0 \\ & 0 & & \\ & & \Sigma_3 & \\ 0 & & & 0 \end{bmatrix},$$

where Σ_1, Σ_2 and Σ_3 are diagonal matrices with positive diagonal elements [176, p.76]. Then we get

$$\check{T}\Phi_{c,1}\check{T}^{-1} = \check{T}\mathcal{G}_{cpc}E^T\mathcal{G}_{cpo}E\check{T}^{-1} = \begin{bmatrix} \Sigma_1^2 & 0 \\ 0 & 0 \end{bmatrix} = \check{T}^{-T}E^T\mathcal{G}_{cpo}E\check{T}^T = \check{T}^{-T}\Phi_{c,4}\check{T}^T,$$

i.e., the matrices $\Phi_{c,1}$ and $\Phi_{c,4}$ are similar to the same diagonal positive semidefinite matrix. Moreover, we have $\Phi_{c,1}^T = \Phi_{c,4}$. Analogously, it can be shown that $\Phi_{c,2}^T = \Phi_{c,3}$ and the matrices $\Phi_{c,2}$, $\Phi_{c,3}$ are diagonalizable and have the same eigenvalues that are real and non-negative.

We will now show that the matrices $\Phi_{c,1}$ and $\Phi_{c,2}$ have the same non-zero eigenvalues. Let $\lambda \neq 0$ be an eigenvalue of $\Phi_{c,1}$ and let $v \neq 0$ be a corresponding eigenvector. We have $\Phi_{c,1} v = \mathcal{G}_{cpc} E^T \mathcal{G}_{cpo} E v = \lambda v \neq 0$. Then $Ev \neq 0$ and $\Phi_{c,2}(Ev) = E\mathcal{G}_{cpc} E^T \mathcal{G}_{cpo}(Ev) = \lambda E v$, that is, Ev is an eigenvector of $\Phi_{c,2}$ corresponding to the eigenvalue λ. □

A similar result is valid for the matrices

$$\begin{aligned} \Psi_{c,1} &:= \mathcal{G}_{cic} A^T \mathcal{G}_{cio} A, \\ \Psi_{c,2} &:= A \mathcal{G}_{cic} A^T \mathcal{G}_{cio}, \\ \Psi_{c,3} &:= \mathcal{G}_{cio} A \mathcal{G}_{cic} A^T, \\ \Psi_{c,4} &:= A^T \mathcal{G}_{cio} A \mathcal{G}_{cic}. \end{aligned} \tag{7.7}$$

Lemma 7.4. *The matrices $\Psi_{c,q}$, $q = 1, \ldots, 4$, given in* (7.7) *are diagonalizable and have the same eigenvalues. These eigenvalues are real and non-negative.*

The matrices $\Phi_{c,q}$ and $\Psi_{c,q}$ play the same role for descriptor systems as the product of the controllability and observability Gramians for standard state space systems [58].

Definition 7.5. Let $\lambda E - A$ be a c-stable pencil and let n_f and n_∞ be the dimensions of the deflating subspaces of $\lambda E - A$ corresponding to the finite and infinite eigenvalues, respectively. The square roots of the n_f largest eigenvalues of the matrix $\Phi_{c,1}$ denoted by ς_j, are called the *proper Hankel singular values* of the continuous-time descriptor system (7.1). The square roots of the n_∞ largest eigenvalues of the matrix $\Psi_{c,1}$ denoted by ϑ_j, are called the *improper Hankel singular values* of system (7.1).

The proper and improper Hankel singular values together form the set of the Hankel singular values of the continuous-time descriptor system (7.1). They are input-output invariants of system (7.1). For $E = I$, the proper Hankel singular values are the classical Hankel singular values of the standard state space system [58].

Since the proper and improper controllability and observability Gramians are symmetric and positive semidefinite, there exist Cholesky factorizations

$$\begin{aligned} \mathcal{G}_{cpc} &= R_p R_p^T, & \mathcal{G}_{cpo} &= L_p^T L_p, \\ \mathcal{G}_{cic} &= R_i R_i^T, & \mathcal{G}_{cio} &= L_i^T L_i, \end{aligned} \tag{7.8}$$

where the matrices R_p, L_p, R_i, $L_i \in \mathbb{R}^{n,n}$ are Cholesky factors [99]. The following lemma gives a connection between the proper and improper Hankel singular values and the standard singular values of the matrices $L_p E R_p$ and $L_i A R_i$.

Lemma 7.6. *Assume that the descriptor system* (7.1) *is c-stable. Consider the Cholesky factorizations* (7.8) *of the Gramians of* (7.1)*. Then the proper Hankel singular values are the n_f largest singular values of the matrix $L_p E R_p$, while the improper Hankel singular values are the n_∞ largest singular values of the matrix $L_i A R_i$.*

Proof. We have

$$\varsigma_j^2 = \lambda_j(\mathcal{G}_{cpc}E^T\mathcal{G}_{cpo}E) = \lambda_j(R_pR_p^TE^TL_p^TL_pE) = \lambda_j(R_p^TE^TL_p^TL_pER_p) = \sigma_j^2(L_pER_p),$$
$$\vartheta_j^2 = \lambda_j(\mathcal{G}_{cic}A^T\mathcal{G}_{cio}A) = \lambda_j(R_iR_i^TA^TL_i^TL_iA) = \lambda_j(R_i^TA^TL_i^TL_iAR_i) = \sigma_j^2(L_iAR_i),$$

where $\lambda_j(\cdot)$ and $\sigma_j(\cdot)$ denote the eigenvalues and the singular values of a matrix ordered decreasingly. □

As a consequence of Corollaries 4.55, 4.56, 4.58 and Lemma 7.6 we obtain the following result.

Corollary 7.7. *Consider the descriptor system* (7.1). *Assume that $\lambda E - A$ is c-stable.*

1. *System* (7.1) *is R-controllable and R-observable if and only if all its proper Hankel singular values are non-zero.*

2. *System* (7.1) *is I-controllable and I-observable if all its improper Hankel singular values are non-zero.*

3. *System* (7.1) *is S-controllable and S-observable if all its Hankel singular values are non-zero.*

4. *System* (7.1) *is C-controllable and C-observable if and only if all its Hankel singular values are non-zero.*

The following example shows that the condition for system (7.1) to be I-controllable and I-observable does not imply that all the improper Hankel singular values of (7.1) are non-zero.

Example 7.8. The descriptor system (7.1) with

$$E = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \qquad C = [\,2,\ 1\,]$$

is I-controllable and I-observable. The improper controllability and observability Gramians have the form

$$\mathcal{G}_{cic} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \mathcal{G}_{cio} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

and, hence, the improper Hankel singular value is $\vartheta = 0$.

The same example can be used to demonstrate that the S-controllable and S-observable descriptor system may have zero Hankel singular values.

7.3 Balancing of descriptor systems

As mentioned above, for a given transfer function $\mathbf{G}(s)$, there are many different realizations. Here we are interesting only in particular realizations that are most useful in the model reduction.

Definition 7.9. A realization $[\,E, A, B, C\,]$ of the transfer function $\mathbf{G}(s)$ is called R-*minimal* if the triplet (E, A, B) is R-controllable and the triplet (E, A, C) is R-observable.

Definition 7.10. A realization $[\,E, A, B, C\,]$ of the c-stable transfer function $\mathbf{G}(s)$ is called *proper balanced* if the proper controllability and observability Gramians $\mathcal{G}_{cpc}$ and $\mathcal{G}_{cpo}$ are equal and diagonal.

We will show that for a R-minimal realization $[\,E, A, B, C\,]$ of the c-stable transfer function $\mathbf{G}(s)$, there exists a system equivalence transformation (W_b^T, T_b) such that the realization

$$[\,W_b^T E T_b,\; W_b^T A T_b,\; W_b^T B,\; C T_b\,] \tag{7.9}$$

is proper balanced.

Consider the Cholesky factors R_p and L_p of the proper controllability and observability Gramians as in (7.8). If (E, A, B) is R-controllable and (E, A, C) is R-observable, then by Corollary 4.58 we have $\operatorname{rank}(\mathcal{G}_{cpc}) = \operatorname{rank}(\mathcal{G}_{cpo}) = n_f$. Compute the QR decompositions

$$R_p^T = Q_c \begin{bmatrix} R^T \\ 0 \end{bmatrix}, \qquad\qquad L_p = Q_o \begin{bmatrix} L \\ 0 \end{bmatrix},$$

where Q_c, Q_o are orthogonal and $R^T, L \in \mathbb{R}^{n_f,n}$ have full rank. Then $\mathcal{G}_{cpc} = R_p R_p^T = RR^T$, $\mathcal{G}_{cpo} = L_p^T L_p = L^T L$ and $\varsigma_j = \sigma_j(LER)$. It follows from Corollary 7.7 that the matrix $LER \in \mathbb{R}^{n_f,n_f}$ is nonsingular. Let

$$LER = U_f \Sigma V_f^T \tag{7.10}$$

be a singular value decomposition of LER, where U_f and V_f are orthogonal matrices and $\Sigma = \operatorname{diag}(\varsigma_1, \ldots, \varsigma_{n_f})$ is nonsingular. Consider the matrices

$$W_b = \left[\, L^T U_f \Sigma^{-1/2},\;\; W_\infty \,\right], \qquad W_b' = \left[\, ERV_f \Sigma^{-1/2},\;\; W_\infty' \,\right] \tag{7.11}$$

and

$$T_b = \left[\, RV_f \Sigma^{-1/2},\;\; T_\infty \,\right], \qquad T_b' = \left[\, E^T L^T U_f \Sigma^{-1/2},\;\; T_\infty' \,\right]. \tag{7.12}$$

Here the columns of matrices W_∞ and T_∞ span, respectively, the left and right deflating subspaces of the pencil $\lambda E - A$ corresponding to the infinite eigenvalues, and matrices W_∞' and T_∞' satisfy $W_\infty^T W_\infty' = (T_\infty')^T T_\infty = I_{n_\infty}$. Clearly, for P_r and P_l as in (2.3), we have $I - P_r = T_\infty (T_\infty')^T$ and $I - P_l = W_\infty' W_\infty^T$. Since

$$\begin{aligned}
(I - P_r) RR^T (I - P_r)^T &= (I - P_r)\mathcal{G}_{cpc}(I - P_r)^T = 0, \\
(I - P_l)^T L^T L (I - P_l) &= (I - P_l)^T \mathcal{G}_{cpo}(I - P_l) = 0,
\end{aligned}$$

we obtain that

$$R^T T'_\infty = 0 \qquad \text{and} \qquad L W'_\infty = 0. \tag{7.13}$$

Then

$$(T'_b)^T T_b = \begin{bmatrix} \Sigma^{-1/2} U_f^T LERV_f \Sigma^{-1/2} & \Sigma^{-1/2} U_f^T LET_\infty \\ (T'_\infty)^T RV_f \Sigma^{-1/2} & (T'_\infty)^T T_\infty \end{bmatrix} = I_n,$$

i.e., the matrices T_b and T'_b are nonsingular and $(T'_b)^T = T_b^{-1}$. Similarly, we can show that the matrices W_b and W'_b are also nonsingular and $(W'_b)^T = W_b^{-1}$.

Using (7.10)-(7.13), we obtain that the proper controllability and observability Gramians of the transformed system (7.9) have the form

$$T_b^{-1} \mathcal{G}_{cpc} T_b^{-T} = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} = W_b^{-1} \mathcal{G}_{cpo} W_b^{-T},$$

where $\Sigma = \operatorname{diag}(\varsigma_1, \ldots, \varsigma_{n_f})$ with the proper Hankel singular values ς_j. Thus, (W_b^T, T_b) with W_b and T_b as in (7.11) and (7.12), respectively, is the balancing transformation and realization (7.9) is proper balanced.

Just as for standard state space systems [58, 119], the balancing transformation for descriptor systems is not unique.

Remark 7.11. Note that the pencil $\lambda E_b - A_b = W_b^T(\lambda E - A)T_b$ is in Weierstrass-like canonical form. Indeed, from (7.10)-(7.12) we have

$$\begin{aligned} E_b &= \begin{bmatrix} \Sigma^{-1/2} U^T LERV \Sigma^{-1/2} & \Sigma^{-1/2} U^T LET_\infty \\ W_\infty^T ERV \Sigma^{-1/2} & W_\infty^T ET_\infty \end{bmatrix} = \begin{bmatrix} I_{n_f} & 0 \\ 0 & E_\infty \end{bmatrix}, \\ A_b &= \begin{bmatrix} \Sigma^{-1/2} U^T LARV \Sigma^{-1/2} & \Sigma^{-1/2} U^T LAT_\infty \\ W_\infty^T ARV \Sigma^{-1/2} & W_\infty^T AT_\infty \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & A_\infty \end{bmatrix}, \end{aligned}$$

where $A_1 = \Sigma^{-1/2} U^T LARV \Sigma^{-1/2}$, $E_\infty = W_\infty^T ET_\infty$ is nilpotent and $A_\infty = W_\infty^T AT_\infty$ is nonsingular. Clearly, the pencil $\lambda E_b - A_b$ is regular, c-stable and has the same index as $\lambda E - A$.

7.4 Balanced truncation

In the previous section we have considered a reduction of an R-minimal realization to proper balanced form. However, computing the proper balanced realization may be ill-conditioned as soon as Σ in (7.10) has small singular values. In addition, if the realization is not R-minimal, then the matrix Σ is singular. In the similar situation for standard state space systems one performs a model reduction by truncating the state components corresponding to the zero and small Hankel singular values without significant changes of the system properties, see, e.g., [119, 156]. This procedure is known as *projection of dynamics* or *balanced truncation*. It can also be applied to the descriptor system (7.1).

The proper controllability and observability Gramians can be used to describe the future output energy

$$\mathbf{E}_y := \int_0^\infty y^T(t)y(t)\,dt$$

and the minimal past proper input energy

$$\mathbf{E}_u := \min_{u\in\mathbb{L}_2^m(\mathbb{R}^-)} \int_{-\infty}^0 u^T(t)u(t)\,dt \tag{7.14}$$

that is needed to reach from $x(-\infty) = 0$ the state $x(0) = x^0 \in \operatorname{Im} P_r$. Here $\mathbb{R}^- = (-\infty, 0)$ and $\mathbb{L}_2^m(\mathbb{R}^-)$ is the Hilbert space of all square integrable functions $f : \mathbb{R} \longrightarrow \mathbb{R}^m$ such that $f(t) = 0$ for $t \geq 0$.

Theorem 7.12. *Consider a descriptor system* (7.1) *that is c-stable and R-minimal. Let* $\mathcal{G}_{cpc}$ *and* $\mathcal{G}_{cpo}$ *be the proper controllability and observability Gramians of* (7.1)*. If* $x^0 \in \operatorname{Im} P_r$ *and* $u(t) = 0$ *for* $t \geq 0$*, then*

$$\mathbf{E}_y = (x^0)^T E^T \mathcal{G}_{cpo} E x^0.$$

Moreover, for $u_{opt}(t) = B^T \mathcal{F}(-t)\mathcal{G}_{cpc}^- x^0$*, we have*

$$\mathbf{E}_{u_{opt}} = (x^0)^T \mathcal{G}_{cpc}^- x^0,$$

where $\mathcal{G}_{cpc}^-$ *is the unique solution of*

$$\begin{aligned} \mathcal{G}_{cpc}\mathcal{G}_{cpc}^-\mathcal{G}_{cpc} &= \mathcal{G}_{cpc}, \\ P_r^T \mathcal{G}_{cpc}^- P_r &= \mathcal{G}_{cpc}^-. \end{aligned} \tag{7.15}$$

Proof. System (7.1) with $x^0 \in \operatorname{Im} P_r$ and $u(t) = 0$ for $t \geq 0$ has a unique solution given by $x(t) = \mathcal{F}(t)Ex^0$. Then $y(t) = C\mathcal{F}(t)Ex^0$ for $t \geq 0$ and, hence,

$$\mathbf{E}_y = \int_0^\infty y^T(t)y(t)\,dt = \int_0^\infty (x^0)^T E^T \mathcal{F}^T(t) C^T C \mathcal{F}(t) E x^0\,dt = (x^0)^T E^T \mathcal{G}_{cpo} E x^0.$$

Consider now the minimization problem (7.14) subject to the constraint for the initial conditions

$$x^0 = \int_{-\infty}^0 \mathcal{F}(-t)Bu(t)\,dt. \tag{7.16}$$

Let $\mu \in \mathbb{R}^n$ be a Lagrange multiplier vector and let

$$L(u(t),\mu) = \int_{-\infty}^0 u^T(t)u(t)\,dt + \mu^T \left(x^0 - \int_{-\infty}^0 \mathcal{F}(-t)Bu(t)\,dt \right)$$

be the Lagrange function. For any variations $\Delta u(t)$ and $\Delta\mu$ we have that

$$\begin{aligned} \Delta L(u(t),\mu) &= 2\int_{-\infty}^0 u^T(t)\Delta u(t)\,dt - \mu^T \int_{-\infty}^0 \mathcal{F}(-t)B\Delta u(t)\,dt \\ &+ \Delta\mu^T \left(x^0 - \int_{-\infty}^0 \mathcal{F}(-t)Bu(t)\,dt \right) = 0 \end{aligned}$$

if and only if (7.16) holds and

$$u^T(t) = \frac{1}{2}\mu^T \mathcal{F}(-t)B = \frac{1}{2}\mu^T P_r \mathcal{F}(-t)B. \tag{7.17}$$

Substitution of (7.17) in (7.16) gives

$$x^0 = \frac{1}{2}\int_{-\infty}^{0} \mathcal{F}(-t)BB^T\mathcal{F}^T(-t)\mu\,dt = \frac{1}{2}\int_{0}^{\infty} \mathcal{F}(t)BB^T\mathcal{F}^T(t)\mu\,dt = \frac{1}{2}\mathcal{G}_{cpc}\mu. \tag{7.18}$$

Using (2.2) and (3.3) we obtain from the projected GCALE (4.42) that the proper controllability Gramian $\mathcal{G}_{cpc}$ has the form

$$\mathcal{G}_{cpc} = T^{-1}\begin{bmatrix} G_1 & 0 \\ 0 & 0 \end{bmatrix} T^{-T}, \tag{7.19}$$

where G_1 is a unique symmetric solution of the Lyapunov equation $JG_1 + G_1J^T = -B_1B_1^T$. Since (E, A, B) is R-controllable, the matrix G_1 is positive definite. In this case equation (7.15) has a unique solution $\mathcal{G}_{cpc}^-$ given by

$$\mathcal{G}_{cpc}^- = T^T\begin{bmatrix} G_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} T. \tag{7.20}$$

It follows from (7.18) that $2\mathcal{G}_{cpc}^- x^0 = \mathcal{G}_{cpc}^-\mathcal{G}_{cpc}\mu = P_r^T\mu$. Hence, for the optimal input

$$u_{opt}(t) = B^T\mathcal{F}^T(-t)\mathcal{G}_{cpc}^- x^0,$$

we have that

$$\begin{aligned} \mathbf{E}_{u_{opt}} &= \int_{-\infty}^{0} u_{opt}^T(t)u_{opt}(t)\,dt = \int_{-\infty}^{0} (x^0)^T(\mathcal{G}_{cpc}^-)^T\mathcal{F}(-t)BB^T\mathcal{F}^T(-t)\mathcal{G}_{cpc}^- x^0\,dt \\ &= (x^0)^T(\mathcal{G}_{cpc}^-)^T\left(\int_{0}^{\infty} \mathcal{F}(t)BB^T\mathcal{F}^T(t)\,dt\right)\mathcal{G}_{cpc}^- x^0 = (x^0)^T\mathcal{G}_{cpc}^- x^0. \end{aligned}$$

□

Remark 7.13. Using (7.19) and (7.20) we obtain the relationships

$$\mathcal{G}_{cpc}\mathcal{G}_{cpc}^- = P_r, \qquad \mathcal{G}_{cpc}^-\mathcal{G}_{cpc} = P_r^T, \qquad \mathcal{G}_{cpc}^-\mathcal{G}_{cpc}\mathcal{G}_{cpc}^- = \mathcal{G}_{cpc}^-.$$

The latter together with the first equation in (7.15) implies that $\mathcal{G}_{cpc}^-$ is a $(1,2)$-pseudo-inverse of $\mathcal{G}_{cpc}$, see [32]. The second equation in (7.15) provides the uniqueness of $\mathcal{G}_{cpc}^-$. However, if $P_r^T = P_r$, then $\mathcal{G}_{cpc}^-$ is the Moore-Penrose inverse [32] of $\mathcal{G}_{cpc}$.

Theorem 7.12 shows that a large input energy $\mathbf{E}_u$ is required to reach from $x(-\infty) = 0$ any state $x(0) = P_r x^0$ which lies in an invariant subspace of the proper controllability Gramian $\mathcal{G}_{cpc}$ corresponding to its small non-zero eigenvalues. Moreover, if x^0 is contained in an invariant subspace of the matrix $E^T\mathcal{G}_{cpo}E$ corresponding to its small non-zero eigenvalues, then the initial value $x(0) = P_r x^0$ has a small effect on the output energy $\mathbf{E}_y$. For the proper balanced system, $\mathcal{G}_{cpc}$ and $E^T\mathcal{G}_{cpo}E$ are diagonal and equal. In this case the states related to the small proper Hankel singular values are less important from the energy point of view and they may be truncated without change system properties significantly.

Let $[\,E, A, B, C\,]$ be a realization (not necessarily R-minimal) of the c-stable transfer function $\mathbf{G}(s)$. Consider the full rank factorizations $\mathcal{G}_{cpc} = R^T R$ and $\mathcal{G}_{cpo} = LL^T$, where the matrices $R \in \mathbb{R}^{n,r_c}$, $L^T \in \mathbb{R}^{n,r_o}$ have full column rank and $r_c = \operatorname{rank}(\mathcal{G}_{cpc}) \leq n_f$, $r_o = \operatorname{rank}(\mathcal{G}_{cpo}) \leq n_f$. Let

$$LER = [\,U_1,\, U_0\,] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_0 \end{bmatrix} [\,V_1,\, V_0\,]^T \tag{7.21}$$

be an "economy size" singular value decomposition of $LER \in \mathbb{R}^{r_o,r_c}$, where $[\,U_1,\, U_0\,] \in \mathbb{R}^{r_o,r}$ and $[\,V_1,\, V_0\,] \in \mathbb{R}^{r_c,r}$ have orthogonal columns,

$$\Sigma_1 = \operatorname{diag}(\varsigma_1, \ldots, \varsigma_{\ell_f}) \qquad \text{and} \qquad \Sigma_0 = \operatorname{diag}(\varsigma_{\ell_f+1}, \ldots, \varsigma_r)$$

with $\varsigma_1 \geq \varsigma_2 \geq \ldots \geq \varsigma_{\ell_f} \gg \varsigma_{\ell_f+1} \geq \ldots \geq \varsigma_r > 0$ and $r = \operatorname{rank}(\mathcal{G}_{cpc}E^T\mathcal{G}_{cpo}E) \leq \min(r_c, r_o)$. Then the reduced order realization can be computed as

$$\left[\begin{array}{c|c} s\widetilde{E} - \widetilde{A} & \widetilde{B} \\ \hline \widetilde{C} & 0 \end{array}\right] = \left[\begin{array}{c|c} W_\ell^T(sE - A)T_\ell & W_\ell^T B \\ \hline CT_\ell & 0 \end{array}\right], \tag{7.22}$$

where

$$W_\ell = \left[\, L^T U_1 \Sigma_1^{-1/2}, \quad W_\infty \,\right] \in \mathbb{R}^{n,\ell}, \qquad T_\ell = \left[\, RV_1\Sigma_1^{-1/2}, \quad T_\infty \,\right] \in \mathbb{R}^{n,\ell} \tag{7.23}$$

and $\ell = \ell_f + n_\infty$. Here W_∞ and T_∞ form the bases of the left and right deflating subspaces, respectively, corresponding to the infinite eigenvalues of $\lambda E - A$.

Note that computing the reduced order descriptor system can be interpreted as performing a system equivalence transformation $(\check{W}, \check{T})$ such that

$$\left[\begin{array}{c|c} \check{W}(sE - A)\check{T} & \check{W}B \\ \hline C\check{T} & 0 \end{array}\right] = \left[\begin{array}{cc|c} sE_f - A_f & 0 & B_f \\ 0 & sE_\infty - A_\infty & B_\infty \\ \hline C_f & C_\infty & 0 \end{array}\right],$$

where the pencil $\lambda E_f - A_f$ has only finite eigenvalues, while all eigenvalues of $\lambda E_\infty - A_\infty$ are infinite, and then reducing the order of the subsystem $[\,E_f,\, A_f,\, B_f,\, C_f\,]$ with nonsingular E_f. Clearly, the reduced order system (7.22) is c-stable, R-minimal and proper balanced. Choosing ℓ_f in (7.21) as a maximal integer such that $\varsigma_{\ell_f} > 0$, this procedure can be used to compute the R-minimal realization of the transfer function $\mathbf{G}(s) = C(sE - A)^{-1}B$. The described decoupling of system matrices is equivalent to the additive decomposition of the transfer function as $\mathbf{G}(s) = \mathbf{G}_p(s) + \mathbf{P}(s)$, where $\mathbf{G}_p(s) = C_f(sE_f - A_f)^{-1}B_f$ is the proper part and $\mathbf{P}(s) = C_\infty(sE_\infty - A_\infty)^{-1}B_\infty$ is the polynomial part of $\mathbf{G}(s)$. The transfer function of the reduced system has the form $\widetilde{\mathbf{G}}(s) = \widetilde{\mathbf{G}}_p(s) + \mathbf{P}(s)$, where $\widetilde{\mathbf{G}}_p(s) = \widetilde{C}_f(s\widetilde{E}_f - \widetilde{A}_f)^{-1}\widetilde{B}_f$ is the reduced subsystem. In this case the difference $\mathbf{G}(s) - \widetilde{\mathbf{G}}(s) = \mathbf{G}_p(s) - \widetilde{\mathbf{G}}_p(s)$ is a proper rational function, and we have the following upper bound on the $\mathbb{H}_\infty$-norm of the error system

$$\|\mathbf{G} - \widetilde{\mathbf{G}}\|_{\mathbb{H}_\infty} := \sup_{\omega\in\mathbb{R}} \|\mathbf{G}(i\omega) - \widetilde{\mathbf{G}}(i\omega)\|_2 \leq 2(\varsigma_{\ell_f+1} + \ldots + \varsigma_{n_f}) \tag{7.24}$$

that has been derived in [46, 58]. Thus, if we remove the states corresponding to small proper Hankel singular values, then the approximation error is small and the reduced order system is a good approximation to (7.1) in the $\mathbb{H}_\infty$-norm.

7.5 Numerical algorithms

To reduce the order of the descriptor system (7.1) we have to compute the full rank factors L and R of the proper observability and controllability Gramians that satisfy the projected generalized Lyapunov equations (4.39) and (4.42), respectively. We also need the matrices W_∞ and T_∞, whose columns span the left and right infinite deflating subspaces of the pencil $\lambda E - A$. The projected generalized Lyapunov equations (4.39) and (4.42) can be solved for the full rank factors via the generalized Schur-Hammarling method, see Algorithms 5.2.1 and 5.2.2. Simultaneously, this method produces the matrices W_∞ and T_∞. Indeed, if the pencil $\lambda E - A$ is reduced to the GUPTRI form (2.4), then W_∞ and T_∞ are computed as

$$W_\infty = V \begin{bmatrix} 0 \\ I_{n_\infty} \end{bmatrix} \qquad \text{and} \qquad T_\infty = U \begin{bmatrix} Y \\ I_{n_\infty} \end{bmatrix}, \tag{7.25}$$

where Y satisfy the generalized Sylvester equation (5.2).

The following algorithm is a generalization of the *square root balanced truncation method* [102, 156] for the descriptor system (7.1).

Algorithm 7.5.1. *Generalized Square Root (GSR) method.*
Input: *A realization* $[\,E,\ A,\ B,\ C\,]$ *such that* $\lambda E - A$ *is c-stable.*
Output: *A reduced order system* $[\,\widetilde{E},\ \widetilde{A},\ \widetilde{B},\ \widetilde{C}\,]$.
Step 1. *Use* Algorithms 5.2.1 *and* 5.2.2 *to compute the full rank factors* L *and* R *of the proper observability and controllability Gramians* $\mathcal{G}_{cpo} = L^T L$ *and* $\mathcal{G}_{cpc} = RR^T$ *as well as the matrices* W_∞ *and* T_∞ *given in* (7.25).
Step 2. *Compute the "economy size" singular value decomposition* (7.21).
Step 3. *Compute the matrices* $W_\ell = [\,L^T U_1 \Sigma_1^{-1/2},\ \ W_\infty\,]$ *and* $T_\ell = [\,RV_1\Sigma_1^{-1/2},\ \ T_\infty\,]$.
Step 4. *Compute the reduced order system* $[\widetilde{E},\ \widetilde{A},\ \widetilde{B},\ \widetilde{C}] = [\,W_\ell^T E T_\ell,\ W_\ell^T A T_\ell,\ W_\ell^T B,\ C T_\ell\,]$.

If the original system (7.1) is highly unbalanced, then the matrices W_ℓ and T_ℓ are ill-conditioned. To avoid accuracy loss in the reduced system, a *square root balancing free* method has been proposed for standard state space systems in [164]. This approach can be generalized for descriptor systems as follows.

Algorithm 7.5.2. *Generalized Square Root Balancing Free (GSRBF) method.*
Input: *A realization* $[E, A, B, C]$ *such that* $\lambda E - A$ *is c-stable.*
Output: *A reduced order system* $[\widetilde{E}, \widetilde{A}, \widetilde{B}, \widetilde{C}]$.
Step 1. *Use* Algorithms 5.2.1 *and* 5.2.2 *to compute the full rank factors* L *and* R *of the proper observability and controllability Gramians* $\mathcal{G}_{cpo} = L^T L$ *and* $\mathcal{G}_{cpc} = RR^T$ *as well as the matrices* W_∞ *and* T_∞ *given in* (7.25).
Step 2. *Compute the "economy size" singular value decomposition* (7.21).
Step 3. *Compute the "economy size" QR decompositions*

$$RV_1 = Q_R R_0, \qquad L^T U_1 = Q_L L_0,$$

where Q_R, $Q_L \in \mathbb{R}^{n,\ell_f}$ *have orthogonal columns and* R_0, $L_0 \in \mathbb{R}^{\ell_f,\ell_f}$ *are upper triangular, nonsingular.*
Step 4. *Compute the reduced order system* $[\widetilde{E}, \widetilde{A}, \widetilde{B}, \widetilde{C}] = [W_\ell^T E T_\ell, W_\ell^T A T_\ell, W_\ell^T B, C T_\ell]$, *where* $W_\ell = [Q_L, \ W_\infty]$ *and* $T_\ell = [Q_R, \ T_\infty]$.

The GSR and GSRBF methods are mathematically equivalent in the sense that they deliver a reduced system with the same transfer function. But the matrices W_ℓ and T_ℓ computed by the GSRBF method are often significantly better conditioned than those computed via the GSR method.

Remark 7.14. In fact, we do not need to compute the full rank Cholesky factors R and L and the matrices W_∞ and T_∞. From (2.4) and (7.25) we have $W_\infty^T E T_\infty = E_\infty$, $W_\infty^T A T_\infty = A_\infty$, $W_\infty^T B = B_\infty$ and $C T_\infty = C_f Y + C_2 = C_\infty$. Moreover, it follows from (2.4), (5.26) and (5.29) that $LER = L_1 E_f R_1$. Thus, computation of the proper Hankel singular values in Step 2 of Algorithms 7.5.1 and 7.5.2 can be performed working only with the matrices L_1, E_f and R_1. This reduces the computational cost and the memory requirement. Note that the singular value decomposition of $L_1 E_f R_1$ may be computed without forming this product explicitly, see [66] for details.

7.6 Numerical examples

In this section we consider numerical examples to illustrate the reliability of the proposed model reduction methods for descriptor systems. All of the following results were obtained on an IBM RS 6000 44P Model 270 with relative machine precision $\epsilon = 2.22 \times 10^{-16}$ using MATLAB mex-functions based on the `GUPTRI` routine [41, 42] and the SLICOT library routines `SB04OD` and `SG03BD` [16, 165].

Example 7.15. Consider the holonomically constrained planar model of a truck [138]. The linearized equation of motion has the form

$$\begin{aligned} \dot{\mathbf{p}}(t) &= \mathbf{v}(t), \\ M\dot{\mathbf{v}}(t) &= K\mathbf{p}(t) + D\mathbf{v}(t) - G^T \boldsymbol{\lambda}(t) + B_2 u(t), \\ 0 &= G\mathbf{p}(t), \end{aligned} \tag{7.26}$$

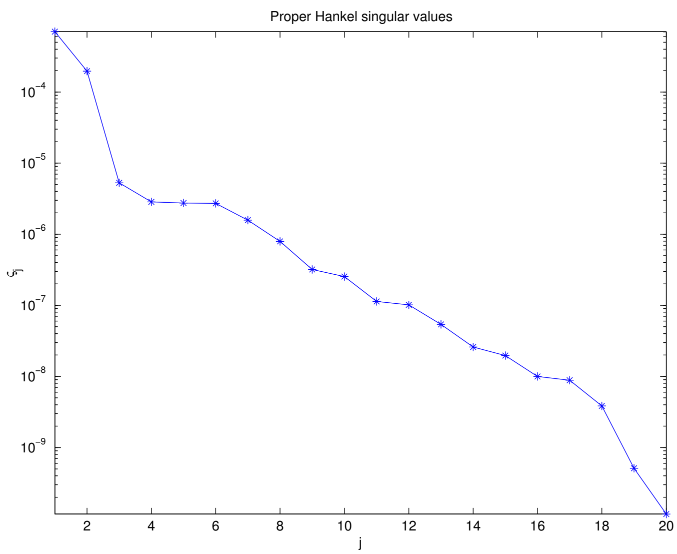


Figure 7.1: Proper Hankel singular values of the linearized truck model

where $\mathbf{p}(t) \in \mathbb{R}^{11}$ is the position vector, $\mathbf{v}(t) \in \mathbb{R}^{11}$ is the velocity vector, $\boldsymbol{\lambda}(t) \in \mathbb{R}$ is the Lagrange multiplier, M is the positive definite mass matrix, K is the stiffness matrix, D is the damping matrix, G is the constraint matrix and B_2 is the input matrix. System (7.26) together with the output equation $y(t) = \mathbf{p}(t)$ forms a descriptor system of order $n = 23$ with $m = 1$ input and $p = 11$ outputs. The dimension of the deflating subspace corresponding to the finite eigenvalues is $n_f = 20$.

Figure 7.1 shows the proper Hankel singular values ς_j. We approximate system (7.26) by a model of order $\ell = 5$. Figure 7.2 illustrates how accurate the reduced order model approximates the original one. We display the amplitude Bode plot of the error system computed as $\|\mathbf{G}(i\omega) - \widetilde{\mathbf{G}}(i\omega)\|_2$ for a frequency range $\omega \in [1, 10^3]$. Comparison of this error with the upper bound $2(\varsigma_3 + \ldots + \varsigma_{20}) = 1.69 \times 10^{-5}$ shows that the error estimate (7.24) is tight. Note that the Bode plots of the original and reduced systems, that is, the spectral norms of the frequency responses $\mathbf{G}(i\omega)$ and $\widetilde{\mathbf{G}}(i\omega)$ are not presented, since they were impossible to distinguish.

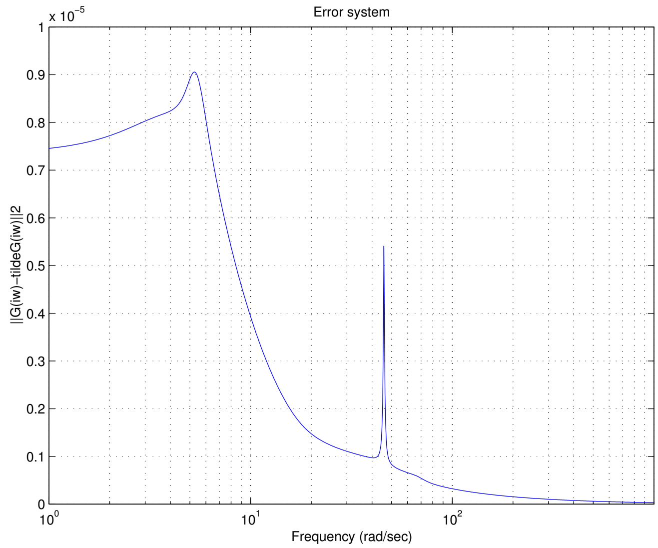


Figure 7.2: Bode plot of the error system for the linearized truck model

Example 7.16. Consider the two dimensional instationary Stokes equation describing the flow of an incompressible fluid

$$\begin{array}{rcll} \dfrac{\partial v}{\partial t} & = & \Delta v - \nabla \rho + f, & (x,t) \in \Omega \times (0, t_f), \\ 0 & = & \nabla \cdot v, & (x,t) \in \Omega \times (0, t_f), \end{array}$$

with appropriate initial and boundary conditions. Here $v(t,x) \in \mathbb{R}^2$ is the velocity vector, $\rho(t,x) \in \mathbb{R}$ is the pressure, $f(t,x) \in \mathbb{R}^2$ is the vector of external forces and $\Omega = [0,1] \times [0,1]$. Using a finite volume semidiscretization method on an uniform staggered grid [19, 170], we obtain the descriptor system

$$\begin{array}{rcl} \dot{\mathbf{v}}(t) & = & A_{11}\mathbf{v}(t) + A_{12}\boldsymbol{\rho}(t) + B_1 u(t), \\ 0 & = & A_{12}^T \mathbf{v}(t), \end{array} \tag{7.27}$$

with the output equation $y(t) = C_2\boldsymbol{\rho}(t)$. Here $\mathbf{v}(t) \in \mathbb{R}^{n_\mathbf{v}}$ is the semidiscretized vector of velocities, $\boldsymbol{\rho}(t) \in \mathbb{R}^{n_\rho}$ is the semidiscretized vector of pressures, $A_{11} = A_{11}^T \in \mathbb{R}^{n_\mathbf{v}, n_\mathbf{v}}$ is the discretized Laplace operator, $A_{12} \in \mathbb{R}^{n_\mathbf{v}, n_\rho}$ is the discretized gradient operator, $B_1 \in \mathbb{R}^{n_\mathbf{v}, m}$ is the input matrix resulting from boundary conditions and $f(t,x)$ with dimensions $n_\mathbf{v} = 480$, $n_{\boldsymbol{\rho}} = 255$, $m = 64$ and $p = 15$. The matrix A_{12} has full column rank. In this case system (7.27) is of index 2 and the dimension of the dynamic part is $n_f = n_\mathbf{v} - n_{\boldsymbol{\rho}} = 225$.

Figure 7.3 shows the proper Hankel singular values of system (7.27). We see that the proper Hankel singular values decay sufficiently fast. The dynamic part of (7.27) has been approximated by a system of order $\ell_f = 10$. The reduced order model is of order $\ell = 520$

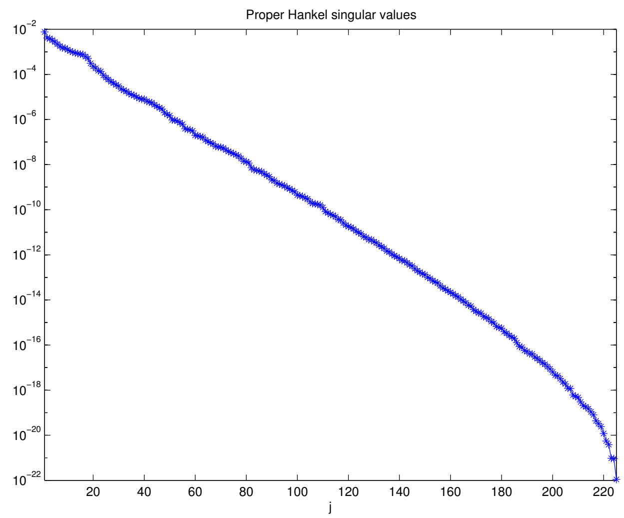


Figure 7.3: Proper Hankel singular values of the semidiscretized Stokes equation

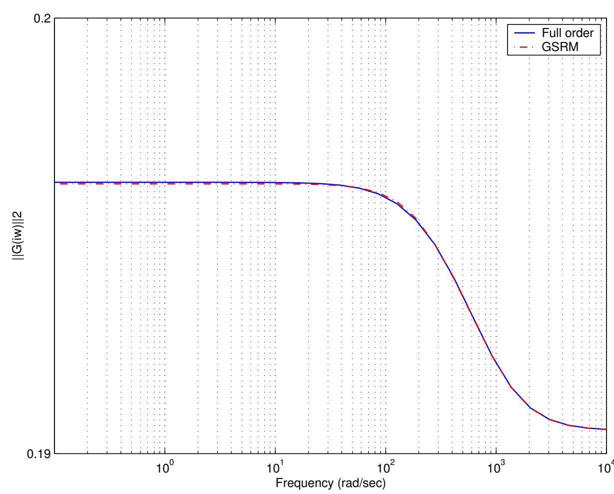


Figure 7.4: Bode plots of the original system and the reduced order system for the semidiscretized Stokes equation

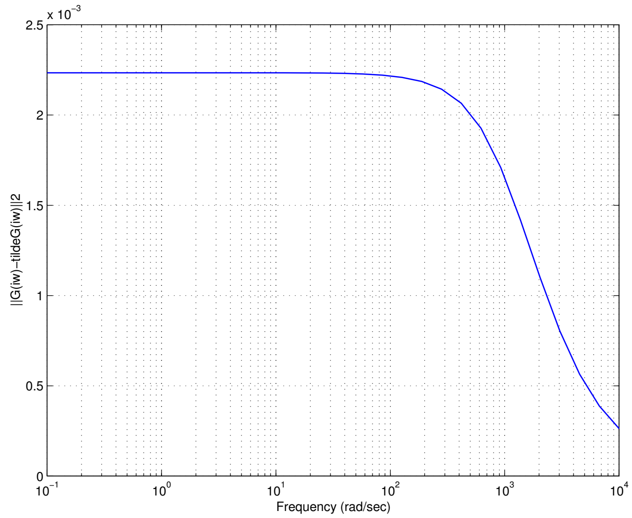


Figure 7.5: Bode plot of the error system for the semidiscretized Stokes equation

and has the form

$$\begin{aligned} \dot{\widetilde{\mathbf{v}}}(t) &= \widetilde{A}_{11}\widetilde{\mathbf{v}}(t) + \widetilde{A}_{12}\widetilde{\boldsymbol{\rho}}(t) + \widetilde{B}_1 u(t), \\ 0 &= \widetilde{A}_{21}\widetilde{\mathbf{v}}(t). \end{aligned} \tag{7.28}$$

One can see that the structure of (7.27) is preserved, but system (7.28) is no more symmetric. The latter is due to the transformation matrices W_ℓ and T_ℓ given in (7.23) include the full rank factors L and R of the solutions of the projected Lyapunov equations (4.39) and (4.42) that are not equal. However, if the output matrix C is the transpose of the input matrix B, then $W_\ell^T = T_\ell$ and the reduced order system will be symmetric.

In Figure 7.4 we compare the spectral norms of the frequency responses of the original system $\mathbf{G}(i\omega)$ and the reduced order system $\widetilde{\mathbf{G}}(i\omega)$ for a frequency rang $\omega \in [10^{-1}, 10^4]$. One can see that the full order system is approximated the reduced order system quite well. The Bode plot of the error systems is presented in Figure 7.5.

Remark 7.17. As Example 7.16 shows, the dimension of the deflating subspaces of the pencil corresponding to the infinite eigenvalues may be much larger than the dimension of the deflating subspaces corresponding to the finite eigenvalues. In this case the algebraic part of the descriptor system is much larger than the dynamic one. It is interesting, whether the order of the algebraic part can be reduced? Formally, we can transform the descriptor system such that the improper controllability and observability Gramians become diagonal and equal. Their diagonal elements are exactly the improper Hankel singular values. What happens if we truncate the states corresponding to small improper Hankel singular values. Is it possible to obtain an error estimate? These questions remain open.

Chapter 8

Conclusions

In this thesis we have presented the theoretical analysis, numerical solution and perturbation theory for generalized continuous-time and discrete-time Lyapunov equations.

The stability analysis for continuous-time and discrete-time singular systems has been considered. It is known that the singular system is asymptotically stable if and only if all the finite eigenvalues of the associated pencil lie in the open left half-plane in the continuous-time case and inside the unit circle in the discrete-time case [36, 123]. We have introduced numerical parameters that estimate the asymptotical decay of solutions of singular systems. These parameters can be used to characterize the property of matrix pencils to have all the finite eigenvalues in the open left half-plane or inside the unit circle without explicitly computing eigenvalues.

An important role in stability theory as well as in many control problems for descriptor systems play generalized Lyapunov equations. We have presented solvability and uniqueness theorems for these equations with a general right-hand side $-G$. However, some difficulties arise if one of the coefficient matrices in the continuous-time case and both the coefficient matrices in the discrete-time case are singular. Such equations may be not solvable and even if solution exists, it is not unique.

In the case of singular E we have studied generalized Lyapunov equations with a special right-hand side $-E^*GE$. For such equations, a generalization of classical Lyapunov stability theorems turned out to be only for pencils of index at most two in the continuous-time case and of index at most one in the discrete-time case.

Further, we have considered projected generalized Lyapunov equations obtained via projection in an appropriate way of the right hand-side and the solution onto the right and left deflating subspaces of the pencil corresponding to the finite eigenvalues. For such equations, necessary and sufficient conditions for existence and uniqueness of solutions have been derived. These conditions are independent of the index of matrix pencils. We have shown that projected generalized Lyapunov equations can be used to characterize the asymptotic stability of singular systems as well as controllability and observability properties of descriptor systems. Moreover, these equations are useful to generalize matrix inertia theorems to matrix pencils. Finally, we have seen that the controllability and observability Gramians of descriptor systems introduced in [11] can be computed by solving projected generalized Lyapunov equations.

Even though the numerical solution of standard Lyapunov equations has been the subject of intense research in many years, e.g. [9, 64, 72, 80, 127, 136], there are not many contributions to numerical methods for generalized Lyapunov equations [17, 55, 117, 125]. In this thesis we have proposed generalizations of the Bartels-Stewart and Hammarling methods for projected generalized Lyapunov equations and studied their numerical properties and complexity. A disadvantage of both methods is that they cost $O(n^3)$ because the computation of the GUPTRI form of a pencil is required. As a consequence, these methods can be used only for problems of small and medium size. Moreover, they do not make use the sparsity of coefficient matrices.

Large scale dense regular generalized Lyapunov equations can be solved via the matrix sign function method or Malyshev algorithm. The latter is applicable also to projected generalized discrete-time Lyapunov equations with nonsingular G in the right-hand side. A generalization of iterative methods like low-rank ADI and Smith methods as well as Krylov subspace methods for projected generalized Lyapunov equations is a subject for further research.

Also, we have developed the perturbation theory for generalized Lyapunov equations. The spectral condition numbers have been introduced and perturbation bounds for solutions of the projected generalized Lyapunov equations have been derived. In the case of nonsingular E, the spectral condition numbers are equivalent to the well-known Frobenius norm based condition numbers. However, from computational point of view the spectral condition numbers have considerable superiority.

Unfortunately, the perturbation bound for projected generalized Lyapunov equations have been obtained under assumption that perturbations in E and A do not change the dimension of the deflating subspaces of the pencil corresponding to the infinite eigenvalues. Moreover, in the continuous-time case we have supposed that the nilpotency structure of the pencil is preserved. The sensitivity theory for general perturbations and backward error analysis for projected Lyapunov equations are still open problems.

Our last topic was the model reduction of descriptor systems. For these systems proper and improper Hankel singular values have been defined and balanced truncation methods have been presented. The proper Hankel singular values can be considered as a measure for the importance of the state components. We have shown that if the c-stable continuous-time descriptor system is in a proper balanced form, that is, if the proper controllability and observability Gramians are diagonal and equal, then a large (small) amount of input energy is required to reach the states corresponding to small (large) proper Hankel singular values and these states generate a small (large) amount of output energy. Balanced truncation methods for descriptor systems are based on the decoupling these systems into dynamic and algebraic parts and reducing the order only the dynamic part by truncation of the states that are related to small proper Hankel singular values. Important properties of these methods are that the stability is preserved in the reduced order system and there is a bound on the approximation error.

Bibliography

[1] E. Anderson, Z. Bai, C.H. Bischof, S. Blackford, J.M. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. *LAPACK Users' Guide, Third Edition.* SIAM, Philadelphia, PA, 1999.

[2] A.C. Antoulas and D.C. Sorensen. Approximation of large-scale dynamical systems: an overview. *Int. J. Appl. Math. Comput. Sci.*, 11(5):1093–1121, 2001.

[3] A.C. Antoulas, D.C. Sorensen, and S. Gugercin. A modified low-rank Smith method for large-scale Lyapunov equations. Technical Report TR01-10, Department of Computational and Applied Mathematics, Rice University, 6100 Main St.-MS 134, Houston, TX 77005-1892, 2001.

[4] A.C. Antoulas, D.C. Sorensen, and S. Gugercin. A survey of model reduction methods for large-scale systems. In V. Olshevsky, editor, *Structured Matrices in Mathematics, Computer Science and Engineering, Vol. I*, Contemporary Mathematics Series, 280. American Mathematical Society, 2001.

[5] A.C. Antoulas, D.C. Sorensen, and Y. Zhou. On the decay rate of Hankel singular values and related issues. Technical Report TR01-09, Department of Computational and Applied Mathematics, Rice University, 6100 Main St.-MS 134, Houston, TX 77005-1892, 2001.

[6] J.D. Aplevich. *Implicit Linear Systems.* Number 152 in Lecture Notes in Control and Information Sciences. Springer-Verlag, New York, 1991.

[7] Z. Bai, J. Demmel, and M. Gu. An inverse free parallel spectral divide and conquer algorithm for nonsymmetric eigenproblems. *Numer. Math.*, 76:279–308, 1997.

[8] L. Balzer. Accelerated convergence of the matrix sign function method of solving Lyapunov, Riccati and other matrix equations. *Internat. J. Control*, 32:1057–1078, 1980.

[9] R.H. Bartels and G.W. Stewart. Solution of the equation $AX + XB = C$. *Comm. ACM*, 15(9):820–826, 1972.

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

[11] D.J. Bender. Lyapunov-like equations and reachability/observability Gramians for descriptor systems. *IEEE Trans. Automat. Control*, 32(4):343–348, 1987.

[12] P. Benner. *Contributions to the Numerical Solution of Algebraic Riccati Equations and Related Eigenvalue Problems*. Logos Verlag, Berlin, 1997.

[13] P. Benner and R. Byers. An arithmetic for matrix pencils. In A. Beghi, L. Finesso, and G. Picci, editors, *Mathematical Theory of Networks and Systems. Proceedings of the MTNS-98 Symposium (Padova, Italy, July, 1998)*, pages 573–576. Il Poligrafo, Padova, 1998.

[14] P. Benner and R. Byers. An arithmetic for rectangular matrix pencils. In O. Gonzalez, editor, *Proceedings of the 1999 IEEE International Symposium on Computer Aided Control System Design (Hawai'i, USA, August 22-27, 1999)*, pages 75–80, 1999.

[15] P. Benner, J.M. Claver, and E.S. Quintana-Ortí. Parallel distribited solvers for large stable generalized Lyapunov equations. *Parallel Process. Lett.*, 9(1):147–158, 1999.

[16] P. Benner, V. Mehrmann, V. Sima, S. Van Huffel, and A. Varga. SLICOT - A subroutine library in systems and control theory. *Appl. Comput. Control Signals Circuits*, 1:499–539, 1999.

[17] P. Benner and E.S. Quintana-Ortí. Solving stable generalized Lyapunov equations with the matrix sign function. *Numerical Algorithms*, 20(1):75–100, 1999.

[18] P. Benner, E.S. Quintana-Ortí, and G. Quintana-Ortí. Balanced truncation model reduction of large-scale dense systems on parallel computers. *Math. Comput. Model. Dyn. Syst.*, 6(4):383–405, 2000.

[19] T.R. Bewley. Flow control: new challenges for a new Renaissance. *Progress in Aerospace Sciences*, 37:21–58, 2001.

[20] S. Bittanti, P. Bolzern, and P. Colaneri. Inertia theorems for Lyapunov and Riccati equations – an updated view. In *Linear Algebra in Signals, Systems, and Control (Boston, MA, 1986)*, pages 11–35. SIAM, 1988.

[21] K.E. Brenan, S.L. Campbell, and L.R. Petzold. *The Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*. Elsevier, North-Holland, New York, N.Y., 1989.

[22] A.Ja. Bulgakov. An efficiently calculable parameter for the stability property of a system of linear differential equations with constant coefficients. *Siberian Math. J.*, 21(3):339–347, 1980.

[23] A.Ja. Bulgakov. The basis of guaranteed accuracy in the problem of separation of invariant subspaces for non-self-adjoint matrices. *Siberian Adv. Math.*, 1(1,2):63–108, 1–56, 1991.

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

[25] R. Byers. Solving the algebraic Riccati equation with the matrix sign function. *Linear Algebra Appl.*, 85:267–279, 1987.

[26] R. Byers. A bisection method for measuring the distance of a stable matrix to the unstable matrices. *SIAM J. Sci. Statist. Comput.*, 9:875–881, 1988.

[27] R. Byers, C. He, and V. Mehrmann. Where is the nearest non-regular pencil. *Linear Algebra Appl.*, 285:81–105, 1998.

[28] R. Byers and N.K. Nichols. On the stability radius of a generalized state-space system. *Linear Algebra Appl.*, 188/189:113–134, 1993.

[29] B.E. Cain. Inertia theory. *Linear Algebra Appl.*, 30:211–240, 1980.

[30] S.L. Campbell. *Singular Systems of Differential Equation, I.* Pitman, San Francisco, 1980.

[31] S.L. Campbell. *Singular Systems of Differential Equation, II.* Reseach Notes in Mathematics, 61. Pitman, San Francisco, 1982.

[32] S.L. Campbell and C.D. Meyer. *Generalized Inverses of Linear Transformations.* Dover Publications, New York, 1979.

[33] D. Carlson and H. Schneider. Inertia theorems: the semidefinite case. *J. Math. Anal. Appl.*, 6:430–446, 1963.

[34] K.E. Chu. The solution of the matrix equations $AXB - CXD = E$ and $(YA - DZ, YC - BZ) = (E, F)$. *Linear Algebra Appl.*, 93:93–105, 1987.

[35] D. Cobb. Controllability, observability, and duality in singular systems. *IEEE Trans. Automat. Control*, 29(12):1076–1082, 1984.

[36] L. Dai. *Singular Control Systems.* Lecture Notes in Control and Information Sciences, 118. Springer-Verlag, Berlin, Heidelberg, 1989.

[37] Ju.L. Daleckii and M.G. Krein. *Stability of Solutions of Differential equations in Banach Space.* Translations of Mathematical Monographs, 43. American Mathematical Society, Providence, RI, 1974.

[38] B.N. Datta. Stability and inertia. *Linear Algebra Appl.*, 302/303:563–600, 1999.

[39] F. Delebecque. On the resolvent approach to the spectral decomposition of a regular matrix pencil. *Linear Algebra Appl.*, 129(2):63–75, 1990.

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

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

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

[43] G. Doetsch. *Guide to the Applications of the Laplace and Z-Transforms.* Van Nostrand Reinhold Company, London, 1971.

[44] D.Y.Hu and L. Reichel. Krylov-subspace methods for the Sylvester equation. *Linear Algebra Appl.*, 172:283–313, 1992.

[45] E. Eich-Soellner and C. Führer. *Numerical methods in multibody dynamics.* B.G. Teubner, Stuttgart, 1998.

[46] D. Enns. Model reduction with balanced realization: an error bound and a frequency weighted generalization. In *Proceedings of the 23rd IEEE Conference on Decision and Control (Las Vegas, 1984)*, pages 127–132. IEEE, New York, 1984.

[47] P. Feldmann and R.W. Freund. Efficient linear circuit analysis by Padé approximation via the Lanzcos process. *IEEE Trans. Computer-Aided Design*, 14:639–649, 1995.

[48] L. Fortuna, G. Nunnari, and A. Gallo. *Model Order Reduction Techniques with Applications in Electrical Engineering.* Springer-Verlag, London, 1992.

[49] V. Frayssé, M. Gueury, F. Nicoud, and V. Toumazou. Spectral portraits for matrix pencils. Technical Report TR/PA/96/19, CERFACS, 42av. Coriolis, 31057 Toulouse Cedex, France, 1996.

[50] P.M. Gahinet, A.J. Laub, C.S. Kenney, and G.A. Hewer. Sensitivity of the stable discrete-time Lyapunov equation. *IEEE Trans. Automat. Control*, 35(11):1209–1217, 1990.

[51] Z. Gajić and M.T.J. Qureshi. *Lyapunov Matrix Equation in System Stability and Control.* Academic Press, San Diego, CA, 1995.

[52] K. Gallivan, E. Grimme, and P. Van Dooren. A rational Lanczos algorithm for model reduction. *Numerical Algorithms*, 12(1-2):33–63, 1996.

[53] F.R. Gantmacher. *Theory of Matrices.* Chelsea, New York, 1959.

[54] J.D. Gardiner and A.J. Laub. A generalization of the matrix-sign-function solution for algebraic Riccati equations. *Internat. J. Control*, 44(3):823–832, 1986.

[55] J.D. Gardiner, A.J. Laub, J.J. Amato, and C.B. Moler. Solution of the Sylvester matrix equation $AXB^T + CXD^T = E$. *ACM Trans. Math. Software*, 18(2):223–231, 1992.

[56] J.D. Gardiner, M.R.Wette, A.J. Laub, J.J. Amato, and C.B. Moler. Algorithm 705: A Fortran-77 software package for solving the Sylvester matrix equation $AXB^T + CXD^T = E$. *ACM Trans. Math. Software*, 18(2):232–238, 1992.

[57] A.R. Ghavimi and A.J. Laub. Residual bounds for discrete-time Lyapunov equations. *IEEE Trans. Automat. Control*, 40(7):1244–1249, 1995.

[58] K. Glover. All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-errors bounds. *Internat. J. Control*, 39(6):1115–1193, 1984.

[59] S.K. Godunov. Problem of the dichotomy of the spectrum of a matrix. *Siberian Math. J.*, 27(5):649–660, 1986.

[60] S.K. Godunov. Spectral portraits of matrices and criteria of spectrum dichotomy. In *Computer arithmetic and enclosure methods. Oldenburg, 1991*, pages 25–35. North-Holland, Amsterdam, 1992.

[61] S.K. Godunov. *Ordinary Differential Equations with Constant Coefficients.* Translations of Mathematical Monographs, 169. American Mathematical Society, Providence, RI, 1997.

[62] S.K. Godunov. *Modern Aspects of Linear Algebra.* Translations of Mathematical Monographs, 175. American Mathematical Society, Providence, RI, 1998.

[63] I. Gohberg, S. Goldberg, and M.A. Kaashoek. *Classes of Linear Operators I.* Birkhäuser, Basel, 1990.

[64] G.H. Golub and C.F. Van Loan. *Matrix Computations. 3rd ed.* The Johns Hopkins University Press, Baltimore, London, 1996.

[65] G.H. Golub, S. Nash, and C. Van Loan. A Hessenberg-Schur method for the problem $AX + XB = C$. *IEEE Trans. Automat. Control*, AC-24:909–913, 1979.

[66] G.H. Golub, K. Sølna, and P. Van Dooren. Computing the SVD of a general matrix product/quotient. *SIAM J. Matrix Anal. Appl.*, 22(1):1–19, 2000.

[67] E. Griepentrog and R. März. *Differential-Algebraic Equations and Their Numerical Treatment.* Teubner-Texte zur Mathematik, 88. B.G. Teubner, Leipzig, 1986.

[68] E.J. Grimme, D.C. Sorensen, and P. Van Dooren. Model reduction of state space systems via an implicitly restarted Lanczos method. *Numerical Algorithms*, 12(1-2):1–31, 1996.

[69] M. Günther and U. Feldmann. CAD-based electric-circuit modeling in industry. I. Mathematical structure and index of network equations. *Surveys Math. Indust.*, 8(2):97–129, 1999.

[70] M. Günther and U. Feldmann. CAD-based electric-circuit modeling in industry. II. Impact of circuit configurations and parameters. *Surveys Math. Indust.*, 8(2):131–157, 1999.

[71] W. Hahn. *Stability of Motion.* Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, 138. Springer-Verlag, Berlin, Heidelberg, 1967.

[72] S.J. Hammarling. Numerical solution of the stable non-negative definite Lyapunov equation. *IMA J. Numer. Anal.*, 2:303–323, 1982.

[73] C. He and G.A. Watson. An algorithm for computing the distance to instability. *SIAM J. Matrix Anal. Appl.*, 20(1):101–116, 1998.

[74] G. Hewer and C. Kenney. The sensitivity of the stable Lyapunov equation. *SIAM J. Contol Optim.*, 26(2):321–344, 1988.

[75] N.J. Higham. *Accuracy and Stability of Numerical Algorithms.* SIAM, Philadelphia, PA, 1996.

[76] N.J. Higham and F. Tisseur. More on pseudospectra for polynomial eigenvalue problems and applications in control theory. Numerical Analysis Report 372, Department of Mathematics, University of Manchester, Manchester M13 9PL, England, January 2001.

[77] D. Hinrichsen and A.J. Pritchard. Real and complex stability radii: a survey. In D. Hinrichsen and B. Mårtensson, editors, *Control of Uncertain Systems, Proc. Int. Workshop (Bremen, 1989). Progr. Systems Control Theory*, volume 6, pages 119–162. Birkhauser, Boston, MA, 1990.

[78] R.A. Horn and C.R. Johnson. *Topics in Matrix Analysis.* Cambridge University Press, Cambridge, 1991.

[79] V. Ionescu, C. Oară, and M. Weiss. *Generalized Riccati Theory and Robust Control: A Popov Function Approach.* John Wiley and Sons, Chichester, UK, 1999.

[80] I.M. Jaimoukha and E.M. Kasenally. Krylov subspace methods for solving large Lyapunov equations. *SIAM J. Numer. Anal.*, 31(1):227–251, 1994.

[81] I. Jonsson and B. Kågström. Recursive bloked algorithms for solving triangular systems – Part I: One-sided and coupled Sylvester-type matrix equations. *ACM Trans. Math. Software*, 28(4):392–415, 2002.

[82] I. Jonsson and B. Kågström. Recursive bloked algorithms for solving triangular systems – Part II: Two-sided and generalized Sylvester and Lyapunov matrix equations. *ACM Trans. Math. Software*, 28(4):416–435, 2002.

[83] T. Kaczorek. *Linear Control Systems. Vol. 1. Analysis of Multivariable Systems*. Industrial Control, Computers and Communications Series, 6. Research Studies Press, Taunton, England, 1992.

[84] B. Kågström. A perturbation analysis of the generalized Sylvester equation $(AR - LB, DR - LE) = (C, F)$. *SIAM J. Matrix Anal. Appl.*, 15(4):1045–1060, 1994.

[85] B. Kågström and P. Poromaa. Computing eigenspaces with specified eigenvalues of a regular matrix pencil (A, B) and condition estimation: Theory, algorithms and software. *Numerical Algorithms*, 12:369–407, 1996.

[86] B. Kågström and P. Poromaa. LAPACK-Style algorithms and software for solving the generalized Sylvester equation and estimating the separation between regular matrix pairs. *ACM Trans. Math. Software*, 22(1):78–103, 1996.

[87] B. Kågström and L. Westin. Generalized Schur methods with condition estimators for solving the generalized Sylvester equation. *IEEE Trans. Automat. Control*, 34:745–751, 1989.

[88] T. Kailath. *Linear Systems*. Prentice-Hall Information and System Sciences Series. Prentice Hall, Englewood Cliffs, NJ, 1980.

[89] L.V. Kantorovich and G.P. Akilov. *Functional Analysis*. Pergamon Press, Oxford, 1982.

[90] T. Kato. *Perturbation Theory for Linear Operators*. Springer-Verlag, New York, 1966.

[91] K. Kautsky, N.K. Nichols, and E.K.-W. Chu. Robust pole assignment in singular control systems. *Linear Algebra Appl.*, 121:9–37, 1989.

[92] C.S. Kenney and A.J. Laub. The matrix sign function. *IEEE Trans. Automat. Control*, 40(8):1330–1348, 1995.

[93] H.W. Knobloch and H. Kwakernaak. *Lineare Kontrolltheorie*. Springer-Verlag, Berlin, 1985. [German].

[94] P.V. Kokotović, R.E. O'Malley, and P. Sannuti. Singular perturbations and order reduction in control theory - an overview. *Automatica*, 12(2):123–132, 1976.

[95] M.M. Konstantinov, V. Mehrmann, and P. Petkov. On properties of Sylvester and Lyapunov operators. *Linear Algebra Appl.*, 312:35–71, 2000.

[96] M.M. Konstantinov, P.Hr. Petkov, D.W. Gu, and V. Mehrmann. Sensitivity of general Lyapunov equations. Technical Report 98-15, Depart. of Engineering, Leicester University, Leicester LE1 7RH, UK, 1998.

[97] A. Kumar and P. Daoutidis. *Control of Nonlinear Differential Algebraic Equation Systems. With Applications to Chemical Processes*. Chapman & Hall/CRC Research Notes in Mathematics, 397. Chapman & Hall/CRC, Boca Raton, FL, 1999.

[98] P. Lancaster and L. Rodman. *The Algebraic Riccati Equation*. Oxford University Press, Oxford, 1995.

[99] P. Lancaster and M. Tismenetsky. *The Theory of Matrices*. Academic Press, Orlando, FL, 2nd edition, 1985.

[100] V.B. Larin and F.A. Aliev. Construction of square root factor for solution of the Lyapunov matrix equation. *Systems Control Lett.*, 20(2):109–112, 1993.

[101] V.B. Larin and F.A. Aliev. Generalized Lyapunov equation and factorization of matrix polynomials. *Systems Control Lett.*, 21(6):485–491, 1993.

[102] A.J. Laub, M.T. Heath, C.C. Paige, and R.C. Ward. Computation of system balancing transformations and other applications of simultaneous diagonalization algorithms. *IEEE Trans. Automat. Control*, AC-32(2):115–122, 1987.

[103] F.L. Lewis. Fundamental, reachability, and observability matrices for discrete descriptor systems. *IEEE Trans. Automat. Control*, AC-30:502–505, 1985.

[104] F.L. Lewis. A survey of linear singular systems. *Circuits Systems Signal Process.*, 5(1):3–36, 1986.

[105] F.L. Lewis. A tutorial on the geometric analysis of linear time-invariant implicit systems. *Automatica*, 28:119–137, 1992.

[106] J.-R. Li, F. Wang, and J. White. An efficient Lyapunov equation-based approach for generating reduced-order models of interconnect. In *Proceedings of the 36th Design Automation Conference (New Orleans, USA, 1999)*, pages 1–6. IEEE, 1999.

[107] Y. Liu and B.D.O. Anderson. Singular perturbation approximation of balanced systems. *Internat. J. Control*, 50:1379–1405, 1989.

[108] R. Loewy. An inertia theorem for Lyapunov's equation and the dimension of a controllability space. *Linear Algebra Appl.*, 260:1–7, 1997.

[109] A. Lu and E. Wachspress. Solution of Lyapunov equations by alternating direction implicit iteration. *Comput. Math. Appl.*, 21(9):43–58, 1991.

[110] D.G. Luenberger. Dynamic equations in descriptor form. *IEEE Trans. Automat. Control*, AC-22(3):312–321, 1977.

[111] A.M. Lyapunov. The general problem of the stability of motion. *Internat. J. Control*, 55(3):521–790, 1992. (Translated by A. T. Fuller from Édouard Davaux's French translation (1907) of the 1892 Russian original).

[112] A. Malyshev. Computing invariant subspaces of a regular linear pencil of matrices. *Siberian Math. J.*, 30(4):559–567, 1990.

[113] A. Malyshev. Parallel algorithm for solving some spectral problems of linear algebra. *Linear Algebra Appl.*, 188/189:489–520, 1993.

[114] R. März. Practical Lyapunov stability criteria for differential algebraic equations. In *Numerical analysis and mathematical modelling*, Banach Cent. Publ., 29, pages 245–266. Polish Academy of Sciences, Inst. of Mathematics, Warszawa, 1994.

[115] R. März. Criteria for the trivial solution of differential algebraic equations with small nonlinearities to be asymptotically stable. *J. Math. Anal. Appl.*, 225:587–607, 1998.

[116] A.G. Mazko. Distribution of the spectrum of a regular matrix pencil with respect to plane curves. *Ukr. Math. Zh.*, 38(1):116–120, 1986.

[117] V. Mehrmann. *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution.* Lecture Notes in Control and Information Sciences, 163. Springer-Verlag, Heidelberg, 1991.

[118] V. Mehrmann. A step toward a unified treatment of continuous and discrete time control problems. *Linear Algebra Appl.*, 241-243:749–779, 1996.

[119] B.C. Moore. Principal component analysis in linear systems: controllability, observability, and model reduction. *IEEE Trans. Automat. Control*, AC-26(1):17–32, 1981.

[120] P.C. Müller. Stability of linear mechanical systems with holonomic constraints. *Appl. Mech. Rev.*, 46(11):160–164, 1993.

[121] G. Obinata and B.D.O. Anderson. *Model Reduction for Control System Design.* Springer-Verlag, London, 2001.

[122] A. Ostrowski and H. Schneider. Some theorems on the inertia of general matrices. *J. Math. Anal. Appl.*, 4(1):72–84, 1962.

[123] D.H. Owens and D.L. Debeljkovic. Consistency and Liapunov stability of linear descriptor systems: A geometric analysis. *IMA J. Numer. Anal.*, 2:139–151, 1985.

[124] D.W. Peaceman and H.H. Rachford. The numerical solution of parabolic and elliptic differential equations. *J. Soc. Indust. Appl. Math.*, 3:28–41, 1955.

[125] T. Penzl. Numerical solution of generalized Lyapunov equations. *Adv. Comput. Math.*, 8(1-2):33–48, 1998.

[126] T. Penzl. *Numerische Lösung großer Lyapunov-Gleichungen.* Logos Verlag, Berlin, 1998. [German].

[127] T. Penzl. A cyclic low-rank Smith method for large sparse Lyapunov equations. *SIAM J. Sci. Comput.*, 21(4):1401–1418, 1999/00.

[128] T. Penzl. Eigenvalue decay bounds for solutions of Lyapunov equations: the symmetric case. *Systems Control Lett.*, 40(2):139–144, 2000.

[129] T. Penzl. Algorithms for model reduction of large dynamical systems. Preprint SFB393/99-40, Fakultät für Mathematik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany, December 1999. Available from `http://www.tu-chemnitz.de/sfb393/sfb99pr.html`.

[130] L. Qiu, B. Bernhardsson, A. Rantzer, E.J. Davison, P.M. Young, and J.C. Doyle. A formula for computation of the real stability radius. *Automatica*, 31(6):879–890, 1995.

[131] L. Qiu and E.J. Davison. The stability robustness of generalized eigenvalues. *IEEE Trans. Automat. Control*, 37(6):886–891, 1992.

[132] P.J. Rabier and W.C. Rheinboldt. *Nonholonomic Motion of Rigid Mechanical Systems from a DAE viewpoint.* SIAM, Philadelphia, PA, 2000.

[133] J.D. Roberts. Linear model reduction and solution of the algebraic Riccati equation by use of the sign function. *Internat. J. Control*, 32(4):677–687, 1980. (Reprint of Technical Report TR-13, CUED/B-Control, Engineering Deparment, Cambrige University, 1971).

[134] H.H. Rosenbrock. Structural properties of linear dynamical systems. *Internat. J. Control*, 20(2):191–202, 1974.

[135] W. Rudin. *Real and Complex Analysis.* McGraw-Hill, New York, 1987.

[136] Y. Saad. Numerical solution of large Lyapunov equations. In M.A. Kaashoek, J.H. Van Schuppen, and A.C.M. Ran, editors, *Signal Processing, Scattering, Operator Theory, and Numerical Methods (Amsterdam, 1989)*, pages 503–511. Birkhäuser, Boston, MA, 1990.

[137] M.G. Safonov and R.Y. Chiang. A Schur method for balanced-truncation model reduction. *IEEE Trans. Automat. Control*, AC-34(7):729–733, 1989.

[138] B. Simeon, F. Grupp, C Führer, and P. Rentrop. A nonlinear truck model and its treatment as a multibody system. *J. Comput. Appl. Math.*, 50:523–532, 1994.

[139] R.A. Smith. Matrix equation $XA + BX = C$. *SIAM J. Appl. Math.*, 16:198–201, 1968.

[140] J. Snyders and M. Zakai. On nonnegative solutions of the equation $AD + DA' = -C$. *SIAM J. Appl. Math.*, 18:704–714, 1970.

[141] G. Starke. *Rationale Minimierungsprobleme in der komplexen Ebene im Zusammenhang mit der Bestimmung optimaler ADI-Parameter.* Ph. thesis, Fakultät für Mathematik, Universität Karlsruhe, 1989. [German].

[142] G. Starke. Optimal alternating direction implicit parameters for nonsymmetric systems of linear equations. *SIAM J. Numer. Anal.*, 28(5):1431–1445, 1991.

[143] G.W. Stewart. Error and perturbation bounds for subspaces associated with certain eigenvalue problems. *SIAM Rev.*, 15:727–764, 1973.

[144] G.W. Stewart. *Matrix Algorithms. Vol. 1: Basic Decompositions.* SIAM, Philadelphia, PA, 1998.

[145] G.W. Stewart and J.-G. Sun. *Matrix Perturbation Theory.* Academic Press, New York, 1990.

[146] T. Stykel. Numerical solution and perturbation theory for generalized Lyapunov equations. *Linear Algebra Appl.*, 349:155–185, 2002.

[147] T. Stykel. On criteria for asymptotic stability of differential-algebraic equations. *Z. Angew. Math. Mech.*, 82(3):147–158, 2002.

[148] T. Stykel. Stability and inertia theorems for generalized Lyapunov equations. *Linear Algebra Appl.*, 355:297–314, 2002.

[149] T. Stykel. On a criterion for asymptotic stability of differential-algebraic equations. Preprint SFB393/99-17, Fakultät für Mathematik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany, August 1999. Available from `http://www.tu-chemnitz.de/sfb393/sfb99pr.html`.

[150] T. Stykel. Model reduction of descriptor systems. Technical Report 720-2001, Institut für Mathematik, Technische Universität Berlin, D-10263 Berlin, Germany, December 2001. Submitted for publication.

[151] T. Stykel. Generalized Lyapunov equations for descriptor systems: stability and inertia theorems. Preprint SFB393/00-38, Fakultät für Mathematik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany, October 2000. Available from `http://www.tu-chemnitz.de/sfb393/sfb00pr.html`.

[152] X. Sun and E.S. Quintana-Ortí. The generalized Newton iteration for the matrix sign function. To appear in *SIAM J. Sci. Comput.*

[153] V.L. Syrmos, P. Misra, and R. Aripirala. On the discrete generalized Lyapunov equation. *Automatica*, 31(2):297–301, 1995.

[154] K. Takaba, N. Morihira, and T. Katayama. A generalized Lyapunov theorem for descriptor system. *Systems Control Lett.*, 24:49–51, 1995.

[155] C. Tischendorf. On the stability of solutions of autonomous index-1 tractable and quasilinear index-2 tractable DAEs. *Circuits Systems Signal Process.*, 13(2-3):139–154, 1995.

[156] M.S. Tombs and I. Postlethweite. Truncated balanced realization of a stable non-minimal state-space system. *Internat. J. Control*, 46(4):1319–1330, 1987.

[157] L. N. Trefethen. Pseudospectra of matrices. In D. F. Griffiths and G. A. Watson, editors, *Numerical Analysis 1991 (Dundee, 1991)*, pages 234–266. Longman Sci. Tech., Harlow, Essex, UK, 1992.

[158] L.N. Trefethen. Computation of pseudospectra. *Acta Numerica*, 8:247–295, 1999.

[159] L.N. Trefethen. Spectra and pseudospectra. In M. Ainworth, J. Levesley, and M. Marletta, editors, *The Graduate Student's Guide to Numerical Analysis*, pages 217–250. Springer-Verlag, Berlin, 1999.

[160] L.N. Trefethen, A.E. Trefethen, S.C. Reddy, and T.A. Driscoll. Hydrodynamic stability without eigenvalues. *Science*, 261:578–584, 1993.

[161] P. Van Dooren. The computation of the Kronecker's canonical form of a singular pencil. *Linear Algebra Appl.*, 27:103–140, 1979.

[162] J.L.M. van Dorsselaer. Pseudospectra for matrix pencils and stability of equilibria. *BIT*, 37(4):833–845, 1997.

[163] C.F. Van Loan. How near is a stable matrix to an unstable matrix? *Contemp. Math.*, 47:465–478, 1985.

[164] A. Varga. Efficient minimal realization procedure based on balancing. In A. EL Moudni, P. Borne, and S.G. Tzafestas, editors, *Proc. of IMACS/IFAC Symposium on Modelling and Control of Technological Systems (Lille, France, May 7-10, 1991)*, volume 2, pages 42–47, 1991.

[165] A. Varga. *Model Reduction Routines for SLICOT*. NICONET Report 1999-8, 1999. Available from `ftp://wgs.esat.kuleuven.ac.be/pub/WGS/REPORTS/nic1999-8.ps.Z`.

[166] G.C. Verghese, B.C. Lévy, and T. Kailath. A generalized state-space for singular systems. *IEEE Trans. Automat. Control*, AC-26(4):811–831, 1981.

[167] E. Wachspress. The ADI minimax problem for complex spectra. *Appl. Math. Lett.*, 1:311–314, 1988.

[168] E. Wachspress. The ADI minimax problem for complex spectra. In D. Kincaid and L. Hayes, editors, *Iterative Methods for Large Linear Systems*, pages 251–271. Academic Press, 1990.

[169] D.S. Watkins. Performance of the QZ algorithm in the presence of infinite eigenvalues. *SIAM J. Matrix Anal. Appl.*, 22(2):364–375, 2000.

[170] J. Weickert. *Applications of the Theory of Differential-Algebraic Equations to Partial Differential Equations of Fluid Dynamics.* Ph.D. thesis, Technische Universität Chemnitz, Chemnitz, 1997.

[171] J.H. Wilkinson. *The Algebraic Eigenvalue Problem.* Clarendon Press, Oxford, 1965.

[172] H.K. Wimmer. On the Ostrowski-Schneider inertia theorem. *J. Math. Anal. Appl.*, 41:164–169, 1973.

[173] H.K. Wimmer. Inertia theorems for matrices, controllability and linear vibrations. *Linear Algebra Appl.*, 8:337–343, 1974.

[174] E.L. Yip and R.F. Sincovec. Solvability, controllability and observability of continuous descriptor systems. *IEEE Trans. Automat. Control*, AC-26(3):702–707, 1981.

[175] L. Zhang, J. Lam, and Q. Zhang. Lyapunov and Riccati equations for discrete-time descriptor systems. *IEEE Trans. Automat. Control*, 44(11):2134–2139, 1999.

[176] K. Zhou, J.C. Doyle, and K. Glover. *Robust and Optimal Control.* Prentice Hall, Upper Saddle River, NJ, 1996.

Die vorliegende Dissertation wurde teilweise veröffentlicht in

[1] T. Stykel. Stability and Inertia Theorems for Generalized Lyapunov Equations. To appear in *Linear Algebra Appl.*, 2002.

[2] T. Stykel. Numerical Solution and Perturbation Theory for Generalized Lyapunov Equations. *Linear Algebra Appl.*, 349(1-3), 2002, 155–185.

[3] T. Stykel. On Criteria for Asymptotic Stability of Differential-Algebraic Equations. *Z. Angew. Math. Mech.*, 82(3), 2002, 147–158.

[4] T. Stykel. Model Reduction of Descriptor Systems. Technical Report 720-2001, Institut für Mathematik, Technische Universität Berlin, December 2001.

[5] T. Stykel. Generalized Lyapunov Equations for Descriptor Systems: Stability and Inertia Theorems. Preprint SFB393/00-38, Fakultät für Mathematik, Technische Universität Chemnitz, October 2000.

[6] T. Stykel. On a Criterion of Asymptotic Stability of Differential-Algebraic Equations. Preprint SFB393/99-17, Fakultät für Mathematik, Technische Universität Chemnitz, August 1999.

LEBENSLAUF

Tatjana Stykel

Berlin, 20.03.2002

Geburtsdatum:	19. Juli, 1973
Geburtsort:	KasMIS, Kasachstan
Staatsangehörigkeit:	Deutsch
Familienstand:	ledig

SCHULBILDUNG und STUDIUM

Schulbildung:	Sekundarschule, KasMIS, Kasachstan	Sept. 1980 - Juni 1990
Studium:	Fakultät für Mechanik und Mathematik, Staatliche Universität Novosibirsk, Rußland	Sept. 1990 - Juni 1994
Diplomarbeit:	*Über die Berechnung der verallgemeinerten Drazin-Matrix.*	
	Diplom-Mathematikerin	Juni 1994
Magisterstudium:	Fakultät für Mechanik und Mathematik, Staatliche Universität Novosibirsk, Rußland	Juli 1994 - Juni 1996
Magisterarbeit:	*Randwertprobleme für die gewöhnlichen Differentialgleichungen, die bezüglich der Ableitungen unlösbar sind.*	
Magistergrad:	Magisterin der Mathematik	Juni 1996
Promotionsstudium:		
	Staatliche Universität Novosibirsk, Rußland	Juli 1996 - Juni 1997
	Technische Universität Chemnitz, Deutschland	Okt. 1998 - März 2001

WEITERBILDUNG

Humboldt-Universität zu Berlin	wissenschaftliche Fortbildung	Okt. 1997 - Sept. 1998
Technische Universität Chemnitz	Weiterbildung "Lehren und Lernen an der Hochschule"	März 2000

BERUFSTÄTIGKEIT

Technische Universität Chemnitz	Wissenschaftliche Mitarbeiterin	März 2000 - Okt. 2000
Technische Universität Berlin	Wissenschaftliche Mitarbeiterin	seit Nov. 2000

STIPENDIEN

International Soros Science Education Program, Open Society Institution	Soros Student Stipendium	1994 - 1996
International Soros Science Education Program, Open Society Institution	Soros Graduate Student Stipendium	1996 - 1997
Deutscher Akademischer Austauschdienst	Jahresstipendium	1997 - 1998
Sächsisches Staatsministerium für Wissenschaft und Kunst, Hochschulsonderprogramm III	Promotionsstipendium	1998 - 2000