MASARYK UNIVERSITY
Faculty of Science

Viera RŮŽIČKOVÁ

Discrete Symplectic Systems and Definiteness of Quadratic Functionals

Ph.D. dissertation

Advisor: doc. RNDr. Roman Hilscher, Ph.D. Brno, 2006

Bibliographical identification

Name and surname of the author: Viera Růžičková

Title of dissertation: Discrete symplectic systems and definiteness of quadratic functionals

Program: Mathematical analysis

Advisor: doc. RNDr. Roman Hilscher, Ph.D.

Year of defense: 2006

Keywords: Discrete symplectic system, Quadratic functional, Nonnegativity, Positivity, Conjoined basis, Riccati equation, Riccati inequality

2000 Mathematics Subject Classification: 39A10, 39A12, 49K99, 49N10.

Preface

This work is a complete presentation of results about the definiteness of discrete quadratic functionals related to discrete symplectic systems. The definiteness is characterized via certain solutions (called conjoined bases) of the corresponding symplectic system, and via implicit and explicit Riccati matrix equations and inequalities. The motivation and history of this topic are included in the introductory chapter.

The first chapter is devoted to preliminary results from the matrix theory, in particular to properties of symplectic matrices and properties of Moore-Penrose generalized inverse. Moreover, the discrete symplectic system and related discrete quadratic functionals are introduced there. The second chapter contains definitions of some important matrices and an augmented symplectic system, and several Picone-type identities. These objects are used in the proofs in the third chapter which contains roundabout theorems with equivalent conditions for the positivity and nonnegativity of discrete quadratic functionals. At the end of each chapter there is a section with notes on the literature.

I would like to thank my advisor Roman Hilscher from whom I learned most of what I know about the problems contained in this work and who helped me to improve it considerably. I am also grateful to other people from the Department of Mathematical Analysis and the Section of Mathematics, especially to Ondřej Došlý for his continuous support throughout my study in Brno. Further I wish to express my thanks to Werner Kratz and the Department of Applied Analysis at the University of Ulm for a friendly welcome and for the opportunity to speak about my work on their seminar during my stay in Ulm.

Finally, a special thank you is due to my parents, sister, grandparents, and other relatives, and to my friends for their support and everything.

Brno, May 2006 Viera Růžičková

Contents

Chapter 0

Introduction

This introductory chapter should respectively give the reader answers on questions like: what is this work about and why are these things being researched, who all has been involved in it, how are the presented results proven, and what did the author herself.

0.1 Introduction and motivation

In the discrete calculus of variations and control problems, a quadratic functional

$$\mathcal{F}_0(x,u) := \sum_{k=0}^{N} \left\{ x_k^T \mathcal{A}_k^T \mathcal{C}_k x_k + 2\, x_k^T \mathcal{C}_k^T \mathcal{B}_k u_k + u_k^T \mathcal{B}_k^T \mathcal{D}_k u_k \right\}$$

arises as second variation [40]. It is useful to know whether it is nonnegativite or positive or not, because its nonnegativity is a necessary optimality condition, while its positivity is a sufficient optimality condition.

With the functional $\mathcal{F}_0$ we associate a linear system, called the *discrete symplectic system*,

$$\begin{pmatrix} x_{k+1} \\ u_{k+1} \end{pmatrix} = \begin{pmatrix} \mathcal{A}_k & \mathcal{B}_k \\ \mathcal{C}_k & \mathcal{D}_k \end{pmatrix} \begin{pmatrix} x_k \\ u_k \end{pmatrix} \tag{S}$$

whose name is derived from the fact that its transition matrix is symplectic. Its first equation, $x_{k+1} = \mathcal{A}_k x_k + \mathcal{B}_k u_k$, is called the *equation of motion* and the pairs of vector sequences $\big(\{x_0, x_1, \ldots, x_{N+1}\}, \{u_0, u_1, \ldots, u_{N+1})\}\big)$, (usually denoted by (x,u)) that solve this equation on the discrete interval

$[0, N]$ are called *admissible*. We are interested in definiteness of $\mathcal{F}_0$ on such admissible pairs.

Discrete symplectic systems cover as a special case *discrete Hamiltonian*[1] *system*

$$\Delta \begin{pmatrix} x_k \\ u_k \end{pmatrix} = \begin{pmatrix} A_k & B_k \\ C_k & -A_k^T \end{pmatrix} \begin{pmatrix} x_{k+1} \\ u_k \end{pmatrix}, \tag{H}$$

where in system (S) we have $\mathcal{A}_k := (I - A_k)^{-1}$, $\mathcal{B}_k := (I - A_k)^{-1} B_k$, $\mathcal{C}_k := -C_k (I - A_k)^{-1}$, and $\mathcal{D}_k = C_k (I - A_k)^{-1} B_k - A_k^T + I$.

Further, Sturm[2]-Liouville[3] difference equation

$$\sum_{\nu=0}^{n} (-1)^\nu \Delta^\nu (r_k^{(\nu)} \Delta^\nu y_{k+n-\nu}) = 0$$

is equivalent to the discrete Hamiltonian system with the transition matrix

$$\begin{pmatrix} A_k & B_k \\ C_k & -A_k^T \end{pmatrix} = \left(\begin{array}{cccc|cccc} 0 & 1 & \dots & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & \frac{1}{r_k^{(n)}} \\ \hline r_k^{(0)} & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & r_k^{(1)} & \dots & 0 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & r_k^{(n-1)} & 0 & \dots & -1 & 0 \end{array}\right).$$

The definiteness of $\mathcal{F}$ is investigated not on all possible pairs of admissible (x, u), but on a subset of them, defined by some additional boundary conditions on x_0 and x_{N+1}. The first type of these conditions, connected with $\mathcal{F}_0$, is $x_0 = 0 = x_{N+1}$, and the functional is then called the *functional with zero endpoints*. Then there are more general types of functionals. One we get, when we add to the sum in $\mathcal{F}_0$ two more quadratic terms, $x_0^T \Gamma_0 x_0$ and $x_{N+1}^T \Gamma_1 x_{N+1}$, and the boundary conditions we define as $\mathcal{M}_0 x_0 = 0 = \mathcal{M}_1 x_{N+1}$. This is called the *functional with separated endpoints*. When we add to the sum

[1] Sir William Rowan Hamilton (∗ August 4, 1805, † September 2, 1865) was an Irish mathematician, physicist, and astronomer.

[2] Jacques Charles François Sturm (∗ September 29, 1803, † December 15, 1855) was a French-Swiss mathematician.

[3] Joseph Liouville (∗ March 24, 1809, † September 8, 1882) was a French mathematician.

in $\mathcal{F}_0$ just one quadratic term, but of double dimension, $\left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right)^T \Gamma \left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right)$, and the boundary conditions we define as $\mathcal{M} \left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right) = 0$, we get another more general type of functional. It is called as the *functional with general endpoints*. The latter one is used for example for problems with periodic endpoints $x_0 = x_{N+1}$.

There are various conditions equivalent to the positivity or nonnegativity for each type of the functional. They are usually collected together in one theorem, which is called a *Reid roundabout theorem.* (See footnote 5 on page 4.) In this work we present together six Reid roundabout theorems, with conditions for the positivity and for the nonnnegativity for all three types of functionals.

Exempli gratia, we show here a characterization of the positivity and nonnegativity of the discrete quadratic functionals via

- the *principal solution* of (S) (for the functional with zero and general endpoints) and via the *natural conjoined basis* of (S) (for the functional with separated endpoints), where the principal solution and the natural conjoined basis of (S) are the matrix solutions of (S) starting with the initial values $(0, I)$ and $(I - \mathcal{M}_0, \Gamma_0 + \mathcal{M}_0)$ respectively,

- *implicit Riccati*[4] *equations*, involving the Riccati operator $R[Q]_k := Q_{k+1}(\mathcal{A}_k + \mathcal{B}_k Q_k) - (\mathcal{C}_k + \mathcal{D}_k Q_k)$ and some other matrices,

- the *explicit Riccati equation* $R[Q]_k = 0$ (only for the positivity, for all three types of the functional),

- the *Riccati inequality* (only for the positivity, for the functional with zero and separated endpoints),

- the positivity and nonnegativity of certain *perturbed functionals*, e.g. of the functional $\mathcal{F}_0(x, u) + \alpha \|x_0\|^2 + \beta \|x_{N+1}\|^2$.

0.2 History and literature

In 1992, L. Erbe and P. Yan introduced linear Hamiltonian difference systems of the form (H) in [27]. The case when B is nonsingular was examined

[4] Jacopo Francesco Riccati ($*$ May 28, 1676, $\dagger$ April 15, 1754) was an Italian mathematician.

first, by C. Ahlbrandt, S. Chen, O. Došlý, L. Erbe, M. Heifetz, J. Hooker, T. Peil, A. Peterson, J. Ridenhour and P. Yan, see [1–3, 21, 28–30, 49–51]. The term "Reid roundabout theorem" for a theorem which gives equivalent conditions for disconjugacy of discrete Hamiltonian system, was first used by C. Ahlbrandt in honour of W. T. Reid [5], who studied this theory in the continuous case, e.g. in [53–55]. The continous case was studied also by W. A. Coppel, e.g. in [20] and W. Kratz, e.g. in [45] and lately in [47].

In 1996, M. Bohner proved in [9] a Reid roundabout theorem for the case when B_k is allowed to be singular. Later it was extended to functionals with general boundary conditions by M. Bohner in [10, 12], by M. Bohner, O. Došlý and W. Kratz in [17] and by R. Hilscher and V. Zeidan in [41, 42]. It was proven with the use of an augmented symplectic system in dimension $2n$, which was already known from the continuous case.

Meanwhile, in 1996, C. Ahlbrandt and A. Peterson showed in [4] that discrete Hamiltonian systems are a special case of discrete symplectic systems and, in 1997, M. Bohner and O. Došlý presented in [13] a Reid roundabout theorem for discrete symplectic systems which gives equivalent conditions for the positivity of discrete quadratic functional $\mathcal{F}$ with zero endpoints. M. Bohner later generalized some of these results to variable endpoints in [11]. The discrete Picone [6] identity was used in the proofs in both cases. Another possible approach is by diagonalizing the matrix representation of $\mathcal{F}$, which was used in [14] by M. Bohner and O. Došlý for the Hamiltonian system (H), and in [35] by R. Hilscher for discrete symplectic system (S). This theory for symplectic systems (positivity) was then completed in 2003 by R. Hilscher and V. Zeidan in [40], where it is also shown that symplectic system (S) is the Euler [7]-Lagrange [8] (or Jacobi [9]) system for the given discrete quadratic functional.

A characterization of the nonnegativity of $\mathcal{F}$ with zero endpoints was derived in 2003 by O. Došlý, R. Hilscher and V. Zeidan in [24] and by

[5] William Thomas Reid ($*$ October 4, 1907, † October 14, 1977) was an American mathematician.

[6] Mauro Picone ($*$ May 2, 1885, † April 11, 1977) was an Italian mathematician.

[7] Leonhard Euler ($*$ April 15, 1707, † September 18, 1783) was a Swiss mathematician and physicist.

[8] Joseph Louis Lagrange ($*$ January 25, 1736, † April 10, 1813) was an Italian-French mathematician and astronomer.

[9] Carl Gustav Jakob Jacobi ($*$ December 10, 1804, † February 18, 1851) was a German mathematician.

M. Bohner, O. Došlý and W. Kratz in [18], and generalized to separated endpoints by M. Bohner, O. Došlý, R. Hilscher and W. Kratz in [16], in the latter by the diagonalization approach.

0.3 Related topics

Other topics on discrete symplectic systems in the current literature include trigonometric systems [5], discrete Prüfer[10] (trigonometric) transformation [15], discrete hyperbolic systems and discrete hyperbolic transformation [26], theory of generalized zeros [13], conjugate intervals [40], coupled intervals [43], Sturmian comparison results [25], and discrete eigenvalue problems [19]. In [56] it is shown that discrete Hamiltonian systems also have a symplectic structure.

Let us further mention that there also exist *variable stepsize symplectic difference systems*,

$$\frac{\Delta \begin{pmatrix} x_k \\ u_k \end{pmatrix}}{\mu_k} = \begin{pmatrix} \mathcal{A}_k & \mathcal{B}_k \\ \mathcal{C}_k & \mathcal{D}_k \end{pmatrix} \begin{pmatrix} x_k \\ u_k \end{pmatrix},$$

see e.g. [24, 40] which can be directly reduced to the system (S), and *time scale symplectic systems*,

$$\begin{pmatrix} x \\ u \end{pmatrix}^{\Delta} = \begin{pmatrix} \mathcal{A}(t) & \mathcal{B}(t) \\ \mathcal{C}(t) & \mathcal{D}(t) \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix},$$

see e.g. [22, 23, 36]. In these two cases, the matrix $I + \mu \left(\begin{smallmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{smallmatrix}\right)$ is now symplectic.

Also, it is demonstrated in [31–33, 57] that symplectic difference schemes are the best way for solving Hamiltonian systems numerically.

0.4 Methods of proofs

In the substantial proofs presented in this work we have to show equivalences of certain statements. These are about the definiteness of a discrete quadratic functional, and about existence and properites of certain matrix solutions of

[10] Ernst Paul Heinz Prüfer ($*$ November 10, 1896, † April 7, 1934) was a German mathematician.

the associated discrete symplectic system (S) or a discrete Riccati equation or inequality.

As all the problems are finite dimensional, most of the work is "playing with matrices", and using the properties of generalized inverses, symplectic matrices, and projections.

Furthermore, we often use a discrete Picone-type identity to write a quadratic functional in the form of a square and show that it is nonnegative, and thus prove the sufficiency of a certain condition for the positivity or nonnegativity. The necessity is proven by finding a pair on which the value of the functional is zero or negative, assuming the condition does not hold.

In the proof of the roundabout theorem for the positivity of the fuctional with separated endpoints we use a transformation to a functional with zero endpoints, i.e. we add one zero element in front of the first one and one zero element after the last one. The functional with general endpoints can be transformed into an augmented functional in double dimension with separated endpoints.

0.5 List of author's results

Author's own results are (in order as they appear in the text):

- ⋄ A new form of the Riccati quotient Q^* for an augmented symplectic system in dimension $2n$ and its relation to the the Riccati quotient Q for symplectic system in dimension n. (Lemmas 2.17, 2.19, page 31.)
- ⋄ Identities about the relation between the value of a functional $\mathcal{F}$ on a pair (x, u) to the value of the same functional $\mathcal{F}$ on another pair $(\bar{x}, \bar{u})$ which satisfies given boundary conditions. (Theorem 2.31 with Corollaries 2.33, 2.34, and Theorem 2.38 with Corollary 2.39, pages 36–39.)
- ⋄ A characterization of the positivity of the quadratic functional with zero endpoints and with separated endpoints via the explicit Riccati inequality. (Statements (vi) and (vii) in Theorems 3.4, 3.14, pages 42, 50. These results are published in [37] and were obtained jointly with R. Hilscher.)
- ⋄ A charecterization of the nonnegativity of the quadratic functional with zero endpoints and with separated endpoints via the implicit Riccati

equation. (Statement (iii) in Theorems 3.42, 3.43, pages 69, 70. These results are contained in [39].)

- A characterization of the nonnegativity of the quadratic functional with general endpoints via the principal solution of the corresponding symplectic system. (Statement (ii) in Theorem 3.49, page 76. These results are published in [37] and were obtained jointly with R. Hilscher.)

- A characterization of the positivity and of the nonnegativity of the quadratic functional with general endpoints via the implicit Riccati equation in terms of the nonaugmented Riccati operator. (Some parts of Theorems 3.54, 3.55, page 81.)

- A characterization of the positivity and of the nonnegativity of all three types of quadratic functionals via the positivity and the nonnegativity of a perturbed quadratic functional. (Some parts of Theorems 3.58, 3.59, 3.64, 3.65, 3.70, 3.71, pages 84–92.)

Chapter 1

Preliminaries

1.1 Notation and definitions

For any matrix $A \in \mathbb{R}^{m\times n}$, by A^T we denote the transpose of A. By $\operatorname{Im} A$, $\operatorname{Ker} A$ we denote respectively the image and the kernel of A, i.e. $\operatorname{Im} A = \{v \in \mathbb{R}^m : v = Ac \text{ for some } c \in \mathbb{R}^n\}$, $\operatorname{Ker} A = \{c \in \mathbb{R}^n : Ac = 0\}$. By $\operatorname{rank} A$ we denote the dimension of $\operatorname{Im} A$. By A^* we *do not* denote the conjugate transpose, but it is in this work a notation for certain matrix (matrices) in $\mathbb{R}^{2n\times 2n}$ arising from a matrix A in $\mathbb{R}^{n\times n}$.

For integers a, b we denote the discrete interval $\{a, a+1, \dots, b\}$ by $[a, b]$. In particular, we will use the intervals $[0, N]$ and $[0, N+1]$.

Further we denote $f_k\big|_0^{N+1} := f_{N+1} - f_0$, for a sequence $\{f_k\}_{k=0}^{N+1}$.

1.2 Matrices and matrix properties

In this section we present various properties of matrices that will be further used when studying discrete symplectic systems. They are stated for real matrices, but all hold for matrices with complex elements as well, when the transpose of a matrix is replaced by the conjugate transpose.

For reader's covenience, some of the proofs are included although all of them can be found in the quoted literature.

1.2.1 Moore-Penrose generalized inverse

For every real matrix A there is a unique matrix B satisfying the four equations

$$ABA = A, \quad BAB = B, \quad (AB)^T = AB, \quad (BA)^T = BA. \tag{1.1}$$

This unique matrix B is known as the *Moore*[1]*-Penrose*[2] *inverse* and we denote it by $A^\dagger$. Some of its properties are

$$(A^\dagger)^\dagger = A, \quad (A^T)^\dagger = (A^\dagger)^T, \quad \operatorname{Ker} A^\dagger = \operatorname{Ker} A^T.$$

Remark 1.1. Since the matrix operations of the transpose and the Moore-Penrose inverse are commutative, we denote by $A^{\dagger T}$ when both are applied to a matrix A.

Full-rank factorization. For every real matrix $A \in \mathbb{R}^{m\times n}$ with $\operatorname{rank} A = r > 0$ there exist matrices $F \in \mathbb{R}^{m\times r}, R \in \mathbb{R}^{r\times n}$ with $\operatorname{rank} F = \operatorname{rank} R = r$, such that $A = FR$ and this formula is called a *full rank factorization* of A.

Explicit formula for generalized inverse. If $A \in \mathbb{R}^{m\times n}$ with $\operatorname{rank} A = r > 0$ has a full-rank factorization $A = FR$, then

$$A^\dagger = R^T(RR^T)^{-1}(F^TF)^{-1}F^T. \tag{1.2}$$

This further implies

$$A^\dagger A = R^T(RR^T)^{-1}R \quad \text{and} \quad AA^\dagger = F(F^TF)^{-1}F^T. \tag{1.3}$$

The Moore-Penrose generalized inverse is a useful tool for describing the relations between image and kernel of a matrix. The following conditions hold

$$\operatorname{Ker} A = \operatorname{Im}(I - A^\dagger A), \quad \operatorname{Im} A = \operatorname{Ker}(I - AA^\dagger), \tag{1.4}$$

$$\operatorname{Ker} V \subseteq \operatorname{Ker} W \quad \Leftrightarrow \quad W = WV^\dagger V \quad \Leftrightarrow \quad W^\dagger = V^\dagger V W^\dagger, \tag{1.5}$$

$$\operatorname{Im} V \subseteq \operatorname{Im} W \quad \Leftrightarrow \quad V = WW^\dagger V \quad \Leftrightarrow \quad V^\dagger = V^\dagger W W^\dagger. \tag{1.6}$$

Lemma 1.2. *Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. The equation $Ax = b$ has a solution if and only if $AA^\dagger b = b$, and then all solutions have the form $x = A^\dagger b + (I - A^\dagger A)\gamma$ for some $\gamma \in \mathbb{R}^n$. Consequently, if there exists at least one solution x of $Ax = b$, then $x = A^\dagger b$ is one of the solutions.*

Proof. It follows from equivalences (1.4). □

[1] Eliakim Hastings Moore ($*$ January 26, 1862, † December 30, 1932) was an American mathematician.

[2] Sir Roger Penrose ($*$ August 8, 1931) is an English mathematical physicist.

1.2.2 Projection

A matrix A is called a *projection*, if A is symmetric and $AA = A$. Some of its properties are

$$A = A^\dagger, \quad \operatorname{Ker} A = \operatorname{Im}(I - A), \quad \operatorname{Im} A = \operatorname{Ker}(I - A).$$

1.2.3 Other matrix properties

Lemma 1.3. *For any real matrix A the following identity*

$$(I + AA^T)^{-1} = I - A(I + A^T A)^{-1} A^T \tag{1.7}$$

holds.

Lemma 1.4. *If $A \in \mathbb{R}^{n\times n}$ is a real symmetric matrix with the smallest eigenvalue $\lambda_{\min}$ and the largest eigenvalue $\lambda_{\max}$, then for any vector $v \in \mathbb{R}^n$ we have*

$$\lambda_{\min} \|v\|^2 \leq v^T A v \leq \lambda_{\max} \|v\|^2. \tag{1.8}$$

1.2.4 Symplectic matrices

Let $n \in \mathbb{N}$ and $\mathcal{J}$ be a real $2n \times 2n$ matrix, $\mathcal{J} = \left(\begin{smallmatrix} 0 & I \\ -I & 0 \end{smallmatrix}\right)$. Some properties of $\mathcal{J}$ are

$$\mathcal{J}^{-1} = \mathcal{J}^T = -\mathcal{J}, \quad \mathcal{J}^2 = -I, \quad \det \mathcal{J} = 1. \tag{1.9}$$

Definition 1.5. A real $2n \times 2n$ matrix $\mathcal{S}$ is called *symplectic* if $\mathcal{S}^T \mathcal{J} \mathcal{S} = \mathcal{J}$.

The simplest examples of symplectic matrices are the $2n \times 2n$ matrices $\mathcal{J}$ and I. In general, symplectic matrices can be characterized by the following.

Lemma 1.6. *If $\mathcal{S}$ has $n \times n$ block entries $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$, i.e. if $\mathcal{S} = \left(\begin{smallmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{smallmatrix}\right)$, then $\mathcal{S}$ is symplectic if and only if one of the following equivalent conditions is satisfied*

$$\mathcal{A}^T\mathcal{C} = \mathcal{C}^T\mathcal{A}, \mathcal{B}^T\mathcal{D} = \mathcal{D}^T\mathcal{B}, \mathcal{A}^T\mathcal{D} - \mathcal{C}^T\mathcal{B} = I, \tag{1.10}$$

$$\mathcal{S}^{-1} = \begin{pmatrix} \mathcal{D}^T & -\mathcal{B}^T \\ -\mathcal{C}^T & \mathcal{A}^T \end{pmatrix}, \tag{1.11}$$

$$\mathcal{S}^{-1} \textit{ is symplectic,} \tag{1.12}$$

$$\mathcal{S}^T = \begin{pmatrix} \mathcal{A}^T & \mathcal{C}^T \\ \mathcal{B}^T & \mathcal{D}^T \end{pmatrix} \textit{ is symplectic,} \tag{1.13}$$

$$\mathcal{D}\mathcal{C}^T = \mathcal{C}\mathcal{D}^T, \mathcal{A}\mathcal{B}^T = \mathcal{B}\mathcal{A}^T, \mathcal{D}\mathcal{A}^T - \mathcal{C}\mathcal{B}^T = I. \tag{1.14}$$

Proof. We can see it from the following calculations

$$\mathcal{S}^T\mathcal{J}\mathcal{S} = \begin{pmatrix} -\mathcal{C}^T\mathcal{A}+\mathcal{A}^T\mathcal{C} & \mathcal{A}^T\mathcal{D}-\mathcal{C}^T\mathcal{B} \\ \mathcal{B}^T\mathcal{C}-\mathcal{D}^T\mathcal{A} & -\mathcal{D}^T\mathcal{B}+\mathcal{B}^T\mathcal{D} \end{pmatrix} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} = \mathcal{J},$$

$$\mathcal{S}^{-1}\mathcal{S} = \begin{pmatrix} \mathcal{D}^T\mathcal{A}-\mathcal{B}^T\mathcal{C} & \mathcal{D}^T\mathcal{B}-\mathcal{B}^T\mathcal{D} \\ -\mathcal{C}^T\mathcal{A}+\mathcal{A}^T\mathcal{C} & -\mathcal{C}^T\mathcal{B}+\mathcal{A}^T\mathcal{D} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix},$$

$$\mathcal{S}\mathcal{S}^{-1} = \begin{pmatrix} \mathcal{A}\mathcal{D}^T-\mathcal{B}\mathcal{C}^T & -\mathcal{A}\mathcal{B}^T+\mathcal{B}\mathcal{A}^T \\ \mathcal{C}\mathcal{D}^T-\mathcal{D}\mathcal{C}^T & -\mathcal{C}\mathcal{B}^T+\mathcal{D}\mathcal{A}^T \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix},$$

$$\mathcal{S}\mathcal{J}\mathcal{S}^T = \begin{pmatrix} -\mathcal{B}\mathcal{A}^T+\mathcal{A}\mathcal{B}^T & -\mathcal{B}\mathcal{C}^T+\mathcal{A}\mathcal{D}^T \\ -\mathcal{D}\mathcal{A}^T+\mathcal{B}\mathcal{B}^T & -\mathcal{D}\mathcal{C}^T+\mathcal{C}\mathcal{D}^T \end{pmatrix} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} = \mathcal{J},$$

$$\mathcal{S}^{-1T}\mathcal{J}\mathcal{S}^{-1} = \begin{pmatrix} \mathcal{C}\mathcal{D}^T-\mathcal{D}\mathcal{C}^T & -\mathcal{C}\mathcal{B}^T+\mathcal{D}\mathcal{A}^T \\ -\mathcal{A}\mathcal{D}^T+\mathcal{B}\mathcal{C}^T & \mathcal{A}\mathcal{B}^T-\mathcal{B}\mathcal{A}^T \end{pmatrix} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} = \mathcal{J}.$$

□

Next two lemmas describe some properties of the eigenvalues and the determinant of a symplectic matrix.

Lemma 1.7. *If $\lambda \in \mathbb{C}$ is an eigenvalue of a symplectic matrix $\mathcal{S}$, then $\frac{1}{\lambda}$ is also an eigenvalue of $\mathcal{S}$. Consequently, if $\lambda = 1$ or $\lambda = -1$ is an eigenvalue of $\mathcal{S}$, then its multiplicity is even.*

Proof. From Definition 1.5 we get $\mathcal{J}^{-1}\mathcal{S}^T\mathcal{J} = \mathcal{S}^{-1}$ i.e. $\mathcal{S}^T$ and $\mathcal{S}^{-1}$ are similar and thus have the same spectrum. So if λ is an eigenvalue of $\mathcal{S}$, then it is also an eigenvalue of $\mathcal{S}^{-1}$, and then $\frac{1}{\lambda}$ is an eigenvalue of $\mathcal{S}$. □

From Definition 1.5 and (1.9) we can see that $(\det \mathcal{S})^2 = 1$. The next lemma shows that the determinant of a symplectic matrix is actually 1.

Lemma 1.8. *If $\mathcal{S}$ is a symplectic $2n \times 2n$ matrix, then $\det \mathcal{S} = 1$.*

Proof. Let $\lambda_1, \dots, \lambda_{2n}$ be the eigenvalues of $\mathcal{S}$, including multiplicities. Since the eigenvalues appear in pairs λ and $\frac{1}{\lambda}$ and, since $\lambda = -1$ has an even multiplicity (provided it is an eigenvalue at all), it follows that $\det \mathcal{S} = \lambda_1 \dots \lambda_{2n} = 1$. □

Remark 1.9. When $n = 1$, we have in the above lemma *if and only if*, i.e. a matrix $\mathcal{S} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is symplectic if and only if $\det \mathcal{S} = a\,d - b\,c = 1$.

The following lemma shows how to perturb a symplectic matrix $\mathcal{S}$ in order to obtain a new symplectic matrix $\underline{\mathcal{S}}$.

Lemma 1.10. *Let $\mathcal{S}$ be a symplectic $2n \times 2n$ matrix and define $\underline{\mathcal{S}} := \mathcal{S} + \mathcal{R}$ with $\mathcal{R} = \left(\begin{smallmatrix} 0 & 0 \\ G & H \end{smallmatrix}\right)$. The matrix $\underline{\mathcal{S}}$ is symplectic if and only if $G^T\mathcal{A}$ and $H^T\mathcal{B}$ are symmetric, and the identity $H^T\mathcal{A} = \mathcal{B}^T G$ holds.*

Proof. We have

$$\begin{aligned} \underline{\mathcal{S}}^T\mathcal{J}\underline{\mathcal{S}} &= (\mathcal{S}+\mathcal{R})^T\mathcal{J}(\mathcal{S}+\mathcal{R}) = \mathcal{S}^T\mathcal{J}\mathcal{S} + \mathcal{S}^T\mathcal{J}\mathcal{R} + \mathcal{R}^T\mathcal{J}\mathcal{S} + \mathcal{R}^T\mathcal{J}\mathcal{R} \\ &= \mathcal{J} + \mathcal{S}^T\mathcal{J}\mathcal{R} + \mathcal{R}^T\mathcal{J}\mathcal{S} = \mathcal{J} + \begin{pmatrix} \mathcal{A}^T G - G^T\mathcal{A} & \mathcal{A}^T H - G^T\mathcal{B} \\ \mathcal{B}^T G - H^T\mathcal{A} & \mathcal{B}^T H - H^T\mathcal{B} \end{pmatrix}. \end{aligned}$$

Hence, $\underline{\mathcal{S}}^T\mathcal{J}\underline{\mathcal{S}} = \mathcal{J}$ if and only if $H^T\mathcal{A} = \mathcal{B}^T G$ and $G^T\mathcal{A}$ and $H^T\mathcal{B}$ are symmetric. $\square$

1.3 Discrete symplectic systems

Let $N \in \mathbb{N}$ and let $X_k \in \mathbb{R}^{n\times n}$, $U_k \in \mathbb{R}^{n\times n}$ be real $n \times n$ matrices and $x_k \in \mathbb{R}^n$ and $u_k \in \mathbb{R}^n$ be real vectors, for $k \in [0, N+1]$ and $\mathcal{S}_k = \left(\begin{smallmatrix} \mathcal{A}_k & \mathcal{B}_k \\ \mathcal{C}_k & \mathcal{D}_k \end{smallmatrix}\right)$ be a symplectic $2n \times 2n$ matrix for $k \in [0, N]$. The pair of matrix sequences $\big(\{X_0, X_1, \dots, X_{N+1}\}, \{U_0, U_1, \dots, U_{N+1})\}\big)$ we denote by (X, U) and the pair of vector sequences $\big(\{x_0, x_1, \dots, x_{N+1}\}, \{u_0, u_1, \dots, u_{N+1})\}\big)$ we denote by (x, u).

In order to simplify the formulae, we sometimes omit the index k, when the formula holds for any $k \in [0, N+1]$.

Definition 1.11 (Discrete symplectic system). The system

$$X_{k+1} = \mathcal{A}_k X_k + \mathcal{B}_k U_k, \quad U_{k+1} = \mathcal{C}_k X_k + \mathcal{D}_k U_k, \quad k \in [0, N]. \tag{S}$$

is called a *discrete symplectic system*.

System (S) can be written as a vector or matrix system

$$\begin{pmatrix} x_{k+1} \\ u_{k+1} \end{pmatrix} = \mathcal{S}_k \begin{pmatrix} x_k \\ u_k \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} X_{k+1} \\ U_{k+1} \end{pmatrix} = \mathcal{S}_k \begin{pmatrix} X_k \\ U_k \end{pmatrix}, \quad k \in [0, N]. \tag{S}$$

Remark 1.12. As every symplectic matrix is invertible (see Lemma 1.6), system (S) has unique solutions for arbitrary initial point $k_0 \in [0, N]$ and initial values at k_0.

Remark 1.13. If a pair (X, U) is a matrix solution of (S), then

$$X_k = \mathcal{D}_k^T X_{k+1} - \mathcal{B}_k^T U_{k+1}, \quad U_k = -\mathcal{C}_k^T X_{k+1} + \mathcal{A}_k^T U_{k+1}. \tag{1.15}$$

This follows from (1.11).

Denote $W_k := \begin{pmatrix} \bar{X}_k \\ \bar{U}_k \end{pmatrix}^T \mathcal{J} \begin{pmatrix} X_k \\ U_k \end{pmatrix} = \bar{X}_k^T U_k - \bar{U}_k^T X_k$. The following theorem shows that if $(\bar{X}, \bar{U})$ and (X, U) are solutions of (S), then W_k is constant and, thus, we can write $W_k \equiv W$. The matrix W is called a *Wronskian*[3] of the solutions $(\bar{X}, \bar{U})$ and (X, U).

Theorem 1.14 (Wronskian identity). *If $(\bar{X}, \bar{U})$ and (X, U) solve (S), then W_k is constant on $[0, N+1]$.*

Proof. We show that $W_k = W_{k+1}$ for $k \in [0, N]$. We have

$$\begin{pmatrix} \bar{X}_{k+1} \\ \bar{U}_{k+1} \end{pmatrix}^T \mathcal{J} \begin{pmatrix} X_{k+1} \\ U_{k+1} \end{pmatrix} = \begin{pmatrix} \bar{X}_k \\ \bar{U}_k \end{pmatrix}^T \mathcal{S}_k^T \mathcal{J} \mathcal{S}_k \begin{pmatrix} X_k \\ U_k \end{pmatrix} = \begin{pmatrix} \bar{X}_k \\ \bar{U}_k \end{pmatrix}^T \mathcal{J} \begin{pmatrix} X_k \\ U_k \end{pmatrix},$$

and this implies $W_{k+1} = W_k \equiv W$ is constant everywhere on $[0, N+1]$. $\square$

Definition 1.15 (Conjoined basis). A matrix solution (X, U) of (S) is called a *conjoined solution* if $\begin{pmatrix} X_k \\ U_k \end{pmatrix}^T \mathcal{J} \begin{pmatrix} X_k \\ U_k \end{pmatrix} = 0$, i.e. if $X_k^T U_k$ symmetric on $[0, N+1]$. If, moreover, $\operatorname{rank} \begin{pmatrix} X_k \\ U_k \end{pmatrix} = \operatorname{rank}(X_k^T, U_k^T) = n$, then it is called a *conjoined basis.*

Remark 1.16. Because of Theorem 1.14, and from the fact that $\begin{pmatrix} X_{k+1} \\ U_{k+1} \end{pmatrix}$ is obtained from $\begin{pmatrix} X_k \\ U_k \end{pmatrix}$ via the multiplication by an invertible matrix (and vice versa), it is enough to check the properties of a conjoined basis at one index k, in particular at the initial point $k = 0$, since then they hold for all $k \in [0, N+1]$.

Definition 1.17 (Normalized conjoined bases). Two conjoined bases $(\bar{X}, \bar{U})$ and (X, U) are called *normalized conjoined bases* if $\begin{pmatrix} \bar{X}_k \\ \bar{U}_k \end{pmatrix}^T \mathcal{J} \begin{pmatrix} X_k \\ U_k \end{pmatrix} = I$, i.e. if $\bar{X}_k^T U_k - \bar{U}_k^T X_k = I$.

[3] Josef Hoëné-Wroński (∗ August 23, 1778, † August 8, 1853) was a Polish eccentric philosopher of mathematics.

Remark 1.18. The identity $\bar{X}^T U - \bar{U}^T X = I$ implies that $\operatorname{rank}\left(\begin{smallmatrix} X \\ U \end{smallmatrix}\right) = \operatorname{rank}\left(\begin{smallmatrix} \bar{X} \\ \bar{U} \end{smallmatrix}\right) = n$.

Lemma 1.19. *Solutions $(\bar{X}, \bar{U})$ and (X, U) are normalized conjoined bases if and only if the matrix $\left(\begin{smallmatrix} \bar{X} & X \\ \bar{U} & U \end{smallmatrix}\right)$ is symplectic.*

Proof. From (1.10) we have that $\left(\begin{smallmatrix} \bar{X} & X \\ \bar{U} & U \end{smallmatrix}\right)$ is symplectic if and only if

$$X^T U = U^T X, \quad \bar{X}^T \bar{U} = \bar{U}^T \bar{X}, \quad \bar{X}^T U - \bar{U}^T X = I, \tag{1.16}$$

which, by using Remark 1.18, is the definition of normalized conjoined bases. $\square$

From other properties (1.10)–(1.14) of symplectic matrices we further get that $(\bar{X}, \bar{U})$ and (X, U) are normalized conjoined bases if and only if

$$X \bar{X}^T = \bar{X} X^T, \quad U \bar{U}^T = \bar{U} U^T, \quad U \bar{X}^T - \bar{U} X^T = I, \tag{1.17}$$

$$\begin{pmatrix} \bar{X} & X \\ \bar{U} & U \end{pmatrix}^T = \begin{pmatrix} \bar{X}^T & \bar{U}^T \\ X^T & U^T \end{pmatrix} \quad \text{is symplectic,} \tag{1.18}$$

$$\begin{pmatrix} \bar{X} & X \\ \bar{U} & U \end{pmatrix}^{-1} = \begin{pmatrix} U^T & -X^T \\ -\bar{U}^T & \bar{X}^T \end{pmatrix} \quad \text{is symplectic.} \tag{1.19}$$

These equivalent conditions further imply that

$$\operatorname{rank} \begin{pmatrix} \bar{X} & X \end{pmatrix} = \operatorname{rank} \begin{pmatrix} \bar{U} & U \end{pmatrix} = n. \tag{1.20}$$

Lemma 1.20. *For any conjoined basis (X, U) there exists another conjoined basis $(\bar{X}, \bar{U})$ such that $(\bar{X}, \bar{U})$ and (X, U) are normalized conjoined bases.*

Proof. We take the (unique) solution $(\bar{X}, \bar{U})$ of (S) with $\bar{X}_0 = U_0(X_0^T X_0 + U_0^T U_0)^{-1}$, $\bar{U}_0 = -X_0(X_0^T X_0 + U_0^T U_0)^{-1}$. Then $(\bar{X}, \bar{U})$ and (X, U) are normalized conjoined bases. $\square$

Definition 1.21. The solution $(\hat{X}, \hat{U})$ of (S) with $\hat{X}_0 = 0$, $\hat{U}_0 = I$ is called the *principal solution* of (S). The solution $(\tilde{X}, \tilde{U})$ of (S) with $\tilde{X}_0 = I$, $\tilde{U}_0 = 0$ is called the *associated solution* of (S).

Remark 1.22. The solutions $(\tilde{X}, \tilde{U})$ and $(\hat{X}, \hat{U})$ from Definition 1.21 are normalized conjoined bases of (S), and $\left(\begin{smallmatrix} \tilde{X}_k & \hat{X}_k \\ \tilde{U}_k & \hat{U}_k \end{smallmatrix}\right) = \mathcal{S}_{k-1} \mathcal{S}_{k-2} \dots \mathcal{S}_0$ for all $k \in [1, N+1]$. Sometimes they are called the *special normalized conjoined bases* of (S).

Remark 1.23. In the literature another associated solution of (S) is often used, namely the solution (X, U) of (S) with $X_0 = -I$, $U_0 = 0$. Then $(\hat{X}, \hat{U})$ and (X, U) are normalized conjoined bases (in the opposite order compared to Definition 1.21). See e.g. [11, 17, 40].

For any normalized conjoined bases $(\bar{X}, \bar{U})$ and (X, U) we can write $\mathcal{S}_k$ in terms of these solutions.

Lemma 1.24. *Let* $(\bar{X}, \bar{U})$ *and* (X, U) *be normalized conjoined bases of* (S)*. Then*

$$\mathcal{S}_k = \begin{pmatrix} \bar{X}_{k+1} & X_{k+1} \\ \bar{U}_{k+1} & U_{k+1} \end{pmatrix} \begin{pmatrix} U_k{}^T & -X_k{}^T \\ -\bar{U}_k^T & \bar{X}_k^T \end{pmatrix}.$$

More specifically,

$$\begin{aligned} \mathcal{A}_k &= \bar{X}_{k+1} U_k^T - X_{k+1} \bar{U}_k^T, \quad & \mathcal{B}_k &= -\bar{X}_{k+1} X_k^T + X_{k+1} \bar{X}_k^T, \\ \mathcal{C}_k &= \bar{U}_{k+1} U_k^T - U_{k+1} \bar{U}_k^T, \quad & \mathcal{D}_k &= -\bar{U}_{k+1} X_k^T + U_{k+1} \bar{X}_k^T. \end{aligned} \tag{1.21}$$

Lemma 1.25. *Let* (X, U) *be a conjoined basis of* (S) *and let* $k \in [0, N]$*. The following conditions are equivalent.*

$$\operatorname{Ker} X_{k+1} \subseteq \operatorname{Ker} X_k, \tag{1.22}$$

$$X_k = X_k X_{k+1}^\dagger X_{k+1}, \tag{1.23}$$

$$\operatorname{Im} \mathcal{B}_k \subseteq \operatorname{Im} X_{k+1}, \tag{1.24}$$

$$\mathcal{B}_k = X_{k+1} X_{k+1}^\dagger \mathcal{B}_k. \tag{1.25}$$

Proof. From condition (1.5) we have the equivalence of conditions (1.22) and (1.23), and from condition (1.6) we have the equivalence of conditions (1.24) and (1.25). Now let condition (1.23) hold and let $(\bar{X}, \bar{U})$ be such that $(\bar{X}, \bar{U})$ and (X, U) are normalized conjoined bases of (S). (It exists by Lemma 1.20.) Using identity (1.23) in the formula for $\mathcal{B}_k$ in (1.21) and the symmetry of $\bar{X}_{k+1} X_{k+1}^T$, see formula (1.17), we get

$$\mathcal{B}_k = -\bar{X}_{k+1} X_{k+1}^T X_{k+1}^{\dagger T} X_k^T + X_{k+1} \bar{X}_k^T = X_{k+1} (\bar{X}_k^T - \bar{X}_{k+1}^T X_{k+1}^{\dagger T} X_k^T),$$

which implies condition (1.24).

Conversely, from identities (1.15) and (1.25) and symmetry of $X_{k+1}^T U_{k+1}$ we have

$$X_k = \mathcal{D}_k^T X_{k+1} - \mathcal{B}_k^T U_{k+1} = (\mathcal{D}_k^T - \mathcal{B}_k^T X_{k+1}^{\dagger T} U_{k+1}^T) X_{k+1},$$

which implies condition (1.22). The proof is complete. $\square$

Condition (1.22) is called the *kernel condition* and it plays an important role in the definition of a focal point. (See Section 3.1.) In this definition, there is also a matrix P, which we introduce in the next lemma.

Lemma 1.26. *If* (X, U) *is conjoined basis of* (S) *with* $\operatorname{Ker} X_{k+1} \subseteq \operatorname{Ker} X_k$ *for some* $k \in [0, N]$*, then the matrix*

$$P_k := X_k X_{k+1}^{\dagger} \mathcal{B}_k, \quad k \in [0, N], \tag{1.26}$$

is symmetric.

Proof. From the formula for $\mathcal{B}_k$ in (1.21) and from (1.23) we get

$$\begin{aligned} P_k &= X_k X_{k+1}^{\dagger} X_{k+1} \bar{X}_k^T - X_k X_{k+1}^{\dagger} \bar{X}_{k+1} X_k^T \\ &= X_k \bar{X}_k^T - X_k X_{k+1}^{\dagger} \bar{X}_{k+1} X_{k+1}^T X_{k+1}^{\dagger T} X_k^T, \end{aligned}$$

where $(\bar{X}, \bar{U})$ and (X, U) are normalized conjoined bases of (S). The last matrix is symmetric due to the symmetry of $X_k \bar{X}_k^T$ and $\bar{X}_{k+1} X_{k+1}^T$. □

1.4 Admissible sequences

Definition 1.27. A pair (x, u) satisfying

$$x_{k+1} = \mathcal{A}_k x_k + \mathcal{B}_k u_k \quad \text{for } k \in [0, N], \tag{1.27}$$

is called *admissible*. Equation (1.27) is called the *equation of motion*.

Remark 1.28. Sometimes we use a more precise term and we say that a pair (x, u) is *admissible with respect to* $(\mathcal{A}, \mathcal{B})$, or that (x, u) is $(\mathcal{A}, \mathcal{B})$-*admissible.*

For a given u and x_0, the equation of motion gives us a unique x such that the pair (x, u) is admissible. This is the content of the next lemma, which uses certain controllability matrices. First we define the transition matrices

$$\Phi_{k,j} := \mathcal{A}_{k-1} \mathcal{A}_{k-2} \dots \mathcal{A}_j \quad \text{for } k > j \quad \text{and} \quad \Phi_{k,k} := I.$$

Since the matrices $\mathcal{A}_k$ may be in general singular, $\Phi_{k,j}$ may also be in general singular. Next, we define the controllability matrices

$$G_0 := 0, \quad G_k := \begin{pmatrix} \Phi_{k,1}\mathcal{B}_0 & \Phi_{k,2}\mathcal{B}_1 & \dots & \Phi_{k,k-1}\mathcal{B}_{k-2} & \mathcal{B}_{k-1} \end{pmatrix} \in \mathbb{R}^{n \times (nk)}$$

and the restriction operator $\mathcal{T}_k : \mathbb{R}^{(N+1)n} \to \mathbb{R}^{nk}$ by

$$\mathcal{T}_k\, \underline{u} := \begin{pmatrix} u_0 \\ \vdots \\ u_{k-1} \end{pmatrix} \quad \text{with} \quad \underline{u} := \begin{pmatrix} u_0 \\ \vdots \\ u_N \end{pmatrix}. \tag{1.28}$$

In fact, $\mathcal{T}_k$ is the $kn \times (N+1)\,n$ matrix $\mathcal{T}_k = \begin{pmatrix} I_{kn\times kn} & 0_{kn\times(N+1-k)n} \end{pmatrix}$ and $\mathcal{T}_{N+1} = I$.

Lemma 1.29. *A pair (x, u) is admissible if and only if*

$$x_k = \begin{pmatrix} \Phi_{k,0} & G_k\mathcal{T}_k \end{pmatrix} \begin{pmatrix} x_0 \\ \underline{u} \end{pmatrix} \quad \text{for all } k \in [0, N+1] \tag{1.29}$$

with the $(N+1)n$-vector $\underline{u}$ defined in (1.28).

Proof. We have

$$\begin{aligned} \mathcal{A}_k x_k + \mathcal{B}_k u_k &= \mathcal{A}_k \begin{pmatrix} \Phi_{k,0} & G_k\mathcal{T}_k \end{pmatrix} \begin{pmatrix} x_0 \\ \underline{u} \end{pmatrix} + \mathcal{B}_k u_k \\ &= \begin{pmatrix} \Phi_{k+1,0} & G_{k+1}\mathcal{T}_{k+1} \end{pmatrix} \begin{pmatrix} x_0 \\ \underline{u} \end{pmatrix}, \end{aligned}$$

which implies the equivalence. □

Now we state an important lemma. It says that kernel condition (1.22) on $[0, N]$ implies the *image condition*, i.e. $x_k \in \operatorname{Im} X_k$ holds for all $k \in [0, N+1]$ and for an admissible (x, u) with $x_0 \in \operatorname{Im} X_0$.

Lemma 1.30. *If (X, U) is conjoined basis of* (S) *with* $\operatorname{Ker} X_{k+1} \subseteq \operatorname{Ker} X_k$ *for all $k \in [0, N]$ and (x, u) is admissible with $x_0 \in \operatorname{Im} X_0$, then $x_k \in \operatorname{Im} X_k$ for all $k \in [0, N+1]$.*

Proof. It suffices to show that $x_k \in \operatorname{Im} X_k$ implies $x_{k+1} \in \operatorname{Im} X_{k+1}$. Let $x_k = X_k c$. Then

$$\begin{aligned} x_{k+1} &= \mathcal{A}_k x_k + \mathcal{B}_k u_k = \mathcal{A}_k X_k c + \mathcal{B}_k U_k c + \mathcal{B}_k(u_k - U_k c) \\ &= X_{k+1} c + \mathcal{B}_k(u_k - U_k c) = X_{k+1}[c + X_{k+1}^{\dagger} \mathcal{B}_k(u_k - U_k c)], \end{aligned}$$

where we used (1.25). Thus, $x_{k+1} \in \operatorname{Im} X_{k+1}$. □

1.5 Discrete quadratic functional

For a pair (x,u) and symplectic $2n\times 2n$ matrices $\mathcal{S}_k = \left(\begin{smallmatrix}\mathcal{A}_k & \mathcal{B}_k\\ \mathcal{C}_k & \mathcal{D}_k\end{smallmatrix}\right)$, $k\in[0,N]$, we define the *discrete quadratic functional*

$$\mathcal{F}_0(x,u) := \sum_{k=0}^{N} \big\{ x_k^T \mathcal{A}_k^T \mathcal{C}_k x_k + 2\, x_k^T \mathcal{C}_k^T \mathcal{B}_k u_k + u_k^T \mathcal{B}_k^T \mathcal{D}_k u_k \big\}. \tag{1.30}$$

By using the last identity in (1.10), we can write $\mathcal{F}_0(x,u)$ in the equivalent form

$$\mathcal{F}_0(x,u) = \sum_{k=0}^{N} \big\{ (\mathcal{A}_k x_k + \mathcal{B}_k u_k)^T (\mathcal{C}_k x_k + \mathcal{D}_k u_k) - x_k^T u_k \big\}. \tag{1.31}$$

Lemma 1.31. *If (x,u) is a solution of* (S) *and $(\bar{x},\bar{u})$ is admissible, then*

$$\mathcal{F}_0(x,u) = x_k^T u_k\big|_0^{N+1}, \tag{1.32}$$

$$\mathcal{F}_0(\bar{x},\bar{u}) = \bar{x}_k^T \bar{u}_k\big|_0^{N+1} + \sum_{k=0}^{N} \bar{x}_{k+1}^T (\mathcal{C}_k \bar{x}_k + \mathcal{D}_k \bar{u}_k - \bar{u}_{k+1}), \tag{1.33}$$

$$\mathcal{F}_0(x+\bar{x}, u+\bar{u}) = x_k^T u_k\big|_0^{N+1} + 2\, \bar{x}_k^T u_k\big|_0^{N+1} + \mathcal{F}_0(\bar{x},\bar{u}). \tag{1.34}$$

Proof. Identities (1.32) and (1.33) can be directly seen from (1.31). For (1.34), we have

$$\mathcal{F}_0(x+\bar{x}, u+\bar{u}) = \mathcal{F}_0(x,u) + \mathcal{F}_0(x,\bar{u}) + \mathcal{F}_0(\bar{x},u) + \mathcal{F}_0(\bar{x},\bar{u}).$$

Now the two middle terms are equal, because

$$\begin{aligned}(\mathcal{A}_k \bar{x}_k + \mathcal{B}_k \bar{u}_k)^T (\mathcal{C}_k x_k + \mathcal{D}_k u_k) - \bar{x}_k^T u_k \\ = (\mathcal{A}_k x_k + \mathcal{B}_k u_k)^T (\mathcal{C}_k \bar{x}_k + \mathcal{D}_k \bar{u}_k) - x_k^T \bar{u}_k,\end{aligned}$$

and together with the first term they can be simplified as in (1.32). $\square$

Another way how to write $\mathcal{F}_0(x,u)$ for an admissible (x,u) is to replace $\mathcal{B}_k u_k$ in the last two terms in (1.30) by $x_{k+1} - \mathcal{A}_k x_k$. There exists a symmetric $n\times n$ matrix $\mathcal{E}_k$ such that $\mathcal{B}_k^T \mathcal{D}_k = \mathcal{B}_k^T \mathcal{E}_k \mathcal{B}_k$, for example $\mathcal{E}_k = \mathcal{B}_k \mathcal{B}_k^{\dagger} \mathcal{D}_k \mathcal{B}_k^{\dagger}$. Denote by

$$\mathcal{G}_k := \begin{pmatrix} \mathcal{A}_k^T \mathcal{E}_k \mathcal{A}_k - \mathcal{A}_k^T \mathcal{C}_k & \mathcal{C}_k^T - \mathcal{A}_k^T \mathcal{E}_k \\ \mathcal{C}_k - \mathcal{E}_k \mathcal{A}_k & \mathcal{E}_k \end{pmatrix}. \tag{1.35}$$

Lemma 1.32. *If (x,u) is admissible, then*

$$\mathcal{F}_0(x,u) = \sum_{k=0}^{N} \begin{pmatrix} x_k \\ x_{k+1} \end{pmatrix}^T \mathcal{G}_k \begin{pmatrix} x_k \\ x_{k+1} \end{pmatrix}. \tag{1.36}$$

Proof. This follows by a direct computation. □

1.5.1 Functionals with boundary conditions

Let $\mathcal{M}_0$, $\mathcal{M}_1$ be real $n\times n$ projections and Γ_0, Γ_1 be symmetric $n\times n$ matrices satisfying $\Gamma_i = (I-\mathcal{M}_i)\,\Gamma_i\,(I-\mathcal{M}_i)$, $i=0,1$, and let $\mathcal{M}$ be real $2n\times 2n$ projection and Γ be symmetric $2n\times 2n$ matrix satisfying $\Gamma = (I-\mathcal{M})\,\Gamma\,(I-\mathcal{M})$. We consider respectively

♣ the *functional with zero endpoints,*

$$\mathcal{F}_0(x,u) = \sum_{k=0}^{N} \left\{ x_k^T \mathcal{A}_k^T \mathcal{C}_k x_k + 2\, x_k^T \mathcal{C}_k^T \mathcal{B}_k u_k + u_k^T \mathcal{B}_k^T \mathcal{D}_k u_k \right\},$$
$$x_0 = 0, x_{N+1} = 0,$$

♡ the *functional with separated endpoints,*

$$\mathcal{F}(x,u) := x_0^T \Gamma_0 x_0 + x_{N+1}^T \Gamma_1 x_{N+1} + \mathcal{F}_0(x,u), \tag{1.37}$$
$$\mathcal{M}_0\, x_0 = 0, \mathcal{M}_1 x_{N+1} = 0,$$

♠ the *functional with general endpoints,*

$$\mathcal{F}(x,u) := \begin{pmatrix} x_0 \\ x_{N+1} \end{pmatrix}^T \Gamma \begin{pmatrix} x_0 \\ x_{N+1} \end{pmatrix} + \mathcal{F}_0(x,u), \tag{1.38}$$

$$\mathcal{M} \begin{pmatrix} x_0 \\ x_{N+1} \end{pmatrix} = 0.$$

Remark 1.33. The functional with zero endpoints is a special case of the functional with separated endpoints, when $\mathcal{M}_0 = \mathcal{M}_1 = I$ and $\Gamma_0 = \Gamma_1 = 0$.

Remark 1.34. The functional with separated endpoints is a special case of the functional with general endpoints, when $\mathcal{M} = \left(\begin{smallmatrix} \mathcal{M}_0 & 0 \\ 0 & \mathcal{M}_1 \end{smallmatrix}\right)$ and $\Gamma = \left(\begin{smallmatrix} \Gamma_0 & 0 \\ 0 & \Gamma_1 \end{smallmatrix}\right)$.

The following lemma is a modification of Lemma 1.29.

Lemma 1.35. *A pair* (x, u) *is admissible and* $\mathcal{M}_0 x_0 = 0$ *if and only if*

$$x_k = \begin{pmatrix} \Phi_{k,0}(I - \mathcal{M}_0) & G_k \mathcal{T}_k \end{pmatrix} \begin{pmatrix} x_0 \\ \underline{u} \end{pmatrix} \quad \textit{for all } k \in [0, N+1], \tag{1.39}$$

with $(N+1)n$*-vector* $\underline{u}$ *defined in* (1.28)*.*

Proof. It follows from Lemma 1.29. □

1.6 Notes

Various matrix properties can be found in [7]. The theory of the Moore-Penrose inverse including Lemma 1.2 is from [6], and Lemma 1.3 is from [34]. The properties of symplectic matrices can be found in [4,7,48]. In particular, Lemmas 1.7, 1.8 are from [48], while Lemma 1.10 is from [19, 37].

Discrete symplectic systems were introduced in [4]. Most of Section 1.3 is from [13], Lemma 1.25 is from [4, 46], while Section 1.4 is from [13, 40]. Finally, Lemma 1.32 can be found in [35] or [37].

Chapter 2

Various important tools

2.1 Picone identity

A Picone-type identity is used when we want to write a quadratic functional $\mathcal{F}$ in the form of a square and to show that $\mathcal{F}$ is nonnegative. (Which happens quite often in the proofs of roundabout theorems in the next chapter.) The Picone identity was discovered by M. Picone [52]. We present here its discrete version, introduced in [9] for Hamiltonian systems (H) and in [13] for discrete symplectic systems. Furthermore, in the next section we present a generalized version involving a parameter $\alpha \in \mathbb{R}^n$, which will be particularly useful for functionals with general endpoints.

First we introduce a symmetric matrix Q (which is closely related to matrix solutions of discrete symplectic system), a Riccati operator $R[Q]$, and another symmetric matrix $\mathcal{P}$, because they all appear in the Picone identity.

2.1.1 Matrix Q

For every pair (X, U) of $n \times n$ matrices with $X^T U$ symmetric there exists symmetric $n \times n$ matrix Q such that

$$QX = UX^\dagger X. \tag{2.1}$$

There are several possible ways how to define the matrix Q.

- Let

$$Q := UX^\dagger + (UX^\dagger)^T (I - XX^\dagger). \tag{2.2}$$

To show that the matrix Q is symmetric we have to use only the symmetry of $X^T U$ and $XX^\dagger$.

• If $\operatorname{rank}(X^T, U^T) = n$, then there exists a pair $(\bar{X}, \bar{U})$ such that $\bar{X}^T \bar{U}$ is symmetric and $\bar{X}^T U - \bar{U}^T X = I$, see Lemma 1.20. Then we can define

$$Q := UX^\dagger - (UX^\dagger \bar{X} - \bar{U})(I - X^\dagger X)U^T. \tag{2.3}$$

This definition of the matrix Q was introduced in [9] and it is more popular. (Used e.g. in [13], [18], [37].)

• Sometimes it suffices to have a symmetric matrix Q with

$$X^T Q X = X^T U, \tag{2.4}$$

which follows from (but is not equivalent to) (2.1). Then we can define

$$Q := XX^\dagger U X^\dagger. \tag{2.5}$$

• If the matrix X is invertible, then all previous definitions reduce to $Q = UX^{-1}$.

2.1.2 Riccati operator and matrix $\mathcal{P}$

Definition 2.1. For symmetric matrices Q_k, $k \in [0, N]$, we define the *discrete Riccati operator* $R[Q]_k$ associated with the symplectic system (S) by

$$R[Q]_k := Q_{k+1}(\mathcal{A}_k + \mathcal{B}_k Q_k) - (\mathcal{C}_k + \mathcal{D}_k Q_k). \tag{2.6}$$

Lemma 2.2. *If* (X, U) *is a conjoined basis of* (S) *and identity* (2.1) *holds for* Q_k*,* X_k*,* U_k*, and* Q_{k+1}*,* X_{k+1}*,* U_{k+1} *in place of* Q, X, U*, then*

$$X_{k+1}^T R[Q]_k X_k = 0,$$

$$X_k^T (\mathcal{A}_k + \mathcal{B}_k Q_k)^T R[Q]_k X_k = 0.$$

If moreover $\operatorname{Ker} X_{k+1} \subseteq \operatorname{Ker} X_k$*, then*

$$R[Q]_k X_k = 0. \tag{2.7}$$

Proof. It is a consequence of the identity

$$R[Q]_k X_k = U_{k+1}(X_{k+1}^\dagger X_{k+1} - I)X_k^\dagger X_k, \tag{2.8}$$

the symmetry of $X_k^T U_k$, and Lemma 1.25. □

Remark 2.3. Equation (2.7) is sometimes called an *implicit discrete Riccati matrix equation.* For another proof of equation (2.7) in Lemma 2.2 see [13].

Lemma 2.4. *If* (X, U) *is a conjoined basis of* (S) *and identity* (2.4) *holds for* Q_k, X_k, U_k, *and* Q_{k+1}, X_{k+1}, U_{k+1} *in place of* Q, X, U, *then*

$$X_{k+1}^T R[Q]_k X_k = (X_{k+1}^T Q_{k+1} - U_{k+1}^T)\mathcal{B}_k(Q_k X_k - U_k).$$

Proof. The following calculation

$$\begin{aligned} X_{k+1}^T R[Q]_k X_k &= X_{k+1}^T[Q_{k+1}X_{k+1} - U_{k+1} + (Q_{k+1}\mathcal{B}_k - \mathcal{D}_k)(Q_k X_k - U_k)] \\ &= X_k^T(Q_k X_k - U_k) + (X_{k+1}^T Q_{k+1} - U_{k+1}^T)\mathcal{B}_k(Q_k X_k - U_k) \end{aligned}$$

implies the identity. □

For symmetric $n \times n$ matrices Q_k we define a symmetric matrix

$$\mathcal{P}_k := \mathcal{D}_k^T \mathcal{B}_k - \mathcal{B}_k^T Q_{k+1}\mathcal{B}_k. \tag{2.9}$$

The next lemma shows that if kernel condition (1.22) holds, then the matrix $\mathcal{P}$ is equal to the matrix P defined by (1.26).

Lemma 2.5. *If* (X, U) *is a conjoined basis of* (S), $\operatorname{Ker} X_{k+1} \subseteq \operatorname{Ker} X_k$ *and* (2.1) *holds for* Q_k, X_k, U_k, *and* $Q_{k+1}, X_{k+1}, U_{k+1}$ *in place of* Q, X, U, *then*

$$\mathcal{P}_k = X_k X_{k+1}^\dagger \mathcal{B}_k.$$

Proof. From (1.25) we have

$$\begin{aligned} \mathcal{P}_k &= (\mathcal{D}_k^T - \mathcal{B}_k^T Q_{k+1})X_{k+1}X_{k+1}^\dagger \mathcal{B}_k \\ &= (\mathcal{D}_k^T X_{k+1} - \mathcal{B}_k^T U_{k+1})X_{k+1}^\dagger \mathcal{B}_k = X_k X_{k+1}^\dagger \mathcal{B}_k, \end{aligned}$$

and hence $\mathcal{P}_k = P_k$ holds. □

Remark 2.6. As the matrix $\mathcal{P}_k$ is symmetric, Lemma 1.26 is a corollary of Lemma 2.5.

2.1.3 Picone identity

Lemma 2.7. *Let (x,u) be admissible, Q_k symmetric for $k \in [0, N+1]$, and $w_k := u_k - Q_k x_k$ for $k \in [0, N]$. Then for $k \in [0, N]$ we have*

$$\begin{aligned} x_k^T \mathcal{A}_k^T \mathcal{C}_k x_k + 2\, x_k^T \mathcal{C}_k^T \mathcal{B}_k u_k + u_k^T \mathcal{B}_k^T \mathcal{D}_k u_k - \Delta(x_k^T Q_k x_k) - w_k^T \mathcal{P}_k w_k \\ = x_k^T (\mathcal{A}_k + \mathcal{B}_k Q_k)^T R[Q]_k x_k - 2x_{k+1}^T R[Q]_k x_k, \end{aligned} \tag{2.10}$$

and

$$(\mathcal{D}_k^T - \mathcal{B}_k^T Q_{k+1})\, x_{k+1} = x_k + \mathcal{P}_k w_k - \mathcal{B}_k^T R[Q]_k x_k. \tag{2.11}$$

Proof. Let $a_k := (\mathcal{A}_k + \mathcal{B}_k Q_k)\, x_k$ and $c_k := (\mathcal{C}_k + \mathcal{D}_k Q_k)\, x_k$. Then we have

$$\begin{aligned} \mathcal{A}_k x_k + \mathcal{B}_k u_k &= (\mathcal{A}_k + \mathcal{B}_k Q_k)\, x_k + \mathcal{B}_k w_k = a_k + \mathcal{B}_k w_k = x_{k+1}, \\ \mathcal{C}_k x_k + \mathcal{D}_k u_k &= (\mathcal{C}_k + \mathcal{D}_k Q_k)\, x_k + \mathcal{D}_k w_k = c_k + \mathcal{D}_k w_k. \end{aligned}$$

Now we prove identity (2.10) by showing

$$\begin{aligned} x_k^T \mathcal{A}_k^T \mathcal{C}_k x_k + {} & 2\, x_k^T \mathcal{C}_k^T \mathcal{B}_k u_k + u_k^T \mathcal{B}_k^T \mathcal{D}_k u_k \\ &= (\mathcal{A}_k x_k + \mathcal{B}_k u_k)^T (\mathcal{C}_k x_k + \mathcal{D}_k u_k) - x_k^T u_k \\ &= a_k^T c_k + w_k^T \mathcal{B}_k^T \mathcal{D}_k w_k + 2 w_k^T \mathcal{B}_k^T c_k - x_k^T Q_k x_k \\ &= (a_k^T + 2 w_k^T \mathcal{B}_k) c_k + w_k^T \mathcal{B}_k^T \mathcal{D}_k w_k + \Delta(x_k^T Q_k x_k) \\ &\quad - (a_k + \mathcal{B}_k w_k)^T Q_{k+1} (a_k + \mathcal{B}_k w_k) \\ &= (a_k^T + 2 w_k^T \mathcal{B}_k)(c_k - Q_{k+1} a_k) \\ &\quad + w_k^T (\mathcal{B}_k^T \mathcal{D}_k - \mathcal{B}_k^T Q_{k+1} \mathcal{B}_k) w_k + \Delta(x_k^T Q_k x_k) \\ &= (a_k^T - 2 x_{k+1}^T)(Q_{k+1} a_k - c_k) + w_k^T \mathcal{P}_k w_k + \Delta(x_k^T Q_k x_k) \\ &= [x_k^T (\mathcal{A}_k + \mathcal{B}_k Q_k) - 2 x_{k+1}^T]^T R[Q]_k x_k + w_k^T \mathcal{P}_k w_k + \Delta(x_k^T Q_k x_k). \end{aligned}$$

Finally we prove identity (2.11) by showing

$$\begin{aligned} (\mathcal{D}_k^T - \mathcal{B}_k^T Q_{k+1}) x_{k+1} &= (\mathcal{D}_k^T - \mathcal{B}_k^T Q_{k+1})(\mathcal{A}_k + \mathcal{B}_k Q_k) x_k + \mathcal{P}_k w_k \\ &= [I + \mathcal{B}_k^T \mathcal{C}_k + \mathcal{D}_k^T \mathcal{B}_k Q_k - \mathcal{B}_k^T Q_{k+1} (\mathcal{A}_k + \mathcal{B}_k Q_k)] x_k + \mathcal{P}_k w_k \\ &= x_k - \mathcal{B}_k^T R[Q]_k x_k + \mathcal{P}_k w_k. \end{aligned}$$

This completes the proof of this lemma. □

Lemma 2.8. *Let the assumptions from Lemma 2.7 hold and let* (X, U) *be a conjoined basis such that* $x_k \in \operatorname{Im} X_k$ *and* $x_{k+1} \in \operatorname{Im} X_{k+1}$*. Assume that identity* (2.1) *holds for* Q_k, X_k, U_k *and* $Q_{k+1}, X_{k+1}, U_{k+1}$ *in place of* Q, X, U*. Then*

$$x_k^T \mathcal{A}_k^T \mathcal{C}_k x_k + 2\, x_k^T \mathcal{C}_k^T \mathcal{B}_k u_k + u_k^T \mathcal{B}_k^T \mathcal{D}_k u_k - \Delta(x_k^T Q_k x_k) - w_k^T \mathcal{P}_k w_k = 0.$$

If moreover $\operatorname{Ker} X_{k+1} \subseteq \operatorname{Ker} X_k$*, then*

$$(\mathcal{D}_k^T - \mathcal{B}_k^T Q_{k+1})\, x_{k+1} = x_k + \mathcal{P}_k w_k. \tag{2.12}$$

Proof. It follows from Lemma 2.2 and Lemma 2.7. □

Theorem 2.9 (Picone identity). *If* (X, U) *is a conjoined basis of* (S) *and* (x, u) *is admissible with* $x_k \in \operatorname{Im} X_k$ *for* $k \in [0, N+1]$*, then*

$$\mathcal{F}_0(x, u) = x_k^T Q_k x_k \Big|_0^{N+1} + \sum_{k=0}^{N} w_k^T \mathcal{P}_k w_k,$$

where $w_k = u_k - Q_k x_k$*,* $\mathcal{P}_k = \mathcal{B}_k^T \mathcal{D}_k - \mathcal{B}_k^T Q_{k+1} \mathcal{B}_k$*, and* Q_k *are symmetric matrices with* $Q_k X_k = U_k X_k^\dagger X_k$*.*

Proof. It follows directly from Lemma 2.8. □

2.2 Generalized Picone identity

In this section we describe a transformation of the system (S) introduced in Section 1.3 into dimension $4n$. Thus we get a new bigger symplectic system, a quadratic functional, etc. and among other things also a new Picone-type identity.

2.2.1 Augmented symplectic system

First let us define the $2n \times 2n$ matrices

$$\mathcal{A}_k^* := \begin{pmatrix} I & 0 \\ 0 & \mathcal{A}_k \end{pmatrix}, \quad \mathcal{B}_k^* := \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{B}_k \end{pmatrix}, \quad \mathcal{C}_k^* := \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{C}_k \end{pmatrix}, \quad \mathcal{D}_k^* := \begin{pmatrix} I & 0 \\ 0 & \mathcal{D}_k \end{pmatrix}, \tag{2.13}$$

and the $4n \times 4n$ matrix $\mathcal{J}_k^* := \left(\begin{smallmatrix} 0 & I \\ -I & 0 \end{smallmatrix}\right)$.

Lemma 2.10. *The matrix* $\mathcal{S}_k^* := \begin{pmatrix} \mathcal{A}_k^* & \mathcal{B}_k^* \\ \mathcal{C}_k^* & \mathcal{D}_k^* \end{pmatrix}$ *is symplectic, i.e.* $\mathcal{S}_k^{*T} \mathcal{J}_k^* \mathcal{S}_k^* = \mathcal{J}_k^*$.

Proof. We have

$$\mathcal{A}_k^{*T} \mathcal{C}_k^* = \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{A}_k^T \mathcal{C}_k \end{pmatrix}, \quad \mathcal{B}_k^{*T} \mathcal{D}_k^* = \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{B}_k^T \mathcal{D}_k \end{pmatrix}, \tag{2.14}$$

$$\mathcal{A}_k^{*T} \mathcal{D}_k^* = \begin{pmatrix} I & 0 \\ 0 & \mathcal{A}_k^T \mathcal{D}_k \end{pmatrix}, \quad \mathcal{C}_k^{*T} \mathcal{B}_k^* = \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{C}_k^T \mathcal{B}_k \end{pmatrix}, \tag{2.15}$$

$$\mathcal{A}_k^{*T} \mathcal{D}_k^* - \mathcal{C}_k^{*T} \mathcal{B}_k^* = \begin{pmatrix} I & 0 \\ 0 & \mathcal{A}_k^T \mathcal{D}_k - \mathcal{C}_k^T \mathcal{B}_k \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}. \tag{2.16}$$

This together with (1.10) imply the statement. □

Let X_k^*, U_k^*, Q_k^* be real $2n \times 2n$ matrices, $k \in [0, N+1]$. Then the system

$$X_{k+1}^* = \mathcal{A}_k^* X_k^* + \mathcal{B}_k^* U_k^*, \quad U_{k+1}^* = \mathcal{C}_k^* X_k^* + \mathcal{D}_k^* U_k^*, \quad k \in [0, N], \tag{S*}$$

is an augmented discrete symplectic system.

We denote by $R^*[Q^*]$ the corresponding augmented Riccati operator, i.e.

$$R^*[Q^*] := Q_{k+1}^* (\mathcal{A}_k^* + \mathcal{B}_k^* Q_k^*) - (\mathcal{C}_k^* + \mathcal{D}_k^* Q_k^*). \tag{2.17}$$

Lemma 2.11. *A pair* (X^*, U^*) *is a solution of* (S*) *if and only if* $X^* = \begin{pmatrix} K & L \\ X & \bar{X} \end{pmatrix}$ *and* $U^* = \begin{pmatrix} M & N \\ U & \bar{U} \end{pmatrix}$*, where* K, L, M, N *are constant matrices and where* (X, U)*,* $(\bar{X}, \bar{U})$ *are solutions of* (S).

Proof. The following identities

$$\mathcal{A}_k^* X_k^* + \mathcal{B}_k^* U_k^* = \begin{pmatrix} K & L \\ \mathcal{A}_k X_k + \mathcal{B}_k U_k & \mathcal{A}_k \bar{X}_k + \mathcal{B}_k \bar{U}_k \end{pmatrix},$$

$$\mathcal{C}_k^* X_k^* + \mathcal{D}_k^* U_k^* = \begin{pmatrix} M & N \\ \mathcal{C}_k X_k + \mathcal{D}_k U_k & \mathcal{C}_k \bar{X}_k + \mathcal{D}_k \bar{U}_k \end{pmatrix},$$

imply the statement. □

Lemma 2.12. *Pairs* $(\bar{X}, \bar{U})$ *and* (X, U) *are normalized conjoined bases of* (S) *if and only if the pair* (X^*, U^*)*, defined by*

$$X^* := \begin{pmatrix} 0 & I \\ X & \bar{X} \end{pmatrix}, \quad U^* := \begin{pmatrix} -I & 0 \\ U & \bar{U} \end{pmatrix}, \tag{2.18}$$

is conjoined basis of (S*).

Proof. We have $\operatorname{rank}\left(\begin{smallmatrix} X^* \\ U^* \end{smallmatrix}\right) = \operatorname{rank}\begin{pmatrix} \begin{smallmatrix} 0 & I \\ X & \bar{X} \\ -I & 0 \\ U & \bar{U} \end{smallmatrix} \end{pmatrix} = 2n$. Furthermore, $X^{*T}U^* = \begin{pmatrix} \begin{smallmatrix} X^TU & X^T\bar{U} \\ -I+\bar{X}^TU & \bar{X}^T\bar{U} \end{smallmatrix} \end{pmatrix}$ and this is a symmetric matrix if and only if $(\bar{X}, \bar{U})$ and (X, U) are normalized conjoined bases of (S). ☐

Remark 2.13. We often use the conjoined basis defined by the principal solution $(\hat{X}, \hat{U})$ and the associated solution $(\tilde{X}, \tilde{U})$ of (S),

$$\hat{X}^* := \begin{pmatrix} 0 & I \\ \hat{X} & \tilde{X} \end{pmatrix}, \quad \hat{U}^* := \begin{pmatrix} -I & 0 \\ \hat{U} & \tilde{U} \end{pmatrix}. \tag{2.19}$$

Lemma 2.14. *Let $\alpha, \beta_k \in \mathbb{R}^n$ and define $2n$-vectors*

$$x_k^* := \begin{pmatrix} \alpha \\ x_k \end{pmatrix},\ k \in [0, N+1], \quad u_k^* := \begin{pmatrix} \beta_k \\ u_k \end{pmatrix},\ k \in [0, N].$$

Then (x^*, u^*) is admissible w.r.t. $(\mathcal{A}^*, \mathcal{B}^*)$ if and only if (x, u) is admissible w.r.t. $(\mathcal{A}, \mathcal{B})$.

Proof. We have $\mathcal{A}_k^* x_k^* + \mathcal{B}_k^* u_k^* = \left(\begin{smallmatrix} I & 0 \\ 0 & \mathcal{A}_k \end{smallmatrix}\right)\left(\begin{smallmatrix} \alpha \\ x_k \end{smallmatrix}\right) + \left(\begin{smallmatrix} 0 & 0 \\ 0 & \mathcal{B}_k \end{smallmatrix}\right)\left(\begin{smallmatrix} \beta_k \\ u_k \end{smallmatrix}\right) = \left(\begin{smallmatrix} \alpha \\ \mathcal{A}_k x_k + \mathcal{B}_k u_k \end{smallmatrix}\right)$, and $\left(\begin{smallmatrix} \alpha \\ \mathcal{A}_k x_k + \mathcal{B}_k u_k \end{smallmatrix}\right) = \left(\begin{smallmatrix} \alpha \\ x_{k+1} \end{smallmatrix}\right)$ if and only if the equation $\mathcal{A}_k x_k + \mathcal{B}_k u_k = x_{k+1}$ holds. ☐

Let $\mathcal{A}^*, \mathcal{B}^*, \mathcal{C}^*, \mathcal{D}^*$ be the $2n \times 2n$ matrices defined in (2.13). Then we introduce the following quadratic functional

$$\mathcal{F}_0^*(x^*, u^*) := \sum_{k=0}^{N} \left\{ x_k^{*T} \mathcal{A}_k^{*T} \mathcal{C}_k^* x_k^* + 2\, x_k^{*T} \mathcal{C}_k^{*T} \mathcal{B}_k^* u_k^* + u_k^{*T} \mathcal{B}_k^{*T} \mathcal{D}_k^* u_k^* \right\}.$$

Lemma 2.15. *Let $x_k^* := \left(\begin{smallmatrix} \alpha_k \\ x_k \end{smallmatrix}\right), k \in [0, N+1]$ and $u_k^* := \left(\begin{smallmatrix} \beta_k \\ u_k \end{smallmatrix}\right), k \in [0, N]$, where $\alpha_k, \beta_k, x_k, u_k$ are arbitrary n-vectors. Then $\mathcal{F}_0^*(x^*, u^*) = \mathcal{F}_0(x, u)$.*

Proof. It follows from identities (2.14) and (2.15). ☐

2.2.2 Big matrix Q^*

In this subsection, let $(\bar{X}, \bar{U})$ and (X, U) be normalized conjoined bases of (S) and let X^* and U^* be the $2n \times 2n$ matrices defined in (2.18).

Lemma 2.16. *The following identities hold*

$$X^{*\dagger} = \begin{pmatrix} -X^\dagger \bar{X} & X^\dagger \\ (I+\bar{X}^T\bar{X})^{-1}(I+\bar{X}^T X X^\dagger \bar{X}) & (I+\bar{X}^T\bar{X})^{-1}\bar{X}^T(I-XX^\dagger) \end{pmatrix} \tag{2.20}$$

and

$$X^{*\dagger}X^* = \begin{pmatrix} X^\dagger X & 0 \\ 0 & I \end{pmatrix}. \tag{2.21}$$

Proof. The four properties of the generalized inverse in (1.1) can be easily verified by a direct computation. □

Let Q be a symmetric matrix and define

$$Q^* := \begin{pmatrix} \bar{X}^T Q \bar{X} - \bar{X}^T \bar{U} & \bar{U}^T - \bar{X}^T Q \\ \bar{U} - Q\bar{X} & Q \end{pmatrix}. \tag{2.22}$$

Lemma 2.17. *Let the matrix Q^* be defined by* (2.22)*. Then Q^* is symmetric and*

(i) $QX = UX^\dagger X$ *if and only if* $Q^*X^* = U^*X^{*\dagger}X^*$,

(ii) $X^TQX = U^TX$ *if and only if* $X^{*T}Q^*X^* = U^{*T}X^*$.

Proof. The symmetry of Q^* follows from the symmetry of Q and $\bar{X}^T\bar{U}$. The equivalence in (i) is obtained from

$$Q^*X^* = \begin{pmatrix} \bar{X}^T(UX^\dagger X - QX) - X^\dagger X & 0 \\ QX & \bar{U} \end{pmatrix}, \quad U^*X^{*\dagger}X^* = \begin{pmatrix} -X^\dagger X & 0 \\ UX^\dagger X & \bar{U} \end{pmatrix},$$

where we used identities (1.16) for a conjoined basis and identity (2.21). The equivalence in (ii) follows from

$$X^{*T}Q^*X^* = \begin{pmatrix} X^TQX & X^T\bar{U} \\ \bar{U}^TX & \bar{X}^T\bar{U} \end{pmatrix}, \quad U^{*T}X^* = \begin{pmatrix} U^TX & U^T\bar{X} - I \\ \bar{U}^TX & \bar{U}^T\bar{X} \end{pmatrix}.$$

This completes the proof. □

Remark 2.18. Similarly as in Subsection 2.1.1, we could define

$$Q^* := U^*X^{*\dagger} + (U^*X^{*\dagger})^T(I - X^*X^{*\dagger}), \tag{2.23}$$

or

$$Q^* := U^* X^{*\dagger} + (U^* X^{*\dagger} \bar{X}^* - \bar{U}^*)(I - X^{*\dagger} X^*) U^{*T}. \tag{2.24}$$

However, in the next lemma we prove that all symmetric matrices Q^* with $Q^* X^* = U^* X^{*\dagger} X^*$ have the form (2.22), and thus these definitions can differ in fact only by the (right lower corner) matrix Q_k. The most often used definition is again the one with the matrix Q_k as defined by (2.3).

Lemma 2.19. *Let* $Q^* = \begin{pmatrix} \dot{Q} & \ddot{Q} \\ \ddot{Q}^T & Q \end{pmatrix}$ *be a symmetric* $2n \times 2n$ *matrix. Then* $Q^* X^* = U^* X^{*\dagger} X^*$ *if and only if* $\dot{Q} = \bar{X}^T Q \bar{X} - \bar{X}^T \bar{U}$, $\ddot{Q} = \bar{U}^T - \bar{X}^T Q$, *and* $QX = UX^\dagger X$.

Proof. The statement follows from

$$Q^* X^* = \begin{pmatrix} \ddot{Q} X & \dot{Q} + \ddot{Q} \bar{X} \\ QX & \ddot{Q}^T + Q\bar{X} \end{pmatrix} = \begin{pmatrix} -X^\dagger X & 0 \\ UX^\dagger X & \bar{U} \end{pmatrix} = U^* X^{*\dagger} X^*.$$

Namely, $QX = UX^\dagger X$, $\ddot{Q}^T = \bar{U} - Q\bar{X}$, and using this in the right upper corner of the above identity yields the rest. ☐

Lemma 2.20. *Let* Q^* *be defined by* (2.22) *and* $R^*[Q^*]$ *be defined by* (2.17) *. Then*

$$R^*[Q^*]_k = \begin{pmatrix} \bar{X}_{k+1}^T R[Q]_k \bar{X}_k & -\bar{X}_{k+1}^T R[Q]_k \\ -R[Q]_k \bar{X}_k & R[Q]_k \end{pmatrix} = \begin{pmatrix} -\bar{X}_{k+1}^T \\ I \end{pmatrix} R[Q]_k \begin{pmatrix} -\bar{X}_k & I \end{pmatrix}, \tag{2.25}$$

$$R^*[Q^*]_k X_k^* = \begin{pmatrix} -\bar{X}_{k+1}^T R[Q]_k X_k & 0 \\ R[Q]_k X_k & 0 \end{pmatrix} = \begin{pmatrix} -\bar{X}_{k+1}^T \\ I \end{pmatrix} R[Q]_k \begin{pmatrix} X_k & 0 \end{pmatrix}, \tag{2.26}$$

$$X_{k+1}^{*T} R^*[Q^*]_k X_k^* = \begin{pmatrix} X_{k+1}^T R[Q]_k X_k & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} X_{k+1}^T \\ 0 \end{pmatrix} R[Q]_k \begin{pmatrix} X_k & 0 \end{pmatrix}. \tag{2.27}$$

Proof. It can be directly computed from the definition of $R^*[Q^*]$ and X^*. ☐

Lemma 2.21. *Let* Q^* *be defined by* (2.22). *Then*

$$R^*[Q^*]_k = 0 \quad \Leftrightarrow \quad R[Q]_k = 0, \tag{2.28}$$

$$R^*[Q^*]_k X_k^* = 0 \quad \Leftrightarrow \quad R[Q]_k X_k = 0, \tag{2.29}$$

$$X_{k+1}^{*T} R^*[Q^*]_k X_k^* = 0 \quad \Leftrightarrow \quad X_{k+1}^T R[Q]_k X_k = 0. \tag{2.30}$$

Proof. It is a corollary of previous Lemma 2.20. $\square$

Lemma 2.22. *Let $Q_k^* = \left(\begin{smallmatrix} \star & \star \\ \star & Q_k \end{smallmatrix}\right)$ be a $2n \times 2n$ matrix and let $\mathcal{P}_k^* := \mathcal{B}_k^{*T}\mathcal{D}_k^* - \mathcal{B}_k^{*T}Q_{k+1}^*\mathcal{B}_k^*$. Then*

$$\mathcal{P}_k^* = \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{B}_k^T\mathcal{D}_k - \mathcal{B}_k^T Q_{k+1}\mathcal{B}_k \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{P}_k \end{pmatrix}.$$

Proof. We compute

$$\mathcal{P}_k^* = \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{B}_k^T \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & \mathcal{D}_k \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{B}_k^T \end{pmatrix}\begin{pmatrix} \star & \star \\ \star & Q_{k+1} \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & \mathcal{B}_k \end{pmatrix} \tag{2.31}$$

$$= \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{B}_k^T\mathcal{D}_k - \mathcal{B}_k^T Q_{k+1}\mathcal{B}_k \end{pmatrix}. \tag{2.32}$$

Hence, the identity is proven. $\square$

Remark 2.23. We often use the matrix $\hat{Q}^*$ defined by (2.22) with the associated solution $(\tilde{X}, \tilde{U})$ of (S), i.e. with $\tilde{X}_0 = I$ and $\tilde{U}_0 = 0$, instead of $(\bar{X}, \bar{U})$, and then we put

$$\hat{Q}^* := \begin{pmatrix} \tilde{X}^T Q \tilde{X} - \tilde{X}^T\tilde{U} & \tilde{U}^T - \tilde{X}^T Q \\ \tilde{U} - Q\tilde{X} & Q \end{pmatrix}. \tag{2.33}$$

2.2.3 Generalized Picone identity

The following two lemmas show the relation between the image and the kernel conditions for a conjoined basis of (S) and the corresponding conjoined basis of (S*).

Lemma 2.24. *Let $x^* := \left(\begin{smallmatrix} \alpha \\ x \end{smallmatrix}\right)$ be a $2n$-vector and X^* be the $2n \times 2n$ matrix defined in* (2.18)*. The following statements are equivalent.*

(i) $x^* \in \operatorname{Im} X^*$, (ii) $x - \bar{X}\alpha \in \operatorname{Im} X$, (iii) $\alpha + U^T x \in \operatorname{Im} X^T$.

Proof. First we show (i) $\Leftrightarrow$ (ii), then (ii) $\Rightarrow$ (iii), and finally (iii) $\Rightarrow$ (i).

(i) $\Leftrightarrow$ $\left(\begin{smallmatrix} \alpha \\ x \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & I \\ X & \bar{X} \end{smallmatrix}\right)\left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right)$ for some $\left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) \in \mathbb{R}^n$ $\Leftrightarrow$ $d = \alpha$ and $x = Xc + \bar{X}\alpha$ $\Leftrightarrow$ $x - \bar{X}\alpha \in \operatorname{Im} X$ $\Leftrightarrow$ (ii).

(ii) $\Leftrightarrow$ $x = Xc + \bar{X}\alpha$ $\Rightarrow$ $U^T x = U^T Xc + U^T\bar{X}\alpha$ $\Leftrightarrow$ $U^T x = X^T Uc + (X^T\bar{U} - I)\alpha$ $\Leftrightarrow$ $U^T x + \alpha = X^T(Uc + \bar{U}\alpha)$ $\Leftrightarrow$ (iii).

(iii) $\Leftrightarrow$ $U^T x + \alpha = X^T d$ $\Rightarrow$ $\bar{X}U^T x + \bar{X}\alpha = \bar{X}X^T d$ $\Leftrightarrow$ $(X\bar{U}^T - I)\,x + \bar{X}\alpha = X\bar{X}^T d$ $\Leftrightarrow$ $X(\bar{U}^T x - \bar{X}^T d) = x - \bar{X}\alpha$ $\Leftrightarrow$ (ii). Thus, all equivalences are proven. $\square$

Lemma 2.25. *Let $(\bar{X},\bar{U})$ and (X,U) be normalized conjoined bases of* (S) *and (X^*,U^*) be the conjoined basis of* (S*) *defined by* (2.18) *via $(\bar{X},\bar{U})$ and (X,U). Then*

$$\operatorname{Ker} X^*_{k+1} \subseteq \operatorname{Ker} X^*_k \quad \Leftrightarrow \quad \operatorname{Ker} X_{k+1} \subseteq \operatorname{Ker} X_k. \tag{2.34}$$

Proof. We have $\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right) \in \operatorname{Ker} X^*_k \Leftrightarrow 0 = X^*_k \left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right) = \left(\begin{smallmatrix} v \\ X_k u + \bar{X}_k v \end{smallmatrix}\right) \Leftrightarrow v = 0, u \in \operatorname{Ker} X_k$. This yields the equivalence in lemma. $\square$

Lemma 2.26. *Let $(\bar{X},\bar{U})$ and (X,U) be normalized conjoined bases of* (S) *and (X^*,U^*) be the conjoined basis of* (S*) *defined by* (2.18) *via $(\bar{X},\bar{U})$ and (X,U), and let $\operatorname{Ker} X_{k+1} \subseteq \operatorname{Ker} X_k$. Then*

$$P^*_k := X^*_k X^{*\dagger}_{k+1} \mathcal{B}^*_k = \begin{pmatrix} 0 & 0 \\ 0 & X_k X^\dagger_{k+1} \mathcal{B}_k \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & P_k \end{pmatrix}. \tag{2.35}$$

Proof. Identity (2.35) can be computed directly with the use of Lemma 2.16 and identity (1.25), applied to the augmented matrices $\mathcal{B}^*_k$ and X^*_k, as the kernel conditions for X and X^* are equivalent by Lemma 2.25. $\square$

Lemma 2.27. *Let $(\bar{X},\bar{U})$ and (X,U) be any normalized conjoined bases of* (S). *For any admissible (x,u) and symmetric Q_k on $[0,N+1]$ and for any $\alpha \in \mathbb{R}^n$ we have*

$$\begin{aligned} x_k^T \mathcal{A}_k^T \mathcal{C}_k x_k + 2\, x_k^T \mathcal{C}_k^T \mathcal{B}_k u_k + u_k^T \mathcal{B}_k^T \mathcal{D}_k u_k - \Delta \left\{ \begin{pmatrix} \alpha \\ x_k \end{pmatrix}^T Q^*_k \begin{pmatrix} \alpha \\ x_k \end{pmatrix} \right\} - \tilde{w}_k^T \mathcal{P}_k \tilde{w}_k \\ = (x_k - \bar{X}_k \alpha)^T (\mathcal{A}_k + \mathcal{B}_k Q_k)^T R[Q]_k (x_k - \bar{X}_k \alpha) \\ -2\, (x_{k+1} - \bar{X}_{k+1} \alpha)^T R[Q]_k (x_k - \bar{X}_k \alpha) \end{aligned}$$

for all $k \in [0,N]$, and the identity

$$\begin{aligned} (\mathcal{D}_k^T - \mathcal{B}_k^T Q_{k+1})\, x_{k+1} = x_k + \mathcal{P}_k \tilde{w}_k + \mathcal{B}_k^T (\bar{U}_{k+1} - Q_{k+1} \bar{X}_{k+1})\, \alpha \\ - \mathcal{B}_k^T R[Q]_k (x_k - \bar{X}_k \alpha) \end{aligned}$$

holds, where $\tilde{w}_k = u_k - \bar{U}_k \alpha - Q_k (x_k - \bar{X}_k \alpha)$, $\mathcal{P}_k = \mathcal{B}_k^T \mathcal{D}_k - \mathcal{B}_k^T Q_{k+1} \mathcal{B}_k$, and Q^*_k is defined by (2.22) *with the matrix Q_k.*

Proof. Let (x,u) be admissible. Then the pair $(\tilde{x},\tilde{u})$, where $\tilde{x}_k := x_k - \bar{X}_k \alpha$ and $\tilde{u}_k := u_k - \bar{U}_k \alpha$, is admissible. The desired identity follows from Lemma 2.7 applied to the pair $(\tilde{x},\tilde{u})$, where we used Lemma 1.31. $\square$

Theorem 2.28 (Generalized Picone identity). *Let* $(\bar{X}, \bar{U})$, (X, U) *be normalized conjoined bases of* (S). *If there exists* $\alpha \in \mathbb{R}^n$ *such that* (x, u) *is admissible with* $x_k - \bar{X}_k\alpha \in \operatorname{Im} X_k$ *for* $k \in [0, N+1]$, *then*

$$\mathcal{F}_0(x, u) = \begin{pmatrix} \alpha \\ x_k \end{pmatrix}^T Q_k^* \begin{pmatrix} \alpha \\ x_k \end{pmatrix} \Bigg|_0^{N+1} + \sum_{k=0}^{N} \tilde{w}_k^T \mathcal{P}_k \tilde{w}_k, \tag{2.36}$$

where $\tilde{w}_k = u_k - \bar{U}_k\alpha - Q_k(x_k - \bar{X}_k\alpha)$, $\mathcal{P}_k = \mathcal{B}_k^T\mathcal{D}_k - \mathcal{B}_k^T Q_{k+1}\mathcal{B}_k$, *and* Q^* *is defined by* (2.22) *with a symmetric matrix* Q *such that* $Q_kX_k = U_kX_k^\dagger X_k$.

If moreover $\operatorname{Ker} X_{k+1} \subseteq \operatorname{Ker} X_k$, *then*

$$(\mathcal{D}_k^T - \mathcal{B}_k^T Q_{k+1})(x_{k+1} - \bar{X}_{k+1}\alpha) = x_k - \bar{X}_k\alpha + \mathcal{P}_k\tilde{w}_k. \tag{2.37}$$

There are (at least) three possible ways of how to prove this theorem. We show here one proof that is via the augmented system (S*) and one that uses identity (1.34) of Lemma 1.31. Another proof can be based on Lemma 2.27.

Proof 1 of Theorem 2.28. We define the pair (x^*, u^*) as in Lemma 2.14. Then (x^*, u^*) is admissible w.r.t. $(\mathcal{A}^*, \mathcal{B}^*)$ and $x^* \in \operatorname{Im} X^*$ by Lemma 2.24. The pair (X^*, U^*), defined by (2.18) via normalized conjoined bases $(\bar{X}, \bar{U})$ and (X, U), is a conjoined basis of (S*), by Lemma 2.12. Then from the Picone identity (Theorem 2.9) we get

$$\mathcal{F}_0^*(x^*, u^*) = x_k^{*T} Q_k^* x_k^* \big|_0^{N+1} + \sum_{k=0}^{N} w_k^{*T} \mathcal{P}_k^* w_k^*,$$

where $w_k^* = u_k^* - Q_k^* x_k^*$, $\mathcal{P}_k^* = \mathcal{B}_k^{*T}\mathcal{D}_k^* - \mathcal{B}_k^{*T} Q_{k+1}^* \mathcal{B}_k^*$, and Q_k^* are symmetric matrices with $Q_k^* X_k^* = U_k^* X_k^{*\dagger} X_k^*$. Further, $\mathcal{P}_k^* = \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{B}_k^T\mathcal{D}_k - \mathcal{B}_k^T Q_{k+1}\mathcal{B}_k \end{pmatrix}$ by Lemma 2.22 and $w_k^{*T}\mathcal{P}_k^* w_k^* = \tilde{w}_k^T \mathcal{P}_k \tilde{w}_k$ with $\tilde{w}_k = (0, I) w_k^* = u_k - \bar{U}_k\alpha - Q_k(x_k - \bar{X}_k\alpha)$. Finally, by Lemma 2.15, we have $\mathcal{F}_0(x, u) = \mathcal{F}_0^*(x^*, u^*)$. Identity (2.37) is obtained from Lemma 2.8 applied to the augmented system (S*), and from Lemma 2.25. □

Proof 2 of Theorem 2.28. We define a pair $(\tilde{x}, \tilde{u})$ by $\tilde{x} := x - \bar{X}\alpha$, $\tilde{u} := u - \bar{U}\alpha$. Such a pair is admissible and $\tilde{x}_k \in \operatorname{Im} X_k$ for $k \in [0, N+1]$ by the definition of $\tilde{x}$ and assumption. Thus, the Picone identity holds for $(\tilde{x}, \tilde{u})$ and we have from Theorem 2.9

$$\mathcal{F}_0(\tilde{x}, \tilde{u}) = \tilde{x}_k^T Q_k \tilde{x}_k \big|_0^{N+1} + \sum_{k=0}^{N} \tilde{w}_k^T \mathcal{P}_k \tilde{w}_k.$$

Simultaneously, by Lemma 1.31, we have

$$\mathcal{F}_0(\tilde{x},\tilde{u}) = \mathcal{F}_0(x,u) + \alpha^T \bar{X}_k^T \bar{U}_k \alpha\Big|_0^{N+1} - 2\, x_k^T \bar{U}_k \alpha\Big|_0^{N+1}.$$

Identity (2.36) then follows from

$$\begin{pmatrix} \alpha \\ x_k \end{pmatrix}^T Q_k^* \begin{pmatrix} \alpha \\ x_k \end{pmatrix} = \tilde{x}^T Q \tilde{x} - \alpha^T \bar{X}^T \bar{U} \alpha + 2\, x^T \bar{U} \alpha,$$

where Q_k^* is defined by (2.22). Identity (2.37) is obtained from Lemma 2.8 applied to $(\tilde{x},\tilde{u})$. ☐

Remark 2.29. For $\alpha = 0$ we get the statement of Theorem 2.9.

2.3 Other identities

In this section we present identities describing the relation between the value of a functional $\mathcal{F}$ on a pair (x,u) to the value of the same functional $\mathcal{F}$ on another pair $(\bar{x},\bar{u})$ which satisfies given boundary conditions. These identities are used later in Section 3.4 when showing the definiteness of a perturbed quadratic functional. It is also possible to use them in the proofs of roundabout theorems for functionals with general endpoints.

2.3.1 Identity for zero endpoints

Recall that $(\hat{X},\hat{U})$ is the principal solution of (S), i.e. $(\hat{X}_0,\hat{U}_0) = (0,I)$, and $(\tilde{X},\tilde{U})$ is the associated solution of (S), i.e. $(\tilde{X}_0,\tilde{U}_0) = (I,0)$.

Lemma 2.30. *Let (x,u) be admissible with $x_{N+1} - \tilde{X}_{N+1} x_0 \in \operatorname{Im} \hat{X}_{N+1}$. Then the pair $(\bar{x},\bar{u})$, defined by*

$$\begin{gathered} \bar{x}_k := x_k - \tilde{X}_k x_0 - \hat{X}_k c, \quad \bar{u}_k := u_k - \tilde{U}_k x_0 - \hat{U}_k c, \\ \textit{where } c := \hat{X}_{N+1}^{\dagger}(x_{N+1} - \tilde{X}_{N+1} x_0), \end{gathered} \tag{2.38}$$

is admissible and $\bar{x}_0 = 0 = \bar{x}_{N+1}$.

Proof. The admissibility of $(\bar{x},\bar{u})$ follows from the fact that it is a sum of the admissible pairs (x,u), $(-\tilde{X}x_0, -\tilde{U}x_0)$, and $(-\hat{X}c, -\hat{U}c)$. Furthermore, $\bar{x}_0 = x_0 - \tilde{X}_0 x_0 - \hat{X}_0 c = 0$ and $\bar{x}_{N+1} = x_{N+1} - \tilde{X}_{N+1} x_0 - \hat{X}_{N+1} c = 0$, by Lemma 1.2. ☐

Theorem 2.31. *Let (x,u) be admissible with $x_{N+1} - \tilde{X}_{N+1}x_0 \in \operatorname{Im}\hat{X}_{N+1}$. Then*

$$\mathcal{F}_0(x,u) = \begin{pmatrix} x_0 \\ x_{N+1} \end{pmatrix}^T \hat{Q}^*_{N+1} \begin{pmatrix} x_0 \\ x_{N+1} \end{pmatrix} + \mathcal{F}_0(\bar{x},\bar{u}), \tag{2.39}$$

where $(\bar{x},\bar{u})$ is defined by (2.38) and $\hat{Q}^*$ is defined by (2.33) with the matrix $Q := \hat{X}\hat{X}^\dagger\hat{U}\hat{X}^\dagger$.

There are again more possible ways how to prove this theorem. We show here one longer, but more direct proof, where we use only identity (1.34), and one shorter, for which we in addition need the augmented quadratic functional introduced in the previous section.

Proof 1 of Theorem 2.31. From identity (1.34) we have

$$\begin{aligned}\mathcal{F}_0(x,u) = {} & (\tilde{X}x_0 + \hat{X}c)^T(\tilde{U}_k x_0 + \hat{U}_k c)\big|_0^{N+1} \\ & + 2\,\bar{x}_k^T(\tilde{U}_k x_0 + \hat{U}_k c)\big|_0^{N+1} + \mathcal{F}_0(\bar{x},\bar{u}).\end{aligned}$$

This is further equal to

$$\begin{aligned}\mathcal{F}_0(x,u) = {} & (\tilde{X}_{N+1}x_0 + \hat{X}_{N+1}c)^T(\tilde{U}_{N+1}x_0 + \hat{U}_{N+1}c) - x_0^T c + \mathcal{F}_0(\bar{x},\bar{u}) \\ = {} & x_0^T\tilde{X}_{N+1}^T\tilde{U}_{N+1}x_0 + 2\,x_0^T\tilde{U}_{N+1}^T\hat{X}_{N+1}c + c^T\hat{X}_{N+1}^T\hat{U}_{N+1}c + \mathcal{F}_0(\bar{x},\bar{u}) \\ = {} & \begin{pmatrix} c \\ x_0 \end{pmatrix}^T \begin{pmatrix} \hat{X}_{N+1}^T\hat{U}_{N+1} & \hat{X}_{N+1}^T\tilde{U}_{N+1} \\ \tilde{U}_{N+1}^T\hat{X}_{N+1} & \tilde{X}_{N+1}^T\tilde{U}_{N+1} \end{pmatrix} \begin{pmatrix} c \\ x_0 \end{pmatrix} + \mathcal{F}_0(\bar{x},\bar{u}).\end{aligned} \tag{2.40}$$

Now, since $\binom{c}{x_0} = \begin{pmatrix} -\hat{X}_{N+1}^\dagger\tilde{X}_{N+1} & \hat{X}_{N+1}^\dagger \\ I & 0 \end{pmatrix}\binom{x_0}{x_{N+1}}$, the first term in identity (2.40) is equal to $\binom{x_0}{x_{N+1}}^T \hat{Q}^*_{N+1}\binom{x_0}{x_{N+1}}$, where

$$\begin{aligned}\hat{Q}^*_{N+1} &= \begin{pmatrix} -\tilde{X}_{N+1} & I \\ I & 0 \end{pmatrix}^T \begin{pmatrix} \hat{X}_{N+1}\hat{X}_{N+1}^\dagger\hat{U}_{N+1}\hat{X}_{N+1}^\dagger & \tilde{U}_{N+1} \\ \tilde{U}_{N+1}^T & \tilde{X}_{N+1}^T\tilde{U}_{N+1} \end{pmatrix} \begin{pmatrix} -\tilde{X}_{N+1} & I \\ I & 0 \end{pmatrix} \\ &= \begin{pmatrix} \tilde{X}_{N+1}^T Q_{N+1}\tilde{X}_{N+1} - \tilde{X}_{N+1}^T\tilde{U}_{N+1} & \tilde{U}_{N+1}^T - \tilde{X}_{N+1}^T Q_{N+1} \\ \tilde{U}_{N+1} - Q_{N+1}\tilde{X}_{N+1} & Q_{N+1} \end{pmatrix},\end{aligned}$$

where $Q_{N+1} = \hat{X}_{N+1}\hat{X}_{N+1}^\dagger\hat{U}_{N+1}\hat{X}_{N+1}^\dagger$. □

Proof 2 of Theorem 2.31. We define augmented pairs $(x^*,u^*) := (\binom{x_0}{x}, \binom{0}{u})$ and $(\bar{x}^*,\bar{u}^*) := \left(x^* - \hat{X}^*\binom{c}{x_0}, u^* - \hat{U}^*\binom{c}{x_0}\right) = (\binom{0}{\bar{x}}, \binom{c}{\bar{u}})$. We have $\bar{x}_0^* =$

$\bar{x}^*_{N+1} = 0$ and $\hat{X}^*_{N+1}\left(\begin{smallmatrix} c \\ x_0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right)$. By Lemma 2.15, $\mathcal{F}_0(x,u) = \mathcal{F}^*_0(x^*,u^*)$ and $\mathcal{F}_0(\bar{x},\bar{u}) = \mathcal{F}^*_0(\bar{x}^*,\bar{u}^*)$. From identity (1.34) we further have

$$\begin{aligned}\mathcal{F}^*_0(x^*,u^*) &= \begin{pmatrix} c \\ x_0 \end{pmatrix}^T \hat{X}^{*T}_k \hat{U}^*_k \begin{pmatrix} c \\ x_0 \end{pmatrix}\Bigg|_0^{N+1} + \mathcal{F}^*_0(\bar{x}^*,\bar{u}^*) \\ &= \begin{pmatrix} x_0 \\ x_{N+1} \end{pmatrix}^T Q^*_{N+1} \begin{pmatrix} x_0 \\ x_{N+1} \end{pmatrix} + \mathcal{F}^*_0(\bar{x}^*,\bar{u}^*),\end{aligned}$$

where Q^*_{N+1} is a symmetric matrix with $\hat{X}^{*T}_{N+1} Q^*_{N+1} \hat{X}^*_{N+1} = \hat{X}^{*T}_{N+1} \hat{U}^*_{N+1}$. □

Remark 2.32. In the second proof we showed even more, in particular we proved that identity (2.39) holds for any symmetric matrix Q^*_{N+1} with $\hat{X}^{*T}_{N+1} Q^*_{N+1} \hat{X}^*_{N+1} = \hat{X}^{*T}_{N+1} \hat{U}^*_{N+1}$. The matrix $\hat{Q}^*_{N+1}$ defined by (2.33) with the matrix $Q_{N+1} = \hat{X}_{N+1} \hat{X}^\dagger_{N+1} \hat{U}_{N+1} \hat{X}^\dagger_{N+1}$ in right lower corner has this property by Lemma 2.17, as Q_{N+1} is the matrix defined by (2.5) via $\hat{X}_{N+1}$ and $\hat{U}_{N+1}$ in place of X and U, it is symmetric, and $\hat{X}^T_{N+1} Q_{N+1} \hat{X}_{N+1} = \hat{X}^T_{N+1} \hat{U}_{N+1}$.

Corollary 2.33. *Let (x,u) be admissible with $x_{N+1} - \tilde{X}_{N+1} x_0 = 0$. Then*

$$\mathcal{F}_0(x,u) = x_0^T \tilde{X}^T_{N+1} \tilde{U}_{N+1} x_0 + \mathcal{F}_0(\bar{x},\bar{u}), \tag{2.41}$$

where $(\bar{x},\bar{u})$ is defined by

$$\bar{x}_k := x_k - \tilde{X}_k x_0, \quad \bar{u}_k := u_k - \tilde{U}_k x_0. \tag{2.42}$$

Proof. It follows from (2.40), because in this case $c = 0$. □

Corollary 2.34. *Let (x,u) be admissible and let $\operatorname{Ker} \hat{X}_{k+1} \subseteq \operatorname{Ker} \hat{X}_k$ for all $k \in [0,N]$. Then identity (2.39) from Theorem 2.31 holds.*

In proof of this corollary we use the following lemma.

Lemma 2.35. *Let (x,u) be an admissible pair and let $\operatorname{Ker} \hat{X}_{k+1} \subseteq \operatorname{Ker} \hat{X}_k$ for all $k \in [0,N]$. Then $x_k - \tilde{X}_k x_0 \in \operatorname{Im} \hat{X}_k$ for all $k \in [0,N+1]$*

Proof. We use Lemma 1.30 with the conjoined basis $(\hat{X},\hat{U})$ and the admissible pair $(x_k - \tilde{X}_k x_0, u_k - \tilde{U}_k x_0)$. For $k = 0$ we have $x_0 - \tilde{X}_0 x_0 = 0 \in \operatorname{Im} \hat{X}_0$, and thus, by Lemma 1.30, the inclusion holds for all $k \in [0,N+1]$. □

Proof of Corollary 2.34. From Lemma 2.35 with $k = N+1$ we get $x_{N+1} - \tilde{X}_{N+1} x_0 \in \operatorname{Im} \hat{X}_{N+1}$ and hence, the statement of Corollary 2.34 follows from Theorem 2.31. □

Remark 2.36. It is possible to use the generalized Picone identity to prove identity (2.39) from Corollary 2.34 with $\hat{Q}^*$ defined by (2.33) with a symmetric matrix Q such that $Q\hat{X} = \hat{U}\hat{X}^\dagger\hat{X}$. In the proof we need the assumption $\operatorname{Ker}\hat{X}_{k+1} \subseteq \operatorname{Ker}\hat{X}_k$ for all $k \in [0, N]$, hence this way of proof cannot be used for Theorem 2.31 itself.

From Theorem 2.9 we get $\mathcal{F}_0(\bar{x}, \bar{u}) = \sum_{k=0}^{N} w_k^T \mathcal{P}_k w_k$, where $w_k = \bar{u}_k - Q_k\bar{x}_k = (u_k - \tilde{U}_k x_0) - Q_k(x_k - \tilde{X}_k x_0) - (\hat{U}_k c - Q_k\hat{X}_k c)$. The kernel condition $\operatorname{Ker}\hat{X}_{k+1} \subseteq \operatorname{Ker}\hat{X}_k$ implies $\mathcal{P}_k(\hat{U}_k c - Q_k\hat{X}_k c) = 0$, and hence, with the use of Lemma 2.35, identity (2.39) follows from Theorem 2.28.

2.3.2 Identity for separated endpoints

Recall that the matrices $\mathcal{M}_0, \mathcal{M}_1$ are defined in Section 1.5.1 and $\hat{X}^*, \hat{U}^*$ are defined in (2.19).

Lemma 2.37. *Let (x, u) be an admissible pair with $\mathcal{M}_1 x_{N+1} - \tilde{X}_{N+1}\mathcal{M}_0 x_0 \in \operatorname{Im}\hat{X}_{N+1}$. Then the pair $(\bar{x}, \bar{u})$, defined by*

$$\begin{gathered}\bar{x}_k := x_k - \tilde{X}_k\mathcal{M}_0 x_0 - \hat{X}_k c, \quad \bar{u}_k := u_k - \tilde{U}_k\mathcal{M}_0 x_0 - \hat{U}_k c, \\ \textit{where } c := \hat{X}_{N+1}^\dagger(\mathcal{M}_1 x_{N+1} - \tilde{X}_{N+1}\mathcal{M}_0 x_0),\end{gathered} \tag{2.43}$$

is admissible and $\mathcal{M}_0\bar{x}_0 = \mathcal{M}_1\bar{x}_{N+1} = 0$.

Proof. The admissibility of $(\bar{x}, \bar{u})$ follows from the fact that it is a sum of the admissible pairs (x, u), $(-\tilde{X}\mathcal{M}_0 x_0, -\tilde{U}\mathcal{M}_0 x_0)$, and $(-\hat{X}c, -\hat{U}c)$. Furthermore, $\bar{x}_0 = (I - \mathcal{M}_0)x_0$, and $\bar{x}_{N+1} = (I - \mathcal{M}_1)x_{N+1}$, by Lemma 1.2. ☐

Theorem 2.38. *Let (x, u) be admissible with $\mathcal{M}_1 x_{N+1} - \tilde{X}_{N+1}\mathcal{M}_0 x_0 \in \operatorname{Im}\hat{X}_{N+1}$. Then*

$$\begin{aligned}\mathcal{F}_0(x, u) = \mathcal{F}_0(\bar{x}, \bar{u}) &- \begin{pmatrix}\mathcal{M}_0 x_0 \\ \mathcal{M}_1 x_{N+1}\end{pmatrix}^T Q^* \begin{pmatrix}\mathcal{M}_0 x_0 \\ \mathcal{M}_1 x_{N+1}\end{pmatrix} \\ &+ 2\begin{pmatrix}x_0 \\ x_{N+1}\end{pmatrix}^T \hat{U}^*_{N+1}\hat{X}^{*\dagger}_{N+1}\begin{pmatrix}\mathcal{M}_0 x_0 \\ \mathcal{M}_1 x_{N+1}\end{pmatrix},\end{aligned} \tag{2.44}$$

where $(\bar{x}, \bar{u})$ is defined in (2.43) and Q^* is a symmetric matrix with $\hat{X}^{*T}_{N+1}Q^*\hat{X}^*_{N+1} = \hat{X}^{*T}_{N+1}\hat{U}^*_{N+1}$.

Proof. We define augmented pairs $(x^*,u^*) := \left(\left(\begin{smallmatrix} x_0 \\ x \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ u \end{smallmatrix}\right)\right)$ and $(\bar{x}^*,\bar{u}^*) := \left(x^* - \hat{X}^* \left(\begin{smallmatrix} c \\ \mathcal{M}_0 x_0 \end{smallmatrix}\right), u^* - \hat{U}^* \left(\begin{smallmatrix} c \\ \mathcal{M}_0 x_0 \end{smallmatrix}\right)\right) = \left(\left(\begin{smallmatrix} (I-\mathcal{M}_0)x_0 \\ \bar{x} \end{smallmatrix}\right), \left(\begin{smallmatrix} c \\ \bar{u} \end{smallmatrix}\right)\right)$. Then we have

$$\hat{X}^*_{N+1} \begin{pmatrix} c \\ \mathcal{M}_0 x_0 \end{pmatrix} = \begin{pmatrix} \mathcal{M}_0 x_0 \\ \mathcal{M}_1 x_{N+1} \end{pmatrix},$$

and

$$\begin{pmatrix} c \\ \mathcal{M}_0 x_0 \end{pmatrix} = \begin{pmatrix} -\hat{X}^\dagger \tilde{X} & \hat{X}^\dagger \\ I & 0 \end{pmatrix} \begin{pmatrix} \mathcal{M}_0 x_0 \\ \mathcal{M}_1 x_{N+1} \end{pmatrix} = \hat{X}^{*\dagger}_{N+1} \begin{pmatrix} \mathcal{M}_0 x_0 \\ \mathcal{M}_1 x_{N+1} \end{pmatrix}.$$

By Lemma 2.15, $\mathcal{F}_0(x,u) = \mathcal{F}_0^*(x^*,u^*)$ and $\mathcal{F}_0(\bar{x},\bar{u}) = \mathcal{F}_0^*(\bar{x}^*,\bar{u}^*)$. From identity (1.34) we further have

$$\begin{aligned} \mathcal{F}_0^*(x^*,u^*) = & \begin{pmatrix} c \\ \mathcal{M}_0 x_0 \end{pmatrix}^T \hat{X}^{*T}_k \hat{U}^*_k \begin{pmatrix} c \\ \mathcal{M}_0 x_0 \end{pmatrix} \Bigg|_0^{N+1} + 2\, \bar{x}^{*T}_k \hat{U}^*_k \begin{pmatrix} c \\ \mathcal{M}_0 x_0 \end{pmatrix} \Bigg|_0^{N+1} \\ & + \mathcal{F}_0^*(\bar{x}^*,\bar{u}^*) \qquad (2.45) \\ = & - \begin{pmatrix} \mathcal{M}_0 x_0 \\ \mathcal{M}_1 x_{N+1} \end{pmatrix}^T Q^*_{N+1} \begin{pmatrix} \mathcal{M}_0 x_0 \\ \mathcal{M}_1 x_{N+1} \end{pmatrix} \\ & + 2 \begin{pmatrix} x_0 \\ x_{N+1} \end{pmatrix}^T \hat{U}^*_{N+1} \hat{X}^{*\dagger}_{N+1} \begin{pmatrix} \mathcal{M}_0 x_0 \\ \mathcal{M}_1 x_{N+1} \end{pmatrix} + \mathcal{F}_0^*(\bar{x}^*,\bar{u}^*), \end{aligned}$$

where Q^*_{N+1} is a symmetric matrix with $\hat{X}^{*T}_{N+1} Q^*_{N+1} \hat{X}^*_{N+1} = \hat{X}^{*T}_{N+1} \hat{U}^*_{N+1}$. $\square$

Corollary 2.39. *Let (x,u) be admissible with $\mathcal{M}_1 x_{N+1} - \tilde{X}_{N+1} \mathcal{M}_0 x_0 = 0$. Then*

$$\mathcal{F}_0(x,u) = (\mathcal{M}_0 x_0)^T \tilde{X}^T_{N+1} \tilde{U}_{N+1} \mathcal{M}_0 x_0 + 2\, x^T_{N+1} \tilde{U}_{N+1} \mathcal{M}_0 x_0 + \mathcal{F}_0(\bar{x},\bar{u}), \quad (2.46)$$

where $(\bar{x},\bar{u})$ is defined by

$$\bar{x}_k := x_k - \tilde{X}_k \mathcal{M}_0 x_0, \quad \bar{u}_k := u_k - \tilde{U}_k \mathcal{M}_0 x_0. \quad (2.47)$$

Proof. It follows from (2.45), because in this case $c = 0$. $\square$

2.4 Notes

The Picone identity in Subsection 2.1.3 is from [13]. The transformation from Subsection 2.2.1 can be found in [17,40,42] and Lemma 2.16 is from [8]. Formula (2.22) for Q^* as well as Lemmas 2.19–2.21, 2.27 are new and some of them will appear in [39]. Section 2.3 is new.

Chapter 3

Definiteness of quadratic functionals

In this chapter we characterize the definiteness of discrete quadratic functionals in terms of the nonexistence of focal points of conjoined bases of the corresponding symplectic system (S), and implicit and explicit Riccati equations and inequalities. The positivity of $\mathcal{F}$ can be characterized also in terms of conjugate intervals, as it is done e.g. in [40], but we omit this characterization here.

In Section 3.1 we deal with the positivity of quadratic functionals and in Section 3.2 with the nonnegativity of quadratic functionals. In both sections, we consider respectively functionals with *zero* endpoints, with *separated* endpoints, and with *general* (or *jointly varying*) endpoints. All functionals are defined in Subsection 1.5.1. In Section 3.3 we compare various forms of implicit Riccati equations for all types of discrete quadratic functionals. Section 3.4 is devoted to perturbation conditions for the nonnegativity and positivity of quadratic functionals.

The functional $\mathcal{F}$ is *nonnegative* (or *nonnegative definite*) if it takes nonnegative values on all admissible pairs (x, u) satisfying the given boundary conditions, while $\mathcal{F}$ is *positive* (or *positive definite*) if it takes positive values on all such admissible pairs (x, u) with $x \not\equiv 0$. Considering the nonnegativity and positivity of $\mathcal{F}$, we will always assume that the corresponding pairs (x, u) are admissible without specifying this any further.

3.1 Positivity of quadratic functionals

We begin this section with the focal point definition for conjoined basis of (S).

Definition 3.1. A conjoined basis (X, U) of (S) has a *focal point* in the interval $(m, m+1]$ if one of the following conditions hold.

(i) $\operatorname{Ker} X_{m+1} \nsubseteq \operatorname{Ker} X_m$,

(ii) $\operatorname{Ker} X_{m+1} \subseteq \operatorname{Ker} X_m$ and $P_m := X_m X_{m+1}^{\dagger} B_m \ngeq 0$.

Remark 3.2. If $\operatorname{Ker} X_{m+1} \subseteq \operatorname{Ker} X_m$, then P_m is symmetric, by Lemma 1.26, and $P_m = \mathcal{P}_m$, by Lemma 2.5.

Remark 3.3. According to the definition of a focal point, a conjoined basis (X, U) of (S) has no focal points in the interval $(m, m+1]$ if and only if

the *kernel condition* $\quad \operatorname{Ker} X_{m+1} \subseteq \operatorname{Ker} X_m$,

and the *P-condition* $\quad P_m \geq 0$

hold. In particular, conditions on no focal points in $(0, N+1]$ are widely used.

3.1.1 Functional with zero endpoints

In this section we state and prove a roundabout theorem for the positivity of the functional with zero endpoints, together with several auxiliary lemmas.

Theorem 3.4. *The following statements are equivalent.*

(i) $\mathcal{F}_0(x, u) > 0$ *over* $x_0 = 0$, $x_{N+1} = 0$, *and* $x \not\equiv 0$.

(ii) *The principal solution* $(\hat{X}, \hat{U})$ *of* (S) *has no focal points in* $(0, N+1]$.

(iii) *The implicit Riccati equation*

$$R[Q]_k G_k = 0, \quad k \in [0, N],$$

has a symmetric solution Q_k *on* $[0, N+1]$ *such that* $\mathcal{P}_k = \mathcal{B}_k^T \mathcal{D}_k - \mathcal{B}_k^T Q_{k+1} \mathcal{B}_k \geq 0$ *for all* $k \in [0, N]$.

(iv) *There exists a conjoined basis* (X, U) *of* (S) *such that* X_k *is invertible for all* $k \in [0, N+1]$, *and* $P_k = X_k X_{k+1}^{-1} \mathcal{B}_k \geq 0$ *on* $[0, N]$.

(v) *There exists a symmetric solution* Q_k *on* $[0, N+1]$ *of the explicit Riccati equation*

$$R[Q]_k = Q_{k+1}(\mathcal{A}_k + \mathcal{B}_k Q_k) - (\mathcal{C}_k + \mathcal{D}_k Q_k) = 0, \quad k \in [0, N], \tag{3.1}$$

with

$$\mathcal{A}_k + \mathcal{B}_k Q_k \textit{ invertible and } (\mathcal{A}_k + \mathcal{B}_k Q_k)^{-1} \mathcal{B}_k \geq 0, \quad k \in [0, N]. \tag{3.2}$$

(vi) *The system*

$$\begin{aligned} X_{k+1} &= \mathcal{A}_k X_k + \mathcal{B}_k U_k, \\ N_k := X_{k+1}^T (U_{k+1} &- \mathcal{C}_k X_k - \mathcal{D}_k U_k) \leq 0, \quad k \in [0, N] \end{aligned} \tag{3.3}$$

has a solution (X, U) *such that* $X_k^T U_k$ *is symmetric and* X_k *is invertible for all* $k \in [0, N+1]$ *and* $P_k = X_k X_{k+1}^{-1} \mathcal{B}_k \geq 0$ *on* $[0, N]$.

(vii) *There exists a symmetric solution* Q_k *on* $[0, N+1]$ *of the Riccati inequality*

$$R[Q]_k (\mathcal{A}_k + \mathcal{B}_k Q_k)^{-1} \leq 0, \quad k \in [0, N] \tag{3.4}$$

such that condition (3.2) *holds.*

Now we separate these statements into several independent lemmas. The proof of Theorem 3.4 is postponed to page 47. In Lemmas 3.5, 3.7, 3.8, we number the corresponding statements accordingly to Theorem 3.4.

Lemma 3.5. *The following statements are equivalent.*

(i) $\mathcal{F}_0(x, u) > 0$ *over* $x_0 = 0$, $x_{N+1} = 0$, *and* $x \not\equiv 0$.

(ii) *The principal solution* $(\hat{X}, \hat{U})$ *of* (S) *has no focal points in* $(0, N+1]$.

This lemma is a corollary of the following one.

Lemma 3.6. *Let* (X, U) *be a conjoined basis of* (S). *The following statements are equivalent.*

(i') $\mathcal{F}_0(x, u) + d^T X_0^T U_0\, d > 0$ *over* $x_0 = X_0\, d$, $x_{N+1} = 0$, *and* $x \not\equiv 0$.

(ii') *(X, U) has no focal points in $(0, N+1]$.*

The necessity of the kernel condition and the P-condition is shown by constructing a counterexample. The sufficiency is proven with the use of the Picone identity introduced in Subsection 2.1.3.

Proof of Lemma 3.6. First we show that (i') implies the kernel condition. Let (i') hold and suppose that there exists $m \in [0, N]$ such that $\operatorname{Ker} X_{m+1} \nsubseteq \operatorname{Ker} X_m$. Then there exists $0 \neq d \in \mathbb{R}^n$ such that $X_{m+1}d = 0$ and $X_m d \neq 0$. Now we define (x, u) by

$$(x_k, u_k) := \begin{cases} (X_k\, d, U_k\, d), & \text{for } k \in [0, m], \\ (0, 0), & \text{for } k \in [m+1, N+1]. \end{cases} \tag{3.5}$$

This (x, u) is admissible, $x_0 = X_0\, d$, $x_{N+1} = 0$, and $x_m = X_m d \neq 0$, thus assumption (i') implies that $\mathcal{F}_0(x, u) + d^T X_0^T U_0\, d > 0$. But

$$\begin{aligned} \mathcal{F}_0(x, u) &= \left\{ \sum_{k=0}^{m} + \sum_{k=m+1}^{N} \right\} \{x_k^T \mathcal{A}_k^T \mathcal{C}_k x_k + 2\, x_k^T \mathcal{C}_k^T \mathcal{B}_k u_k + u_k^T \mathcal{B}_k^T \mathcal{D}_k u_k\} \\ &= \sum_{k=0}^{m} d^T \{X_k^T \mathcal{A}_k^T \mathcal{C}_k X_k + 2\, X_k^T \mathcal{C}_k^T \mathcal{B}_k U_k + U_k^T \mathcal{B}_k^T \mathcal{D}_k U_k\}\, d, \end{aligned}$$

and, by Lemma 1.31, identity (1.32), we further have

$$\mathcal{F}_0(x, u) = d^T X_{m+1}^T U_{m+1} d - d^T X_0^T U_0\, d = -d^T X_0^T U_0\, d.$$

Thus, $\mathcal{F}_0(x, u) + d^T X_0^T U_0\, d = 0$, which is a contradiction with $\mathcal{F}_0(x, u) + d^T X_0^T U_0\, d > 0$.

Now we prove that (i') together with the kernel condition imply the P-condition. The matrix P is symmetric, by Lemma 1.26. Suppose that there exists $m \in [0, N]$ such that $P_m \ngeq 0$. Then there exists $c \in \mathbb{R}^n$ such that $c^T P c < 0$. Define $d := X_{m+1}^\dagger \mathcal{B}_m c$. Then $X_m d = X_m X_{m+1}^\dagger \mathcal{B}_m c = P_m c \neq 0$. Now we define (x, u) by

$$x_k := \begin{cases} X_k\, d, & \text{for } k \in [0, m], \\ 0, & \text{for } k > m, \end{cases}$$

$$u_k := \begin{cases} U_k\, d, & \text{for } k \in [0, m-1], \\ -\mathcal{A}_m^T X_{m+1}^{\dagger T} X_m^T c, & \text{for } k = m, \\ 0, & \text{for } k > m. \end{cases}$$

Again, this (x,u) is admissible, $x_0 = X_0\, d$, $x_{N+1} = 0$, and $x_m = X_m d \neq 0$. Thus, assumption (i') implies that $\mathcal{F}_0(x,u) + d^T X_0^T U_0\, d > 0$. But when we use identity (1.33) and equivalence (1.5), we get

$$\begin{aligned}
\mathcal{F}_0(x,u) + d^T X_0^T U_0\, d &= \left\{ \sum_{k=0}^{m-2} + \sum_{k=m}^{N} \right\} (\mathcal{C}_k x_k + \mathcal{D}_k u_k - u_{k+1})^T x_{k+1} \\
&\quad + (\mathcal{C}_{m-1} x_{m-1} + \mathcal{D}_{m-1} u_{m-1} - u_m)^T x_m \\
&= (\mathcal{C}_{m-1} x_{m-1} + \mathcal{D}_{m-1} u_{m-1} - u_m)^T x_m \\
&= [d^T (\mathcal{C}_{m-1} X_{m-1} + \mathcal{D}_{m-1} U_{m-1})^T + c^T X_m X_{m+1}^\dagger \mathcal{A}_m] X_m d \\
&= d^T U_m^T X_m d + c^T X_m X_{m+1}^\dagger (X_{m+1} - \mathcal{B}_m U_m)\, d \\
&= d^T U_m^T X_m d + c^T X_m d - c^T P_m U_m d = c^T P_m c < 0.
\end{aligned}$$

This is again a contradiction and the first part of the proof is complete.

Now suppose that the kernel condition and the P-condition hold and let (x,u) be admissible with $x_0 = X_0\, d$, $x_{N+1} = 0$, $x \not\equiv 0$. Then, by Lemma 1.30, we have $x_k \in \operatorname{Im} X_k$ for all $k \in [0, N+1]$, and from Theorem 2.9 (Picone identity) we get

$$\mathcal{F}_0(x,u) = x_k^T Q_k x_k \big|_0^{N+1} + \sum_{k=0}^{N} w_k^T \mathcal{P}_k w_k = -d^T X_0^T U_0\, d + \sum_{k=0}^{N} w_k^T \mathcal{P}_k w_k, \quad (3.6)$$

where $\mathcal{P}_k = \mathcal{B}_k^T \mathcal{D}_k - \mathcal{B}_k^T Q_{k+1} \mathcal{B}_k = P_k \geq 0$, by Lemma 2.5, and where Q_k is defined by (2.2). This implies $\mathcal{F}_0(x,u) + d^T X_0^T U_0\, d \geq 0$. It remains to show that $\mathcal{F}_0(x,u) + d^T X_0^T U_0\, d \neq 0$. Suppose that $\mathcal{F}_0(x,u) + d^T X_0^T U_0\, d = 0$. Then the nonnegativity of $\mathcal{P}$ and equality (3.6) impliy that $w_k^T \mathcal{P}_k w_k = 0$, i.e. $\mathcal{P}_k w_k = 0$ for all $k \in [0, N]$. From identity (2.12) in Lemma 2.8 we get $(\mathcal{D}_k^T - \mathcal{B}_k^T Q_{k+1})\, x_{k+1} = x_k$. This together with $x_{N+1} = 0$ imply that $x \equiv 0$, which is a contradiction. $\square$

Lemma 3.7 (Riccati equivalence). *The following statements are equivalent.*

(iv) *There exists a conjoined basis (X, U) of* (S) *such that X_k is invertible for all $k \in [0, N+1]$, and $P_k \geq 0$ on $[0, N]$.*

(v) *There exists a symmetric solution Q_k on $[0, N+1]$ of the Riccati equation* (3.1) *such that condition* (3.2) *holds.*

Proof. Let (X, U) be conjoined basis with X_k invertible for all $k \in [0, N+1]$. We put $Q = UX^{-1}$ and show that it has all the properties in (v). The symmetry of Q follows from the symmetry of $X^T U$ and the identity $UX^{-1} = X^{-1T} X^T U X^{-1}$. Further,

$$R[Q]_k = U_{k+1} X_{k+1}^{-1} (\mathcal{A}_k X_k + \mathcal{B}_k U_k) X_k^{-1} - (\mathcal{C}_k X_k + \mathcal{D}_k U_k) X_k^{-1} = 0,$$

$$\mathcal{A}_k + \mathcal{B}_k Q_k = (\mathcal{A}_k X_k + \mathcal{B}_k U_k) X_k^{-1} = X_{k+1} X_k^{-1},$$

thus the matrix $\mathcal{A}_k + \mathcal{B}_k Q_k$ is invertible. The identity

$$(\mathcal{A}_k + \mathcal{B}_k Q_k)^{-1} \mathcal{B}_k = X_k X_{k+1}^{-1} \mathcal{B}_k = P_k \tag{3.7}$$

implies that $(\mathcal{A}_k + \mathcal{B}_k Q_k)^{-1} \mathcal{B}_k \geq 0$.

Conversely, let Q be a symmetric solution of the Riccati equation such that condition (3.2) holds. We define X as the solution of the linear equation $X_{k+1} = (\mathcal{A}_k + \mathcal{B}_k Q_k) X_k$, $k \in [0, N]$, with $X_0 = I$, and define $U := QX$. Then $X^T U = X^T Q X$ is symmetric and we have

$$\begin{aligned} X_{k+1} &= \mathcal{A}_k X_k + \mathcal{B}_k U_k, \\ U_{k+1} &= Q_{k+1} X_{k+1} = Q_{k+1} (\mathcal{A}_k + \mathcal{B}_k Q_k) X_k = \mathcal{C}_k X_k + \mathcal{D}_k U_k, \end{aligned}$$

thus (X, U) is conjoined basis of (S). The invertibility of X follows from the invertibility of $\mathcal{A}_k + \mathcal{B}_k Q_k$, and $P_k \geq 0$ follows from identity (3.7). The proof is complete. □

Lemma 3.8. *The following statements are equivalent.*

(vi) *System* (3.3) *has a solution* (X, U) *such that* $X_k^T U_k$ *is symmetric and* X_k *is invertible for all* $k \in [0, N+1]$, *and* $P_k \geq 0$ *on* $[0, N]$.

(vii) *Discrete Riccati inequality* (3.4) *has a symmetric solution* Q_k *on* $[0, N+1]$ *such that condition* (3.2) *holds.*

Proof. The proof is similar to the previous one. Let (X, U) be a solution of system (3.3) with X_k invertible and $X_k^T U_k$ symmetric for all $k \in [0, N+1]$. We put $Q = UX^{-1}$ again and the only thing different from the previous proof is the inequality. We have

$$\begin{aligned} R[Q]_k (\mathcal{A}_k + \mathcal{B}_k Q_k)^{-1} &= U_{k+1} X_{k+1}^{-1} - (\mathcal{C}_k X_k + \mathcal{D}_k U_k)(\mathcal{A}_k X_k + \mathcal{B}_k U_k)^{-1} \\ &= (U_{k+1} - \mathcal{C}_k X_k - \mathcal{D}_k U_k) X_{k+1}^{-1} = X_{k+1}^{-1T} N_k X_{k+1}^{-1} \leq 0. \end{aligned}$$

Conversely, let Q be a symmetric solution of the Riccati inequality (3.4) such that condition (3.2) holds. We again define X as the solution of linear equation $X_{k+1} = (\mathcal{A}_k + \mathcal{B}_k Q_k)X_k, k \in [0, N]$ with $X_0 = I$, and define $U := QX$. Now the only difference is the inequality in system (3.3). We have

$$\begin{aligned} N_k &= X_{k+1}^T(U_{k+1} - \mathcal{C}_k X_k - \mathcal{D}_k U_k) = X_{k+1}^T[Q_{k+1}X_{k+1} - (\mathcal{C}_k + \mathcal{D}_k Q_k)X_k] \\ &= X_{k+1}^T[Q_{k+1}(\mathcal{A}_k + \mathcal{B}_k Q_k) - (\mathcal{C}_k + \mathcal{D}_k Q_k)]X_k \\ &= X_{k+1}^T R[Q]_k(\mathcal{A}_k + \mathcal{B}_k Q_k)^{-1}X_{k+1} \le 0. \end{aligned}$$

Thus, this lemma is proven. □

Proof of Theorem 3.4. Equivalence (i) ⇔ (ii) is Lemma 3.5, equivalence (iv) ⇔ (v) is Lemma 3.7, and equivalence (vi) ⇔ (vii) is Lemma 3.8. We first prove implications (ii) ⇒ (iii) ⇒ (i) and then (iv) ⇒ (i) and (vi) ⇒ (i). Statement (v) implies (vii) trivially. Finally we prove that statement (i) implies (iv).

(ii) ⇒ (iii): Let Q be defined by (2.2) or (2.3) by the principal solution $(\hat{X}, \hat{U})$. Further, let $\underline{u} \in \mathbb{R}^{(N+1)n}$ and define $x_k := G_k \mathcal{T}_k \underline{u}$, $k \in [1, N+1]$ and $x_0 = 0$. Such a pair (x, u) is admissible, by Lemma 1.29. Thus, by Lemma 1.30, there exists $c_k \in \mathbb{R}^n$ such that $x_k = \hat{X}_k c_k$ for all $k \in [0, N+1]$. Then $R[Q]_k G_k \mathcal{T}_k \underline{u} = R[Q]_k \hat{X}_k c_k = 0$, where we used Lemma 2.2.

(iii) ⇒ (i): Let (x, u) be admissible with $x_0 = x_{N+1} = 0$, $x \not\equiv 0$. Lemma 1.35 yields that $x_k = G_k \mathcal{T}_k \underline{u}$ for $k \in [0, N+1]$, and thus, by condition (iii), there exists Q such that $0 = R[Q]_k G_k \mathcal{T}_k \underline{u} = R[Q]_k x_k$. Then, by Lemma 2.7, we have $\mathcal{F}_0(x, u) = \sum_{k=0}^N w_k^T \mathcal{P}_k w_k \ge 0$. Now, if $\mathcal{F}_0(x, u) = 0$, then again as in the proof of Lemma 3.5, the equation $(\mathcal{D}_k^T - \mathcal{B}_k^T Q_{k+1})\, x_{k+1} = x_k$ holds for all $k \in [0, N]$ and thus, $x \equiv 0$.

(iv) ⇒ (i): The invertibility of X implies the kernel condition, thus (X, U) has no focal points in $(0, N+1]$ and statement (i) follows from Lemma 3.6.

(vi) ⇒ (i): We put $F_k := X_{k+1}^{-1T} N_k X_{k+1}^{-1} \le 0$ for $k \in [0, N]$ and define

$$\underline{\mathcal{A}}_k := \mathcal{A}_k, \quad \underline{\mathcal{B}}_k := \mathcal{B}_k, \quad \underline{\mathcal{C}}_k := \mathcal{C}_k + F_k \mathcal{A}_k, \quad \underline{\mathcal{D}}_k := \mathcal{D}_k + F_k \mathcal{B}_k,$$

and

$$\underline{\mathcal{S}}_k := \begin{pmatrix} \underline{\mathcal{A}}_k & \underline{\mathcal{B}}_k \\ \underline{\mathcal{C}}_k & \underline{\mathcal{D}}_k \end{pmatrix} = \mathcal{S}_k + \mathcal{R}_k \quad \text{with } \mathcal{R}_k := \begin{pmatrix} 0 & 0 \\ F_k \mathcal{A}_k & F_k \mathcal{B}_k \end{pmatrix}. \tag{3.8}$$

The matrix $\underline{\mathcal{S}}_k$ is symplectic, by Lemma 1.10, and it defines symplectic system denoted here by ($\underline{\mathrm{S}}$). The pair (X, U) solves the system ($\underline{\mathrm{S}}$), which follows from the calculations

$$\begin{aligned}\underline{\mathcal{A}}_k X_k + \underline{\mathcal{B}}_k U_k &= \mathcal{A}_k X_k + \mathcal{B}_k U_k = X_{k+1},\\ \underline{\mathcal{C}}_k X_k + \underline{\mathcal{D}}_k U_k &= \mathcal{C}_k X_k + \mathcal{D}_k U_k + F_k(\mathcal{A}_k X_k + \mathcal{B}_k U_k)\\ &= \mathcal{C}_k X_k + \mathcal{D}_k U_k + X_{k+1}^{-1T} N_k = U_{k+1}.\end{aligned}$$

Moreover, $X_k^T U_k$ is symmetric, X_k is invertible for all $k \in [0, N+1]$, and $\underline{P}_k = P_k \geq 0$ on $[0, N]$, hence condition (iv) holds for ($\underline{\text{S}}$). We already proved (iv) $\Rightarrow$ (i), and hence the functional $\underline{\mathcal{F}}_0(x,u) > 0$ over $x_0 = 0 = x_{N+1}$, $x \not\equiv 0$. Furthermore, the symmetric matrix $\underline{\mathcal{E}}_k := \mathcal{E}_k + F_k$ satisfies

$$\underline{\mathcal{D}}_k^T \underline{\mathcal{B}}_k = \mathcal{D}_k^T \mathcal{B}_k + \mathcal{B}_k^T F_k \mathcal{B}_k = \mathcal{B}_k^T (\mathcal{E}_k + F_k)\, \mathcal{B}_k = \underline{\mathcal{B}}_k^T \underline{\mathcal{E}}_k \underline{\mathcal{B}}_k$$

and we have

$$\begin{aligned}&\begin{pmatrix} \mathcal{A}_k^T (\mathcal{E}_k + F_k)\, \mathcal{A}_k - \mathcal{A}_k^T(\mathcal{C}_k + F_k \mathcal{A}_k) & \mathcal{C}_k^T + \mathcal{A}_k^T F_k - \mathcal{A}_k^T (\mathcal{E}_k + F_k) \\ \mathcal{C}_k + F_k \mathcal{A}_k - (\mathcal{E}_k + F_k)\, \mathcal{A}_k & \mathcal{E}_k + F_k \end{pmatrix}\\ &\quad - \begin{pmatrix} \mathcal{A}_k^T \mathcal{E}_k \mathcal{A}_k - \mathcal{A}_k^T \mathcal{C}_k & \mathcal{C}_k^T - \mathcal{A}_k^T \mathcal{E}_k \\ \mathcal{C}_k - \mathcal{E}_k \mathcal{A}_k & \mathcal{E}_k \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & F_k \end{pmatrix} \leq 0.\end{aligned}$$

Thus, Lemma 1.32 yields $\underline{\mathcal{F}}_0(x,u) \leq \mathcal{F}_0(x,u)$.

(i) $\Rightarrow$ (iv): We take the conjoined basis (X, U) with $X_0 = I$ and $U_0 = I + \tilde{X}_{N+1}^T \tilde{U}_{N+1} - \tilde{X}_{N+1}^T \hat{X}_{N+1} \hat{X}_{N+1}^\dagger \hat{U}_{N+1} \hat{X}_{N+1}^\dagger \tilde{X}_{N+1}$ and show that it has no focal points. Then the kernel condition implies that X_k is invertible for all $k \in [0, N+1]$. We use Lemma 3.6 and show that inequality in statement (i') holds for all (x,u) admissible with $x_{N+1} = 0$ and $x_0 = d \neq 0$. (For $x_0 = 0$ it holds by (i).) Since the kernel condition holds for the principal solution by the implication (i) $\Rightarrow$ (ii), identity (2.39) yields that

$$\mathcal{F}_0(x,u) + x_0^T U_0 x_0 = x_0^T x_0 + \mathcal{F}_0(\bar{x}, \bar{u}), \tag{3.9}$$

where $(\bar{x}, \bar{u})$ is defined by (2.38), and $\bar{x}_0 = 0 = \bar{x}_{N+1}$, by Lemma 2.30 and Lemma 2.35. Thus, $\mathcal{F}_0(\bar{x}, \bar{u}) \geq 0$ by (i) (equality holds iff $\bar{x} \equiv 0$). Further, since $x_0 \neq 0$, the inequality in statement (i') of Lemma 3.6 holds by (3.9). The proof is complete. □

Remark 3.9. An alternative proof of implication (i) $\Rightarrow$ (iv) uses the results for functionals with separated endpoints (Subsection 3.1.2) and for perturbed functionals (Section 3.4).

3.1.2 Functional with separated endpoints

In this subsection, the functional $\mathcal{F}$ has separated endpoints, i.e. it is given by formula (1.37). A roundabout theorem for the functional with separated endpoints, Theorem 3.14, is analogical to the presented roundabout theorem for the functional with zero endpoints, Theorem 3.4. The difference is that we take the solution (X, U) of (S) with $X_0 = I - \mathcal{M}_0$, and $U_0 = \Gamma_0 + \mathcal{M}_0$ instead of the principal solution $(\hat{X}, \hat{U})$. This solution is called the *natural conjoined basis* of (S).

Remark 3.10. Note that when the initial endpoint is zero, i.e. when $\mathcal{M}_0 = I$, then the natural conjoined basis reduces to the principal solution.

The next difference is that statement (i) in Theorem 3.14 below now says that the functional is positive on a larger set of admissible pairs, which means that also in other statements we have to add more conditions. These are the final and the initial endpoint constraints,

$$X_{N+1}^T(\Gamma_1 X_{N+1} + U_{N+1}) > 0 \quad \text{on } \operatorname{Ker} \mathcal{M}_1 X_{N+1} \setminus \operatorname{Ker} X_{N+1} \tag{3.10}$$
$$\Gamma_1 + Q_{N+1} > 0 \quad \text{on } \operatorname{Ker} \mathcal{M}_1 \cap \operatorname{Im} X_{N+1}, \tag{3.11}$$

$$X_{N+1}^T(\Gamma_1 X_{N+1} + U_{N+1}) > 0 \quad \text{on } \operatorname{Ker} \mathcal{M}_1 X_{N+1}, \tag{3.12}$$
$$\Gamma_1 + Q_{N+1} > 0 \quad \text{on } \operatorname{Ker} \mathcal{M}_1, \tag{3.13}$$

$$X_0^T(\Gamma_0 X_0 - U_0) > 0 \quad \text{on } \operatorname{Ker} \mathcal{M}_0 X_0, \tag{3.14}$$
$$\Gamma_0 - Q_0 > 0 \quad \text{on } \operatorname{Ker} \mathcal{M}_0. \tag{3.15}$$

Remark 3.11. Inequalities (3.10), (3.11) are used with the natural conjoined basis (or the principal solution) which may be singular at $k = N+1$. On the other hand, conditions (3.12)–(3.15) are used for a conjoined basis (X, U) with X_k invertible on $[0, N+1]$.

Remark 3.12. Condition (3.10) is equivalent to the condition

$$X_{N+1}^T(\Gamma_1 X_{N+1} + U_{N+1}) \geq 0 \quad \text{on } \operatorname{Ker} \mathcal{M}_1 X_{N+1},$$
$$\operatorname{Ker} (I - \mathcal{M}_1)(\Gamma_1 X_{N+1} + U_{N+1}) \cap \operatorname{Ker} \mathcal{M}_1 X_{N+1} \subseteq \operatorname{Ker} X_{N+1}.$$

Remark 3.13. If identity (2.4) holds for $X_{N+1}, U_{N+1}, Q_{N+1}$ in place of X, U, Q, then condition (3.10) is equivalent to condition (3.11). If moreover the matrix X_{N+1} is invertible, i.e. $Q_{N+1} = U_{N+1} X_{N+1}^{-1}$, then condition (3.12) is equivalent to condition (3.13). Similarly, if the matrix X_0 is invertible and $Q_0 = U_0 X_0^{-1}$, then condition (3.14) is equivalent to condition (3.15).

Theorem 3.14. *The following statements are equivalent.*

(i) $\mathcal{F}(x,u) > 0$ *over* $\mathcal{M}_0\, x_0 = 0$, $\mathcal{M}_1 x_{N+1} = 0$, *and* $x \not\equiv 0$.

(ii) *The natural conjoined basis* (X, U) *of* (S) *has no focal points in* $(0, N+1]$ *and satisfies* (3.10).

(iii) *The implicit Riccati equation*

$$R[Q]_k \left(\Phi_{k,0}(I - \mathcal{M}_0) \quad G_k \right) = 0, \quad k \in [0, N],$$

has a symmetric solution Q_k *on* $[0, N+1]$ *such that* $Q_0 = \Gamma_0$, *and* $\mathcal{P}_k = \mathcal{B}_k^T \mathcal{D}_k - \mathcal{B}_k^T Q_{k+1} \mathcal{B}_k \geq 0$ *for all* $k \in [0, N]$, *and satisfying* (3.11).

(iv) *There exists a conjoined basis* (X, U) *of* (S) *such that* X_k *is invertible for all* $k \in [0, N+1]$, *and* $P_k = X_k X_{k+1}^{-1} \mathcal{B}_k \geq 0$ *on* $[0, N]$, *and satisfying* (3.12) *and* (3.14).

(v) *There exists a symmetric solution* Q_k *on* $[0, N+1]$ *of explicit Riccati equation* (3.1) *such that condition* (3.2) *holds, and satisfying* (3.13) *and* (3.15).

(vi) *The system*

$$\begin{aligned} X_{k+1} &= \mathcal{A}_k X_k + \mathcal{B}_k U_k, \\ N_k := X_{k+1}^T (U_{k+1} - \mathcal{C}_k X_k - \mathcal{D}_k U_k) &\leq 0, \quad k \in [0, N], \end{aligned}$$

has a solution (X, U) *such that* $X_k^T U_k$ *is symmetric and* X_k *is invertible for all* $k \in [0, N+1]$, $P_k = X_k X_{k+1}^{-1} \mathcal{B}_k \geq 0$ *on* $[0, N]$, *and satisfying* (3.12) *and* (3.14).

(vii) *The Riccati inequality* (3.4) *has a symmetric solution* Q_k *on* $[0, N+1]$ *satisfying* (3.2), (3.13) *and* (3.15).

Remark 3.15. Note that, by Remark 3.10, Theorem 3.4 is a corollary of Theorem 3.14. But as we use it in the proof of Theorem 3.14, we had to prove it first.

First we prove the equivalence of (i) and (ii), stated separately in the next lemma. In the proof we again use Lemma 3.6 and the Picone identity.

Lemma 3.16. *The following statements are equivalent.*

(i) $\mathcal{F}(x,u) > 0$ *over* $\mathcal{M}_0\, x_0 = 0$, $\mathcal{M}_1 x_{N+1} = 0$, *and* $x \not\equiv 0$.

(ii) *The natural conjoined basis* (X,U) *of* (S) *has no focal points in* $(0, N+1]$ *and satisfies* (3.10).

Proof. First let (i) hold. Then, from Lemma 3.6 for the natural conjoined basis (X,U) and $\alpha = x_0$, we have that (X,U) has no focal points in $(0, N+1]$. It remains to show (3.10). Let $d \in \operatorname{Ker} \mathcal{M}_1 X_{N+1}$, $X_{N+1} d \neq 0$ and take admissible $(X\,d, U\,d)$. Then $\mathcal{M}_0\, X_0\, d = 0$, $\mathcal{M}_1 X_{N+1} d = 0$, $X d \not\equiv 0$, and thus $\mathcal{F}(X\,d, U\,d) > 0$. By Lemma 1.31, we have

$$\begin{aligned}\mathcal{F}(X\,d, U\,d) &= d^T \Gamma_0\, d + d^T X_{N+1}^T \Gamma_1 X_{N+1} d + d^T X_{N+1}^T U_{N+1} d - d^T \Gamma_0\, d \\ &= d^T (X_{N+1}^T \Gamma_1 X_{N+1} + X_{N+1}^T U_{N+1})\, d,\end{aligned}$$

and we get that the inequality $d^T (X_{N+1}^T \Gamma_1 X_{N+1} + X_{N+1}^T U_{N+1})\, d > 0$ holds.

Now suppose that (ii) holds and let (x,u) be admissible with $\mathcal{M}_0\, x_0 = 0$, $\mathcal{M}_1 x_{N+1} = 0$, and $x \not\equiv 0$. Then, by Lemma 1.30, $x_k \in \operatorname{Im} X_k$ for all $k \in [0, N+1]$, and from Theorem 2.9 (Picone identity) we get

$$\begin{aligned}\mathcal{F}(x,u) &= x_0^T \Gamma_0\, x_0 + x_{N+1}^T \Gamma_1 x_{N+1} + x_k^T Q_k x_k \big|_0^{N+1} + \sum_{k=0}^{N} w_k^T \mathcal{P}_k w_k \\ &= x_{N+1}^T \Gamma_1 x_{N+1} + x_{N+1}^T Q_{N+1} x_{N+1} + \sum_{k=0}^{N} w_k^T \mathcal{P}_k w_k \\ &= d^T (X_{N+1}^T \Gamma_1 X_{N+1} + X_{N+1}^T U_{N+1})\, d + \sum_{k=0}^{N} w_k^T \mathcal{P}_k w_k, \qquad (3.16)\end{aligned}$$

where d is such that $X_{N+1} d = x_{N+1}$ and $\mathcal{P}_k = \mathcal{B}_k^T \mathcal{D}_k - \mathcal{B}_k^T Q_{k+1} \mathcal{B}_k = P_k \geq 0$, by Lemma 2.5. This together with (3.10) imply $\mathcal{F}(x,u) > 0$. Case when $x_{N+1} = 0$ is shown in Lemma 3.6. The proof is complete. □

The proof of some parts of this roundabout theorem is based on transforming the system (S) into a system with zero endpoints with the same value of the functional $\mathcal{F}$.

We define $n \times n$ matrices $\tilde{\mathcal{A}}_k, \tilde{\mathcal{B}}_k, \tilde{\mathcal{C}}_k, \tilde{\mathcal{D}}_k$ on $[-1, N+1]$ by

$$
\begin{gathered}
\tilde{\mathcal{A}}_k := \begin{cases} \mathcal{M}_0, & \text{for } k = -1, \\ \mathcal{A}_k, & \text{for } k \in [0, N], \\ I, & \text{for } k = N+1, \end{cases} \\
\tilde{\mathcal{B}}_k := \begin{cases} I - \mathcal{M}_0, & \text{for } k = -1, \\ \mathcal{B}_k, & \text{for } k \in [0, N], \\ I - \mathcal{M}_1, & \text{for } k = N+1, \end{cases} \\
\tilde{\mathcal{C}}_k := \begin{cases} \mathcal{M}_0 - I, & \text{for } k = -1, \\ \mathcal{C}_k, & \text{for } k \in [0, N-1], \\ \mathcal{C}_N + [\Gamma_1 - (I - \mathcal{M}_1)]\mathcal{A}_N, & \text{for } k = N, \\ 0, & \text{for } k = N+1, \end{cases} \\
\tilde{\mathcal{D}}_k := \begin{cases} \Gamma_0 + \mathcal{M}_0, & \text{for } k = -1, \\ \mathcal{D}_k, & \text{for } k \in [0, N-1], \\ \mathcal{D}_N + [\Gamma_1 - (I - \mathcal{M}_1)]\mathcal{B}_N, & \text{for } k = N, \\ I, & \text{for } k = N+1. \end{cases}
\end{gathered}
\tag{3.17}
$$

Then the matrix $\tilde{\mathcal{S}}_k := \left(\begin{smallmatrix} \tilde{\mathcal{A}}_k & \tilde{\mathcal{B}}_k \\ \tilde{\mathcal{C}}_k & \tilde{\mathcal{D}}_k \end{smallmatrix}\right)$ with coefficients $\tilde{\mathcal{A}}_k, \tilde{\mathcal{B}}_k, \tilde{\mathcal{C}}_k, \tilde{\mathcal{D}}_k$ defined by (3.17) is symplectic for $k \in [-1, N+1]$, and the system

$$
\begin{pmatrix} X_{k+1} \\ U_{k+1} \end{pmatrix} = \tilde{\mathcal{S}}_k \begin{pmatrix} X_k \\ U_k \end{pmatrix}, \quad k \in [-1, N+1], \tag{$\tilde{\mathrm{S}}$}
$$

is a symplectic system.

Remark 3.17. Another way how to transform the variable endpoint at $k = 0$ into a zero endpoint at $k = -1$ is used for the Hamiltonian case in [17] and for the symplectic case in [40]. The only difference is in the coefficients $\tilde{\mathcal{A}}_{-1}$ and $\tilde{\mathcal{C}}_{-1}$. Namely, in [40], $\tilde{\mathcal{A}}_{-1} := [\Gamma_0 + \mathcal{M}_0 - \varepsilon(I - \mathcal{M}_0)]^{-1}$ and $\tilde{\mathcal{C}}_{-1} := \varepsilon[\Gamma_0 + \mathcal{M}_0 - \varepsilon(I - \mathcal{M}_0)]^{-1}$. Then $\mathcal{A}_{-1}$ is invertible, which was crucial for the Hamiltonian case in [17]. But, as it is known, for the symplectic case the invertibility of $\mathcal{A}_{-1}$ is not necessary.

Lemma 3.18. *Let (x, u) be an admissible pair w.r.t. $(\mathcal{A}, \mathcal{B})$ on $[0, N+1]$.*

Then the pair $(\tilde{x}, \tilde{u})$ *defined by*

$$(\tilde{x}, \tilde{u}) := \begin{cases} (\mathcal{M}_0 x_0, x_0), & \text{for } k = -1, \\ (x_k, u_k), & \text{for } k \in [0, N], \\ (x_{N+1}, -x_{N+1}), & \text{for } k = N+1, \\ (\mathcal{M}_1 x_{N+1}, \gamma), & \text{for } k = N+2, \end{cases} \tag{3.18}$$

where $\gamma \in \mathbb{R}^n$ *is arbitrary, is admissible w.r.t.* $(\tilde{\mathcal{A}}, \tilde{\mathcal{B}})$ *on* $[-1, N+2]$ *and*

$$\mathcal{F}(x, u) = \tilde{\mathcal{F}}_0(\tilde{x}, \tilde{u}) := \sum_{k=-1}^{N+1} \left\{ \tilde{x}_k^T \tilde{\mathcal{A}}_k^T \tilde{\mathcal{C}}_k \tilde{x}_k + 2\, \tilde{x}_k^T \tilde{\mathcal{C}}_k^T \tilde{\mathcal{B}}_k \tilde{u}_k + \tilde{u}_k^T \tilde{\mathcal{B}}_k^T \tilde{\mathcal{D}}_k \tilde{u}_k \right\}. \tag{3.19}$$

Proof. A pair (x, u) is admissible w.r.t. $(\mathcal{A}, \mathcal{B})$ on $[0, N+1]$ if and only if the pair $(\tilde{x}, \tilde{u})$ is admissible w.r.t. $(\tilde{\mathcal{A}}, \tilde{\mathcal{B}})$ on $[0, N+1]$. Further, we have

$$\tilde{\mathcal{A}}_{-1} \tilde{x}_{-1} + \tilde{\mathcal{B}}_{-1} \tilde{u}_{-1} = x_0 = \tilde{x}_0,$$

and

$$\tilde{\mathcal{A}}_{N+1} \tilde{x}_{N+1} + \tilde{\mathcal{B}}_{N+1} \tilde{u}_{N+1} = \mathcal{M}_1 x_{N+1} = \tilde{x}_{N+2}.$$

Finally, we have

$$\begin{aligned} \tilde{\mathcal{F}}_0(\tilde{x}, \tilde{u}) &= x_0^T (\Gamma_0 + \mathcal{M}_0)(I - \mathcal{M}_0)\, x_0 + \mathcal{F}_0(x, u) \\ &\quad + x_N^T [\Gamma_1 - (I - \mathcal{M}_1)\mathcal{A}_N]^T \mathcal{A}_N x_N + 2 x_N^T [\Gamma_1 - (I - \mathcal{M}_1)\mathcal{A}_N]^T \mathcal{B}_N u_N \\ &\quad + u_N^T [\Gamma_1 - (I - \mathcal{M}_1)\mathcal{B}_N]^T \mathcal{B}_N u_N + x_{N+1}^T (I - \mathcal{M}_1)\, x_{N+1} \\ &= x_0^T \Gamma_0 x_0 + \mathcal{F}_0(x, u) + x_{N+1}^T \Gamma_1 x_{N+1} = \mathcal{F}(x, u). \end{aligned}$$

Thus, this lemma is proven. ☐

Lemma 3.19. *Let* $(\tilde{x}, \tilde{u})$ *be an admissible pair w.r.t.* $(\tilde{\mathcal{A}}, \tilde{\mathcal{B}})$ *on* $[-1, N+2]$ *with* $\tilde{x}_{-1} = 0 = \tilde{x}_{N+2}$*. Then the pair* (x, u) *defined by*

$$(x, u) := \begin{cases} (\tilde{x}_k, \tilde{u}_k), & \text{for } k \in [0, N], \\ (\tilde{x}_{N+1}, \gamma), & \text{for } k = N+1, \end{cases} \tag{3.20}$$

where $\gamma \in \mathbb{R}^n$ *is arbitrary, is admissible w.r.t.* $(\mathcal{A}, \mathcal{B})$ *on* $[0, N+1]$ *with* $\mathcal{M}_0 x_0 = 0 = \mathcal{M}_1 x_{N+1}$*, and equation* (3.19) *holds.*

Proof. A pair $(\tilde{x}, \tilde{u})$ is admissible w.r.t. $(\tilde{\mathcal{A}}, \tilde{\mathcal{B}})$ on $[0, N+1]$ if and only if the pair (x, u) is admissible w.r.t. $(\mathcal{A}, \mathcal{B})$ on $[0, N+1]$. Further, we have

$$x_0 = \tilde{x}_0 = \tilde{\mathcal{A}}_{-1}\tilde{x}_{-1} + \tilde{\mathcal{B}}_{-1}\tilde{u}_{-1} = (I - \mathcal{M}_0)\tilde{u}_{-1},$$

and

$$0 = \tilde{x}_{N+2} = \tilde{\mathcal{A}}_{N+1}\tilde{x}_{N+1} + \tilde{\mathcal{B}}_{N+1}\tilde{u}_{N+1} = x_{N+1} + (I - \mathcal{M}_1)\tilde{u}_{N+1},$$

thus, $\mathcal{M}_0 x_0 = 0 = \mathcal{M}_1 x_{N+1}$. Finally, we have

$$\begin{aligned}
\tilde{\mathcal{F}}_0(\tilde{x}, \tilde{u}) &= \tilde{u}_{-1}^T(\Gamma_0 + \mathcal{M}_0)(I - \mathcal{M}_0)\,\tilde{u}_{-1} + \mathcal{F}_0(x, u) \\
&\quad + x_N^T[\Gamma_1 - (I - \mathcal{M}_1)\mathcal{A}_N]^T \mathcal{A}_N x_N + 2x_N^T[\Gamma_1 - (I - \mathcal{M}_1)\mathcal{A}_N]^T \mathcal{B}_N u_N \\
&\quad + u_N^T[\Gamma_1 - (I - \mathcal{M}_1)\mathcal{B}_N]^T \mathcal{B}_N u_N + x_{N+1}^T(I - \mathcal{M}_1)\,x_{N+1} \\
&= x_0^T \Gamma_0 x_0 + \mathcal{F}_0(x, u) + x_{N+1}^T \Gamma_1 x_{N+1} = \mathcal{F}(x, u).
\end{aligned}$$

Thus, this lemma is proven. □

Lemma 3.20. *The functional $\mathcal{F}$ is nonnegative over $\mathcal{M}_0 x_0 = 0 = \mathcal{M}_1 x_{N+1}$ if and only if the functional $\tilde{\mathcal{F}}_0$ defined in equation* (3.19) *is nonnegative over $\tilde{x}_{-1} = 0 = \tilde{x}_{N+2}$.*

Proof. First we show the implication to the right. Let $(\tilde{x}, \tilde{u})$ be admissible w.r.t. $(\tilde{\mathcal{A}}, \tilde{\mathcal{B}})$ on $[-1, N+2]$ with $\tilde{x}_{-1} = 0 = \tilde{x}_{N+2}$. Then, by Lemma 3.19, the pair (x, u) defined by (3.20) is admissible w.r.t. $(\mathcal{A}, \mathcal{B})$ with $\mathcal{M}_0 x_0 = 0 = \mathcal{M}_1 x_{N+1}$. Thus, $\mathcal{F}(x, u) \geq 0$, and hence, by Lemma 3.19, $\tilde{\mathcal{F}}_0(\tilde{x}, \tilde{u}) = \mathcal{F}(x, u) \geq 0$.

Conversely, let (x, u) be admissible w.r.t. $(\mathcal{A}, \mathcal{B})$ on $[0, N+1]$ with $\mathcal{M}_0 x_0 = 0 = \mathcal{M}_1 x_{N+1}$. Then the pair $(\tilde{x}, \tilde{u})$ defined by (3.18) is admissible w.r.t. $(\tilde{\mathcal{A}}, \tilde{\mathcal{B}})$ on $[-1, N+2]$ with $\tilde{x}_{-1}\mathcal{M}_0 x_0 = 0$, $\tilde{x}_{N+2} = \mathcal{M}_1 x_{N+1} = 0$, thus, $\tilde{\mathcal{F}}_0(\tilde{x}, \tilde{u}) \geq 0$, and hence, by Lemma 3.18, $\mathcal{F}(x, u) = \tilde{\mathcal{F}}_0(\tilde{x}, \tilde{u}) \geq 0$. □

Lemma 3.21. *The functional $\mathcal{F}$ is positive over $\mathcal{M}_0 x_0 = 0 = \mathcal{M}_1 x_{N+1}$ and $x \not\equiv 0$ if and only if the functional $\tilde{\mathcal{F}}_0$ defined in equation* (3.19) *is positive over $\tilde{x}_{-1} = 0 = \tilde{x}_{N+2}$ and $\tilde{x} \not\equiv 0$.*

Proof. The proof is same as the proof of Lemma 3.20, with ">" instead of "≥", since for all pairs (x, u) and $(\tilde{x}, \tilde{u})$ with $\tilde{x}_{-1} = 0 = \tilde{x}_{N+2}$, and $x_k = \tilde{x}_k$, $k \in [0, N+1]$ we have $x \equiv 0$ if and only if $\tilde{x} \equiv 0$. □

In the following two lemmas we prove that statement (iv) in Theorem 3.14 is equivalent to statement (iv) in Theorem 3.4 with system ($\tilde{\text{S}}$) in place of (S) and that statement (vi) in Theorem 3.14 implies statement (vi) in Theorem 3.4 with system ($\tilde{\text{S}}$) in place of (S).

Lemma 3.22. *There exists a conjoined basis (X, U) of* (S) *such that X_k is invertible for all $k \in [0, N+1]$ and $X_k X_{k+1}^{-1} \mathcal{B}_k \geq 0$ on $[0, N]$, and satisfying* (3.12) *and* (3.14) *if and only if there exists a conjoined basis $(\tilde{X}, \tilde{U})$ of* ($\tilde{\text{S}}$) *such that $\tilde{X}_k$ is invertible for all $k \in [-1, N+2]$, and $\tilde{X}_k \tilde{X}_{k+1}^{-1} \tilde{\mathcal{B}}_k \geq 0$ on $[0, N]$.*

Proof. We can assume that $U_k = \tilde{U}_k$ for $k \in [0, N]$ and $X_k = \tilde{X}_k$ for $k \in [0, N+1]$. Then X_k and $\tilde{X}_k$ are invertible on $[0, N+1]$. Then it suffices to prove the following equivalences.

(a) $\tilde{X}_{-1} = (\Gamma_0 + \mathcal{M}_0)\, X_0 - (I - \mathcal{M}_0) U_0$ is invertible, and
$\tilde{P}_{-1} = [(\Gamma_0 + \mathcal{M}_0) X_0 - (I - \mathcal{M}_0) U_0] X_0^{-1} (I - \mathcal{M}_0) \geq 0$
if and only if $\Gamma_0 - U_0 X_0^{-1} > 0$ on $\operatorname{Ker} \mathcal{M}_0$,

(b) $\tilde{X}_{N+2} = (\Gamma_1 + \mathcal{M}_1) X_{N+1} + (I - \mathcal{M}_1) U_{N+1}$ is invertible, and
$\tilde{P}_{N+1} = X_{N+1} [(\Gamma_1 + \mathcal{M}_1) X_{N+1} + (I - \mathcal{M}_1) U_{N+1}]^{-1} (I - \mathcal{M}_1) \geq 0$
if and only if $\Gamma_1 + U_{N+1} X_{N+1}^{-1} > 0$ on $\operatorname{Ker} \mathcal{M}_1$.

Both proofs are almost identical, we show here the proof of equivalence (b).

The matrix $\tilde{X}_{N+2}$ is invertible if and only if the matrix $(I - \mathcal{M}_1)(\Gamma_1 + U_{N+1} X_{N+1}^{-1}) + \mathcal{M}_1$ is invertible. And invertibility of this matrix is further equivalent to

$$u^T (\Gamma_1 + U_{N+1} X_{N+1}^{-1}) u \neq 0 \quad \text{for all } u \in \operatorname{Ker} \mathcal{M}_1, u \neq 0.$$

Next, $\tilde{P}_{N+1} = [(I - \mathcal{M}_1)(\Gamma_1 + U_{N+1} X_{N+1}^{-1}) + \mathcal{M}_1]^{-1} (I - \mathcal{M}_1)$, and the nonnegativity of the matrix $\tilde{P}_{N+1}$ is equivalent to the nonnegativity of the matrix

$$\begin{aligned} (I - \mathcal{M}_1)[(I - \mathcal{M}_1)(\Gamma_1 + U_{N+1} X_{N+1}^{-1}) + \mathcal{M}_1]^T& \\ = (I - \mathcal{M}_1)(\Gamma_1 + U_{N+1} X_{N+1}^{-1})^T (I - \mathcal{M}_1)&, \end{aligned}$$

which is nonnegative if and only if $(\Gamma_1 + U_{N+1} X_{N+1}^{-1}) \geq 0$ on $\operatorname{Im}(I - \mathcal{M}_1) = \operatorname{Ker} \mathcal{M}_1$. $\square$

Lemma 3.23. *If the system*

$$\begin{aligned} X_{k+1} &= \mathcal{A}_k X_k + \mathcal{B}_k U_k, \\ X_{k+1}^T(U_{k+1} - \mathcal{C}_k X_k - \mathcal{D}_k U_k) &\leq 0, \quad k \in [0, N], \end{aligned}$$

has a solution (X, U) such that $X_k^T U_k$ is symmetric and X_k is invertible for all $k \in [0, N+1]$, $P_k \geq 0$ on $[0, N]$, and satisfying (3.12) and (3.14), then the system

$$\begin{aligned} \tilde{X}_{k+1} &= \tilde{\mathcal{A}}_k \tilde{X}_k + \tilde{\mathcal{B}}_k \tilde{U}_k, \\ \tilde{X}_{k+1}^T(\tilde{U}_{k+1} - \tilde{\mathcal{C}}_k \tilde{X}_k - \tilde{\mathcal{D}}_k \tilde{U}_k) &\leq 0, \quad k \in [-1, N+1], \end{aligned}$$

has a solution $(\tilde{X}, \tilde{U})$ such that $\tilde{X}_k^T \tilde{U}_k$ is symmetric and $\tilde{X}_k$ is invertible for all $k \in [-1, N+2]$, and $\tilde{P}_k \geq 0$ on $[-1, N+1]$.

Proof. We define $\tilde{X}_k := X_k$ for $k \in [0, N+1]$, $\tilde{U}_k := U_k$ for $k \in [0, N]$ and

$$\begin{aligned} \tilde{X}_{-1} &:= \tilde{\mathcal{D}}_{-1}^T X_0 - \tilde{\mathcal{B}}_{-1}^T U_0 = (\Gamma_0 + \mathcal{M}_0) X_0 - (I - \mathcal{M}_0) U_0, \\ \tilde{U}_{-1} &:= -\tilde{\mathcal{C}}_{-1}^T X_0 + \tilde{\mathcal{A}}_{-1}^T U_0, \\ \tilde{U}_{N+1} &:= U_{N+1} + [\Gamma_1 + (I - \mathcal{M}_1)] X_{N+1}, \\ \tilde{X}_{N+2} &:= \tilde{\mathcal{A}}_{N+1} \tilde{X}_{N+1} + \tilde{\mathcal{B}}_{N+1} \tilde{U}_{N+1} = (\Gamma_1 + \mathcal{M}_1) X_{N+1} + (I - \mathcal{M}_1) U_{N+1}, \\ \tilde{U}_{N+2} &:= \tilde{\mathcal{C}}_{N+1} \tilde{X}_{N+1} + \tilde{\mathcal{D}}_{N+1} \tilde{U}_{N+1} = \tilde{U}_{N+1}. \end{aligned}$$

Then

$$\begin{aligned} &\tilde{X}_0^T(\tilde{U}_0 - \tilde{\mathcal{C}}_{-1} \tilde{X}_{-1} - \tilde{\mathcal{D}}_{-1} \tilde{U}_{-1}) = 0, \\ &\tilde{X}_{N+1}^T(\tilde{U}_{N+1} - \tilde{\mathcal{C}}_N \tilde{X}_N - \tilde{\mathcal{D}}_N \tilde{U}_N) = X_{N+1}^T(U_{N+1} - C_N X_N - D_N U_N) \leq 0, \\ &\tilde{X}_{N+2}^T(\tilde{U}_{N+2} - \tilde{\mathcal{C}}_{N+1} \tilde{X}_{N+1} - \tilde{\mathcal{D}}_{N+1} \tilde{U}_{N+1}) = 0, \end{aligned}$$

and equivalences (a), (b) from previous Lemma 3.22 hold. □

Proof of Theorem 3.14. Equivalence (i) ⇔ (ii) follows from Lemma 3.16, equivalence (i) ⇔ (iv) and implication (vi) ⇒ (i) follow from Theorem 3.4 and Lemmas 3.21, 3.22, 3.23. Implication (v) ⇒ (vii) holds trivially, and equivalences (iv) ⇔ (v) and (vi) ⇔ (vii) follow from Lemma 3.7 and Lemma 3.8, and from equivalences (3.12) ⇔ (3.13), and (3.14) ⇔ (3.15) where X is invertible and $Q = UX^{-1}$. It remains to show implications (ii) ⇒ (iii), and (iii) ⇒ (i).

(ii) $\Rightarrow$ (iii): Let Q be defined by (2.2) or (2.3) by the natural conjoined basis (X, U). Then $Q_0 = \Gamma_0$, $\mathcal{P}_k = P_k \geq 0$ and (3.10) implies (3.11). Furthermore, let $\binom{\alpha}{\underline{u}} \in \mathbb{R}^{(N+1)n}$ and define $x_k := (\Phi_{k,0}(I - \mathcal{M}_0), G_k \mathcal{T}_k) \binom{\alpha}{\underline{u}}$, $k \in [0, N+1]$. Such a pair (x, u) is admissible, by Lemma 1.35, and $\mathcal{M}_0 x_0 = 0$. Thus, by Lemma 1.30, there exists $c_k \in \mathbb{R}^n$ such that $x_k = X_k c_k$ for all $k \in [0, N+1]$. Then

$$R[Q]_k \begin{pmatrix} \Phi_{k,0}(I - \mathcal{M}_0) & G_k \mathcal{T}_k \end{pmatrix} \begin{pmatrix} \alpha \\ \underline{u} \end{pmatrix} = R[Q]_k X_k c_k = 0,$$

where we used Lemma 2.2.

(iii) $\Rightarrow$ (i): Let (x, u) be admissible with $\mathcal{M}_0 x_0 = \mathcal{M}_1 x_{N+1} = 0$, $x \not\equiv 0$. Then Lemma 1.35 yields that $x_k = (\Phi_{k,0}(I - \mathcal{M}_0), G_k \mathcal{T}_k) \binom{x_0}{\underline{u}}$, $k \in [0, N+1]$ and, thus, by (iii), there exists Q such that $0 = R[Q]_k(\Phi_{k,0}(I - \mathcal{M}_0), G_k \mathcal{T}_k) \binom{x_0}{\underline{u}} = R[Q]_k x_k$, $Q_0 = \Gamma_0$, $\mathcal{P}_k \geq 0$ for all $k \in [0, N]$, and (3.11) holds. Then, by identity (2.10) in Lemma 2.7, we have

$$\begin{aligned} \mathcal{F}(x, u) &= x_0^T \Gamma_0 \, x_0 + x_{N+1}^T \Gamma_1 x_{N+1} + x_k^T Q_k x_k \big|_0^{N+1} + \sum_{k=0}^{N} w_k^T \mathcal{P}_k w_k \\ &= x_{N+1}^T (Q_{N+1} + \Gamma_1) \, x_{N+1} + \sum_{k=0}^{N} w_k^T \mathcal{P}_k w_k \geq 0, \end{aligned}$$

and the equality occurs only if $x_{N+1} = 0$ and $w_k^T \mathcal{P}_k w_k = 0$ for all $k \in [0, N]$. Then, by identity (2.11) in Lemma 2.7, we have $(\mathcal{D}_k^T - \mathcal{B}_k^T Q_{k+1}) \, x_{k+1} = x_k$ for all $k \in [0, N]$. And this together with $x_{N+1} = 0$ imply $x \equiv 0$. The proof is complete. $\square$

3.1.3 Functional with general endpoints

In this subsection, let $\mathcal{F}$ be the quadratic functional defined in (1.38). In the roundabout theorem for the positivity of the functional with general endpoints stated below we use the principal solution $(\hat{X}, \hat{U})$ as in the case of the functional with zero endpoints, but we again have to add some endpoint constraints. They are now formulated in terms of $2n \times 2n$ matrices, which is unavoidable, because instead of the $n \times n$ matrices Γ_0 and Γ_1 there is now the $2n \times 2n$ matrix Γ in the functional $\mathcal{F}$.

Let (S*) be the symplectic system in dimension $2n$ introduced in Subsection 2.2.1, and let $(\hat{X}^*, \hat{U}^*)$ be the conjoined basis of (S*) defined by (2.19) via the principal solution $(\hat{X}, \hat{U})$ and the associated solution $(\tilde{X}, \tilde{U})$ of (S), and let $\hat{Q}^*$ be the matrix defined by (2.33) via a matrix Q.

Theorem 3.24. *The following statements are equivalent.*

(i) $\mathcal{F}(x,u) > 0$ *over* $\mathcal{M} \left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix} \right) = 0$, *and* $x \not\equiv 0$.

(ii) *The principal solution* $(\hat{X}, \hat{U})$ *of* (S) *has no focal points in* $(0, N+1]$ *and satisfies the final endpoint inequality*

$$\hat{X}^{*T}_{N+1}(\Gamma\, \hat{X}^*_{N+1} + \hat{U}^*_{N+1}) > 0 \quad \textit{on } \operatorname{Ker} \mathcal{M} \hat{X}^*_{N+1} \setminus \operatorname{Ker} \hat{X}^*_{N+1}. \tag{3.21}$$

(iii) *The implicit Riccati equation*

$$R^*[Q^*]_k \begin{pmatrix} I & 0 \\ \Phi_{k,0} & G_k \end{pmatrix} = 0, \quad k \in [0, N],$$

has a symmetric solution $Q^*_k = \left(\begin{smallmatrix} \star & \star \\ \star & Q_k \end{smallmatrix} \right)$ *on* $[0, N+1]$ *such that* $Q^*_0 = 0$, *and* $\mathcal{P}_k = \mathcal{B}^T_k \mathcal{D}_k - \mathcal{B}^T_k Q_{k+1} \mathcal{B}_k \geq 0$ *for all* $k \in [0, N]$, *and satisfying the final endpoint inequality*

$$\Gamma + Q^*_{N+1} > 0 \quad \textit{on } \operatorname{Ker} \mathcal{M} \cap \operatorname{Im} \hat{X}^*_{N+1}. \tag{3.22}$$

(iv) *There exists a conjoined basis* (X^*, U^*) *of* (S*) *such that* X^*_k *is invertible for all* $k \in [0, N+1]$, *and* $P^*_k = X^*_k X^{*-1}_{k+1} \mathcal{B}^*_k \geq 0$ *on* $[0, N]$, *and satisfying*

$$X^{*T}_{N+1}(\Gamma\, X^*_{N+1} + U^*_{N+1}) > 0 \quad \textit{on } \operatorname{Ker} \mathcal{M} X^*_{N+1}, \tag{3.23}$$

and

$$U^*_0 X^{*-1}_0 < 0 \quad \textit{on } \operatorname{Im} \begin{pmatrix} I \\ I \end{pmatrix}. \tag{3.24}$$

(v) *There exists a symmetric solution* $Q^*_k = \left(\begin{smallmatrix} \star & \star \\ \star & Q_k \end{smallmatrix} \right)$ *on* $[0, N+1]$ *of the augmented Riccati equation* $R^*[Q^*]_k = 0$ *with* $Q^*_0 < 0$ *on* $\operatorname{Im} \left(\begin{smallmatrix} I \\ I \end{smallmatrix} \right)$, *condition* (3.2) *holds, and satisfying*

$$\Gamma + Q^*_{N+1} > 0 \quad \textit{on } \operatorname{Ker} \mathcal{M}. \tag{3.25}$$

The proof of Theorem 3.24 is based on transforming the quadratic functional $\mathcal{F}$ and system (S) into the quadratic functional

$$\mathcal{F}^*(x^*, u^*) := x_0^{*T}\Gamma_0^*\, x_0^* + x_{N+1}^{*T}\Gamma_1^*\, x_{N+1}^* + \mathcal{F}_0^*(x^*, u^*)$$

and augmented system (S*), where the endpoints of $\mathcal{F}$ are separated, i.e. $\mathcal{M}_0^* x_0^* = 0$ and $\mathcal{M}_1^* x_{N+1}^* = 0$. Here $\mathcal{M}_0^* := \frac{1}{2}\left(\begin{smallmatrix} I & -I \\ -I & I \end{smallmatrix}\right)$ and $\mathcal{M}_1^* := \mathcal{M}$ are $2n \times 2n$ projections, and $\Gamma_0^* := 0$ and $\Gamma_1^* := \Gamma$ are symmetric $2n \times 2n$ matrices.

Remark 3.25. There is also possible a different transformation to dimension $2n$, with the $2n \times 2n$ matrices $\mathcal{A}^\# := \left(\begin{smallmatrix} \mathcal{A} & 0 \\ 0 & I \end{smallmatrix}\right)$, $\mathcal{B}^\# := \left(\begin{smallmatrix} \mathcal{B} & 0 \\ 0 & 0 \end{smallmatrix}\right)$, $\mathcal{C}^\# := \left(\begin{smallmatrix} \mathcal{C} & 0 \\ 0 & 0 \end{smallmatrix}\right)$, $\mathcal{D}^\# := \left(\begin{smallmatrix} \mathcal{D} & 0 \\ 0 & I \end{smallmatrix}\right)$, $\mathcal{M}_0^\# := \mathcal{M}$, $\mathcal{M}_1^\# := \frac{1}{2}\left(\begin{smallmatrix} I & -I \\ -I & I \end{smallmatrix}\right)$, $\Gamma_0^\# := \Gamma$, $\Gamma_1^\# := 0$. Then we get another roundabout theorem for functional with general endpoints, which is a generalization of the roundabout theorem for functional with separated edpoints. In contrast to this, Theorem 3.24 does not generalize Theorem 3.14.

Now we state some auxiliary lemmas about how the admissible pairs, boundary conditions and functionals for systems (S) and (S*) are related.

Lemma 3.26. *Let (x, u) be an admissible pair defined on $[0, N+1]$, $\beta_k \in \mathbb{R}^n$ and*

$$x_k^* := \begin{pmatrix} x_0 \\ x_k \end{pmatrix}, \ k \in [0, N+1], \quad u_k^* := \begin{pmatrix} \beta_k \\ u_k \end{pmatrix}, \ k \in [0, N].$$

Then

$$\mathcal{F}^*(x^*, u^*) = \mathcal{F}(x, u).$$

Proof. It follows from the identity

$$x_0^{*T}\Gamma_0^*\, x_0^* + x_{N+1}^{*T}\Gamma_1^*\, x_{N+1}^* = \begin{pmatrix} x_0 \\ x_{N+1} \end{pmatrix}^T \Gamma \begin{pmatrix} x_0 \\ x_{N+1} \end{pmatrix}$$

and from Lemma 2.15. □

Lemma 3.27. *Let (x, u) and (x^*, u^*) be defined as in the previous lemma and let $\mathcal{M}_0^* := \frac{1}{2}\left(\begin{smallmatrix} I & -I \\ -I & I \end{smallmatrix}\right)$ and $\mathcal{M}_1^* := \mathcal{M}$. Then $\mathcal{M}_0^* x_0^* = 0$ and $\mathcal{M}_1^* x_{N+1}^* = 0$ if and only if $\mathcal{M}\left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right) = 0$.*

Proof. Since $x_0^* = \left(\begin{smallmatrix} x_0 \\ x_0 \end{smallmatrix}\right)$ and $x_{N+1}^* = \left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right)$, we have $\mathcal{M}_0^* x_0^* = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$ and $\mathcal{M}_1^* x_{N+1}^* = \mathcal{M}\left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right)$. □

Lemma 3.28. *Let* (X, U) *be a solution of* (S) *with* $\operatorname{Ker} X_{k+1} \subseteq \operatorname{Ker} X_k$. *Then for any constant nonsingular* $n \times n$ *matrix* E *and the solution* $(\bar{X}, \bar{U}) := (XE, UE)$ *of* (S) *we have*

$$\operatorname{Ker} \bar{X}_{k+1} \subseteq \operatorname{Ker} \bar{X}_k \tag{3.26}$$

and

$$P_k = \bar{P}_k, \tag{3.27}$$

where $P_k = X_k X_{k+1}^\dagger \mathcal{B}_k$ *and* $\bar{P}_k = \bar{X}_k \bar{X}_{k+1}^\dagger \mathcal{B}_k$.

Proof. Inclusion (3.26) is trivial. To prove identity (3.27) we use the identities $X_k = X_k X_{k+1}^\dagger X_{k+1}$ and $\bar{X}_{k+1} \bar{X}_{k+1}^\dagger \mathcal{B}_k = \mathcal{B}_k$ from Lemma 1.25. We have

$$\bar{P}_k = X_k E (X_{k+1} E)^\dagger \mathcal{B}_k = X_k X_{k+1}^\dagger X_{k+1} E (X_{k+1} E)^\dagger \mathcal{B}_k = X_k X_{k+1}^\dagger B_k.$$

The proof is complete. □

Proof of Theorem 3.24. (i) $\Leftrightarrow$ (ii): From Lemma 3.26 and Lemma 3.27 we get that the positivity of the functional $\mathcal{F}$ over $\mathcal{M} \left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix} \right) = 0$ is equivalent to the positivity of $\mathcal{F}^*$ over $\mathcal{M}_0^* x_0^* = 0$ and $\mathcal{M}_1^* x_{N+1}^* = 0$. We can now apply Theorem 3.14 to this transformed augmented functional. Thus, we get that the positivity of $\mathcal{F}$ is equivalent to the following.

(ii*) The augmented natural conjoined basis (X^*, U^*) of (S*) given by the initial conditions $X_0^* = I - \mathcal{M}_0^* = \frac{1}{2} \left(\begin{smallmatrix} I & I \\ I & I \end{smallmatrix} \right)$ and $U_0^* = \Gamma_0^* + \mathcal{M}_0^* = \frac{1}{2} \left(\begin{smallmatrix} I & -I \\ -I & I \end{smallmatrix} \right)$ satisfies the P^*-condition

$$P_k^* \geq 0 \quad \text{for all } k \in [0, N], \tag{3.28}$$

the kernel* condition

$$\operatorname{Ker} X_{k+1}^* \subseteq \operatorname{Ker} X_k^* \quad \text{for all } k \in [0, N], \tag{3.29}$$

and the corresponding augmented final endpoint inequality

$$X_{N+1}^{*T} (\Gamma_1^* X_{N+1}^* + U_{N+1}^*) > 0 \quad \text{on } \operatorname{Ker} \mathcal{M}_1^* X_{N+1}^* \setminus \operatorname{Ker} X_{N+1}^*. \tag{3.30}$$

Now, by Lemma 3.28 with $E := \left(\begin{smallmatrix} -I & I \\ I & I \end{smallmatrix} \right)$, this is equivalent to the following.

(ii*a) The conjoined basis $(\hat{X}^*, \hat{U}^*)$ of (S*) given by the initial conditions $\hat{X}_0^* = \left(\begin{smallmatrix} 0 & I \\ 0 & I \end{smallmatrix} \right)$ and $\hat{U}_0^* = \left(\begin{smallmatrix} -I & 0 \\ I & 0 \end{smallmatrix} \right)$ satisfies conditions (3.28), (3.29) and (3.30) with $\hat{X}^*, \hat{U}^*$ in place of X^*, U^*.

Further, by Lemma 2.25, Lemma 2.26, and from $\mathcal{M} = \mathcal{M}_1^*$, $\Gamma = \Gamma_1^*$, this is equivalent to (ii).

(i) $\Leftrightarrow$ (iii): Again, by Theorem 3.14, applied to the transformed augmented functional, we get (i) is equivalent to the following.

(iii*) The augmented implicit Riccati equation

$$R[Q^*]_k \left(\Phi^*_{k,0}(I-\mathcal{M}_0^*) \quad G_k^*\right) = 0, \quad k \in [0,N],$$

has a symmetric solution Q_k^* on $[0,N+1]$ such that $Q_0^* = \Gamma_0^*$, and $\mathcal{P}_k^* \geq 0$ for all $k \in [0,N]$, and satisfying $\Gamma_1^* + Q_{N+1}^* > 0$ on $\operatorname{Ker}\mathcal{M}_1^* \cap \operatorname{Im} X_{N+1}^*$,

where $G_k^* = \left(\left(\begin{smallmatrix} 0 & 0 \\ 0 & \Phi_{k,1}\mathcal{B}_0 \end{smallmatrix}\right) \; \left(\begin{smallmatrix} 0 & 0 \\ 0 & \Phi_{k,2}\mathcal{B}_1 \end{smallmatrix}\right) \; \cdots \; \left(\begin{smallmatrix} 0 & 0 \\ 0 & \Phi_{k,k-1}\mathcal{B}_{k-2} \end{smallmatrix}\right) \; \left(\begin{smallmatrix} 0 & 0 \\ 0 & \mathcal{B}_{k-1} \end{smallmatrix}\right) \right)$, $\Phi^*_{k,0} = \left(\begin{smallmatrix} I & 0 \\ 0 & \Phi_{k,0} \end{smallmatrix}\right)$, and where for $\underline{u}^* := (u_0^{*T}, u_1^{*T}, \dots, u_N^{*T})^T$, $u_k^* = \left(\begin{smallmatrix} \beta_k \\ u_k \end{smallmatrix}\right)$, $x_0^* = \left(\begin{smallmatrix} \alpha \\ x_0 \end{smallmatrix}\right)$ we get

$$R[Q^*]_k \left(\Phi^*_{k,0}(I-\mathcal{M}_0^*) \quad G_k^*\mathcal{T}_k^*\right) \begin{pmatrix} x_0^* \\ \underline{u}^* \end{pmatrix} = R[Q^*]_k \begin{pmatrix} \frac{1}{2}(\alpha + x_0) \\ \frac{1}{2}\Phi_{k,0}(\alpha + x_0) + G_k\mathcal{T}_k\underline{u} \end{pmatrix}$$
$$= R[Q^*]_k \begin{pmatrix} I & 0 \\ \Phi_{k,0} & G_k \end{pmatrix} \begin{pmatrix} \frac{1}{2}(\alpha + x_0) \\ \mathcal{T}_k\underline{u} \end{pmatrix}.$$

(i) $\Leftrightarrow$ (iv): By Theorem 3.14 we get that statement (i) is equivalent to the following.

(iv*) There exists a conjoined basis (X^*, U^*) of (S*) such that X_k^* is invertible for all $k \in [0, N+1]$, and $P_k^* \geq 0$ on $[0,N]$ and satisfying $X_{N+1}^{*T}(\Gamma_1^* X_{N+1}^* + U_{N+1}^*) > 0$ on $\operatorname{Ker}\mathcal{M}_1^* X_{N+1}^*$ and $X_0^{*T}(\Gamma_0^* X_0^* - U_0^*) > 0$ on $\operatorname{Ker}\mathcal{M}_0^* X_0^*$.

This is equivalent to (iv) for $\Gamma_0^* = 0$, $\Gamma_1^* = \Gamma$, $\mathcal{M}_0^* = \frac{1}{2}\left(\begin{smallmatrix} I & -I \\ -I & I \end{smallmatrix}\right)$ and $\mathcal{M}_1^* = \mathcal{M}$.

(i) $\Leftrightarrow$ (v): Again, by Theorem 3.14, statement (i) is equivalent to the following.

(v*) There exists a symmetric solution Q_k^* on $[0, N+1]$ of the Riccati equation $R^*[Q^*]_k = 0$ with $\mathcal{A}_k^* + \mathcal{B}_k^* Q_k^*$ invertible and $(\mathcal{A}_k^* + \mathcal{B}_k^* Q_k^*)^{-1}\mathcal{B}_k^* \geq 0$ for all $k \in [0,N]$, and satisfying $\Gamma_1^* + Q_{N+1}^* > 0$ on $\operatorname{Ker}\mathcal{M}_1^*$, and $\Gamma_0^* - Q_0^* > 0$ on $\operatorname{Ker}\mathcal{M}_0^*$.

Let $Q_k^* = \left(\begin{smallmatrix} \star & \star \\ \star & Q_k \end{smallmatrix}\right)$. Since $\mathcal{A}_k^* + \mathcal{B}_k^* Q_k^* = \left(\begin{smallmatrix} I & 0 \\ \star & \mathcal{A}_k + \mathcal{B}_k Q_k \end{smallmatrix}\right)$, the invertibility of the matrix $\mathcal{A}_k^* + \mathcal{B}_k^* Q_k^*$ is equivalent to the invertibility of the matrix $\mathcal{A}_k + \mathcal{B}_k Q_k$,

and since $(\mathcal{A}_k^* + \mathcal{B}_k^* Q_k^*)^{-1}\mathcal{B}_k^* = \left(\begin{smallmatrix} 0 & 0 \\ 0 & (\mathcal{A}_k+\mathcal{B}_k Q_k)^{-1}\mathcal{B}_k \end{smallmatrix}\right)$, the nonnegativity of the matrix $(\mathcal{A}_k^* + \mathcal{B}_k^* Q_k^*)^{-1}\mathcal{B}_k^*$ is equivalent to the nonnegativity of the matrix $(\mathcal{A}_k+\mathcal{B}_k Q_k)^{-1}\mathcal{B}_k$. Finally, the condition $\Gamma_0^* - Q_0^* > 0$ on $\operatorname{Ker}\mathcal{M}_0^*$ is equivalent to $\left(\begin{smallmatrix} I & I \end{smallmatrix}\right)^T Q_0^* \left(\begin{smallmatrix} I \\ I \end{smallmatrix}\right) < 0$.
This completes the proof of theorem. □

Alternative proof of equivalence (i) ⇔ (ii) in Theorem 3.24. (i) ⇒ (ii): The positivity of $\mathcal{F}$ over $\mathcal{M}\left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right) = 0$ implies the positivity of $\mathcal{F}_0$ over $x_0 = 0 = x_{N+1}$ and hence Lemma 3.5 yields that the principal solution has no focal points. It remains to show condition (3.21). Let $\left(\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix}\right) \in \operatorname{Ker}\mathcal{M} \cap \operatorname{Im}\hat{X}_{N+1}^*$. Then there exists $c \in \mathbb{R}^n$ such that $\beta = \tilde{X}_{N+1}\alpha + \hat{X}_{N+1}c$. We define an admissible pair (x, u) by

$$x_k := \tilde{X}_k\alpha + \hat{X}_k c, \quad u_k := \tilde{U}_k\alpha + \hat{U}_k c, \quad k \in [0, N+1].$$

We have $x_0 = \alpha$, $x_{N+1} = \beta$, and identity from Corollary 2.34 used on this pair yields

$$\mathcal{F}(x,u) = \begin{pmatrix} x_0 \\ x_{N+1} \end{pmatrix}^T (\Gamma + \hat{Q}_{N+1}^*) \begin{pmatrix} x_0 \\ x_{N+1} \end{pmatrix} + \mathcal{F}_0(\bar{x}, \bar{u}),$$

where $(\bar{x}, \bar{u}) \equiv (0,0)$ and $\hat{Q}_{N+1}^*$ is a symmetric matrix with $\hat{X}^{*T}\hat{Q}_{N+1}^*\hat{X}^* = \hat{X}^{*T}\hat{U}^*$. Thus, from positivity of $\mathcal{F}$ we get $\Gamma + \hat{Q}_{N+1}^* > 0$ on $\operatorname{Ker}\mathcal{M} \cap \operatorname{Im}\hat{X}_{N+1}^*$, which implies condition (3.21).

(ii) ⇒ (i): From Lemma 2.35 and Lemma 2.24 we have that $\left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right) \in \operatorname{Im}\hat{X}_{N+1}^*$. The positivity of $\mathcal{F}$ then follows from Lemma 3.5 and Corollary 2.34. (Or, alternatively, from Theorem 2.28.) □

Remark 3.29. Each statement in Theorem 3.24 is further equivalent e.g. to the following, see also Section 3.3.

(iv') *There exists conjoined basis* (X, U) *of* (S) *such that* $X_0 = I$, X_k *is invertible on* $[0, N+1]$, *and* $P_k = X_k X_{k+1}^{-1}\mathcal{B}_k \geq 0$ *on* $[0, N]$, *and satisfying*

$$\Gamma + \begin{pmatrix} -\bar{U}_0^T X_{N+1}^{-1}\bar{X}_{N+1} - \bar{U}_0 - \bar{U}_0^T - U_0 & \bar{U}_0^T X_{N+1}^{-1} \\ X_{N+1}^{-1T}\bar{U}_0 & U_{N+1}X_{N+1}^{-1} \end{pmatrix} > 0 \quad \text{on } \operatorname{Ker}\mathcal{M}, \tag{3.31}$$

where $(\bar{X}, \bar{U})$ *is a solution of* (S) *with* $\bar{X}_k^T\bar{U}_k$ *symmetric and* $\bar{X}_0 = 0$.

(v') *There exists a symmetric solution Q_k on $[0, N+1]$ of the Riccati equation $R[Q]_k = 0$ satisfying (3.2), and*

$$\Gamma + Q^*_{N+1} - \begin{pmatrix} \bar{U}_0 + \bar{U}_0^T + U_0 & 0 \\ 0 & 0 \end{pmatrix} > 0 \quad on \ \operatorname{Ker} \mathcal{M}, \tag{3.32}$$

where $(\bar{X}, \bar{U})$ is a solution of (S) with $\bar{X}_k^T \bar{U}_k$ symmetric and $\bar{X}_0 = 0$ and Q^* is defined by (2.22) via $(\bar{X}, \bar{U})$ and Q.

Proof. First we show (iv) $\Rightarrow$ (iv'), then (iv') $\Rightarrow$ (v'), and finally (v') $\Rightarrow$ (v).

(iv) $\Rightarrow$ (iv'): From (iv) we have that there exists a conjoined basis (X^*, U^*) of (S*) with no focal points in $(0, N+1]$ such that X_k^* is invertible for all $k \in [0, N+1]$, and satisfying condition (3.23) and (3.24). Without loss of generality we can assume that $X_0^* = I$. From Lemma 2.11 we have that

$$X_k^* = \begin{pmatrix} I & 0 \\ \bar{X}_k & X_k \end{pmatrix}, \quad U_k^* = \begin{pmatrix} M & N \\ \bar{U}_k & U_k \end{pmatrix}, \quad k \in [0, N+1],$$

where M, N are constant matrices and (X, U), $(\bar{X}, \bar{U})$ are solutions of (S) with $X_0 = I$ and $\bar{X}_0 = 0$. The symmetry of $X_0^{*T} U_0^*$ yields that $N = \bar{U}_0^T$ and M is symmetric, and the symmetry of $X_k^{*T} U_k^*$ further yields that $X_k^T U_k$ and $\bar{X}_k^T \bar{U}_k$ are symmetric and

$$\bar{U}_0^T = \bar{U}_k^T X_k - \bar{X}_k^T U_k, \quad \text{for } k \in [0, N+1]. \tag{3.33}$$

The matrix X_k^* is invertible if and only if X_k is invertible. Further,

$$X_k^{*-1} = \begin{pmatrix} I & 0 \\ -X_k^{-1} \bar{X}_k & X_k^{-1} \end{pmatrix}, \quad P_k^* = \begin{pmatrix} 0 & 0 \\ 0 & P_k \end{pmatrix},$$

and

$$U_k^* X_k^{*-1} = \begin{pmatrix} M - \bar{U}_0^T X_k^{-1} \bar{X}_k & \bar{U}_0^T X_k^{-1} \\ \bar{U}_k - U_k X_k^{-1} \bar{X} & U_k X_k^{-1} \end{pmatrix}.$$

Thus, $P_k^* \geq 0$ implies $P_k \geq 0$ and conditions (3.23), (3.24) imply

$$\Gamma + \begin{pmatrix} M - \bar{U}_0^T X_{N+1}^{-1} \bar{X}_{N+1} & \bar{U}_0^T X_{N+1}^{-1} \\ X_{N+1}^{-1T} \bar{U}_0 & U_{N+1} X_{N+1}^{-1} \end{pmatrix} > 0 \quad \text{on } \operatorname{Ker} \mathcal{M}, \tag{3.34}$$

and

$$M + \bar{U}_0 + \bar{U}_0^T + U_0 < 0. \tag{3.35}$$

From conditions (3.34), (3.35) we get inequality (3.34).

(iv') $\Rightarrow$ (v'): As in the proof of Lemma 3.7, we have that for $Q = UX^{-1}$ one has $R[Q]_k = 0$ and condition (3.2) holds. Further, identity (3.33) is equivalent to

$$\bar{U}_0^T X_k^{-1} = \bar{U}_k^T - \bar{X}_k^T Q_k, \quad \text{for } k \in [0, N+1].$$

Hence, condition (3.31) is equivalent to (3.32).

(v') $\Rightarrow$ (v): From (3.32) we have that there exists $\varepsilon > 0$ such that

$$\Gamma + \begin{pmatrix} \bar{X}_{N+1}^T Q_{N+1} \bar{X}_{N+1} - \bar{X}_{N+1}^T \bar{U}_{N+1} & \bar{U}_{N+1}^T - \bar{X}_{N+1}^T Q \\ \bar{U}_{N+1} - Q_{N+1} \bar{X}_{N+1} & Q_{N+1} \end{pmatrix} \tag{3.36}$$
$$- \begin{pmatrix} \bar{U}_0 + \bar{U}_0^T + U_0 & 0 \\ 0 & 0 \end{pmatrix} > \varepsilon \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \quad \text{on } \operatorname{Ker} \mathcal{M}.$$

The solution is

$$Q^* := \begin{pmatrix} \bar{X}^T Q \bar{X} - \bar{X}^T \bar{U} & \bar{U}^T - \bar{X}^T Q \\ \bar{U} - Q\bar{X} & Q \end{pmatrix} - \begin{pmatrix} \varepsilon I + \bar{U}_0 + \bar{U}_0^T + Q_0 & 0 \\ 0 & 0 \end{pmatrix}.$$

The identity $R^*[M^* + \left(\begin{smallmatrix} M & 0 \\ 0 & 0 \end{smallmatrix}\right)] = R^*[M^*]$ holds for any $2n \times 2n$ matrix M^* and $n \times n$ matrix M. Thus, by Lemma 2.21, we have $R^*[Q^*] = 0$. Further, $Q_0^* = \left(\begin{smallmatrix} -\varepsilon I - \bar{U}_0 - \bar{U}_0^T - Q_0 & \bar{U}_0^T \\ \bar{U}_0 & Q_0 \end{smallmatrix}\right) < 0$ on $\operatorname{Im}\left(\begin{smallmatrix} I \\ I \end{smallmatrix}\right)$, and inequality (3.36) implies condition (3.25). $\square$

3.1.4 Examples

In the following examples we show a situation when a symmetric Q_k solves Riccati inequality (3.4) and satisfies condition (3.2), but it does not solve Riccati equation (3.1).

Example 3.30. Let $\mathcal{A}_k \equiv 0$, $\mathcal{B}_k \equiv -\mathcal{C}^{T-1}$, $\mathcal{C}_k \equiv \mathcal{C}$, $\mathcal{D}_k \equiv -\mathcal{C}^{T-1} - \mathcal{C} - K$, where $\mathcal{C}$ is a constant nonsingular matrix, $K \neq 0$, and $\mathcal{C}K^T = K\mathcal{C}^T \geq 0$. Then $Q_k \equiv I$ satisfies condition (3.2), since $\mathcal{A}_k + \mathcal{B}_k Q_k = -\mathcal{C}^{T-1}$ is invertible, $\mathcal{P}_k \equiv I > 0$, and $R[Q]_k(\mathcal{A}_k + \mathcal{B}_k Q_k)^{-1} = -K\mathcal{C}^T \leq 0$, while the Riccati equation is $R[Q]_k = K \neq 0$. Another (more specific) example can be obtained when we take e.g. $\mathcal{C} = K = I$.

Example 3.31. Let $\mathcal{A}_k$ and $\mathcal{C}_k$ be invertible, $\mathcal{B}_k \equiv 0$, and $\mathcal{D}_k = \mathcal{A}_k^{T-1}$, with $\mathcal{C}_k^T \mathcal{A}_k > 0$. Then $Q_k \equiv 0$ satisfies condition (3.2), since $\mathcal{A}_k + \mathcal{B}_k Q_k = \mathcal{A}_k$ is invertible, $\mathcal{P}_k \equiv 0$, and $R[Q]_k(\mathcal{A}_k + \mathcal{B}_k Q_k)^{-1} = -\mathcal{C}_k \mathcal{A}_k^{-1} < 0$, while the Riccati equation is $R[Q]_k = -\mathcal{C}_k \neq 0$. However, in this simple example we can directly verify that $\mathcal{F} > 0$ over free endpoints.

3.2 Nonnegativity of quadratic functionals

The main difference between roundabout theorems for the positivity and nonnegativity of $\mathcal{F}$ is that the kernel condition $\operatorname{Ker} X_{k+1} \subseteq \operatorname{Ker} X_k$ is not necessary for the nonnegativity, while the image condition $x_k \in \operatorname{Im} X_k$ is used instead. The P-condition remains, but also in a weakened form. Further, in the literature there are no statements about the solvability of the explicit Riccati equation and inequality for the case of nonnegative quadratic functionals.

Before we state in Subsections 3.2.3, 3.2.4 and 3.2.5 Reid roundabout theorems for the nonnegativity, we introduce in the following two subsections matrices M and T and corresponding augmented matrices M^* and T^*, since the matrix T appears in these statements. These matrices were for the first time defined in [46].

3.2.1 Matrices M and T

In this subsection, (X, U) is a conjoined basis of (S), Q_k is a symmetric matrix with $Q_k X_k = U_k X_k^\dagger X_k$, $P_k = X_k X_{k+1}^\dagger \mathcal{B}_k$ and $\mathcal{P}_k = \mathcal{B}_k^T \mathcal{D}_k + \mathcal{B}_k^T Q_{k+1} \mathcal{B}_k$.

Let us define the $n \times n$ matrices

$$M_k := (I - X_{k+1} X_{k+1}^\dagger)\, \mathcal{B}_k, \quad T_k := I - M_k^\dagger M_k. \tag{3.37}$$

In the following two lemmas we show some properties of these matrices.

Lemma 3.32. *The matrix $M_k = 0$ if and only if the kernel condition $\operatorname{Ker} X_{k+1} \subseteq \operatorname{Ker} X_k$ holds.*

Proof. It follows directly from Lemma 1.25. □

Lemma 3.33. *The following identities hold.*

$$M_k T_k = 0, \tag{3.38}$$
$$\mathcal{B}_k T_k = X_{k+1} X_{k+1}^{\dagger} \mathcal{B}_k T_k, \tag{3.39}$$
$$X_{k+1}^T M_k = 0, \tag{3.40}$$
$$M_k^{\dagger} X_{k+1} = 0, \tag{3.41}$$
$$T_k X_k = T_k X_k X_{k+1}^{\dagger} X_{k+1}, \tag{3.42}$$
$$T_k P_k T_k = T_k \mathcal{P}_k T_k. \tag{3.43}$$

Proof. Identities (3.38), (3.39) and (3.40) hold by the definition of M_k and T_k, identity (3.41) follows from (3.40) and from the property of a generalized inverse $\operatorname{Ker} A^{\dagger} = \operatorname{Ker} A^T$. Identities (3.42) and (3.43) are proven below,

$$\begin{aligned} T_k X_k &= T_k(\mathcal{D}_k^T X_{k+1} - \mathcal{B}_k^T U_{k+1}) = T_k(\mathcal{D}_k^T X_{k+1} - \mathcal{B}_k^T X_{k+1}^{\dagger T} X_{k+1}^T U_{k+1}) \\ &= T_k(\mathcal{D}_k^T - \mathcal{B}_k^T X_{k+1}^{\dagger T} U_{k+1}^T) X_{k+1} = T_k(\mathcal{D}_k^T X_{k+1} - \mathcal{B}_k^T U_{k+1}) X_{k+1}^{\dagger} X_{k+1} \\ &= T_k X_k X_{k+1}^{\dagger} X_{k+1}, \end{aligned}$$

$$\begin{aligned} P_k T_k &= X_k X_{k+1}^{\dagger} \mathcal{B}_k T_k = (\mathcal{D}_k^T X_{k+1} - \mathcal{B}_k^T U_{k+1}) X_{k+1}^{\dagger} \mathcal{B}_k T_k \\ &= \mathcal{D}_k^T \mathcal{B}_k T_k - \mathcal{B}_k^T Q_{k+1} X_{k+1} X_{k+1}^{\dagger} \mathcal{B}_k T_k = (\mathcal{D}_k^T \mathcal{B}_k - \mathcal{B}_k^T Q_{k+1} \mathcal{B}_k) T_k = \mathcal{P}_k T_k. \end{aligned}$$

Thus, $T_k P_k T_k = T_k \mathcal{P}_k T_k$ and this lemma is proven. □

Remark 3.34. The matrix $T_k P_k T_k$ is symmetric, by identity (3.43).

Remark 3.35. Identity (3.42) is equivalent to $\operatorname{Ker} X_{k+1} \subseteq \operatorname{Ker} T_k X_k$, by equivalence (1.5).

In the rest of this subsection we prove several auxiliary lemmas that are used in proofs of roundabout theorems for the nonnegativity.

Lemma 3.36. *Let (x, u) be such that $x_{k+1} = \mathcal{A}_k x_k + \mathcal{B}_k u_k$, $x_k \in \operatorname{Im} X_k$, and $x_{k+1} \in \operatorname{Im} X_{k+1}$. Then $M_k w_k = 0$, where $w_k := u_k - Q_k x_k$. This further implies that $w_k = T_k w_k$.*

Proof. Let $x_k = X_k c_k$ and $x_{k+1} = X_{k+1} d_k$. Then

$$\begin{aligned} & M_k(u_k - Q_k x_k) = M_k(u_k - U_k X_k^{\dagger} x_k) = (I - X_{k+1} X_{k+1}^{\dagger}) \mathcal{B}_k (u_k - U_k X_k^{\dagger} x_k) \\ & = (I - X_{k+1} X_{k+1}^{\dagger})(X_{k+1} d - \mathcal{A}_k X_k c - (X_{k+1} - \mathcal{A}_k X_k) X_k^{\dagger} X_k c_k) = 0. \end{aligned}$$

Thus, this lemma is proven. □

Lemma 3.37. *Let $m \in [0, N]$ be fixed. Then there exists an $n \times n$ matrix S' that $X_{m+1}S' = 0$ and $U_{m+1}S' = M_m$.*

Proof. Let us take a pair $(\bar{X}_{m+1}, \bar{U}_{m+1})$ such that $\begin{pmatrix} \bar{X}_{m+1} & X_{m+1} \\ \bar{U}_{m+1} & U_{m+1} \end{pmatrix}$ is symplectic. This is possible by Lemma 1.20, since we can put

$$\begin{aligned} \bar{X}_{m+1} :=&\ U_{m+1}\begin{pmatrix} X_{m+1}^T & U_{m+1}^T \end{pmatrix}\begin{pmatrix} X_{m+1} \\ U_{m+1} \end{pmatrix}, \\ \bar{U}_{m+1} :=&\ -X_{m+1}\begin{pmatrix} X_{m+1}^T & U_{m+1}^T \end{pmatrix}\begin{pmatrix} X_{m+1} \\ U_{m+1} \end{pmatrix}. \end{aligned}$$

Now for $S' := \bar{X}_{m+1}^T M_m$ we have

$$\begin{aligned} X_{m+1}S' =&\ X_{m+1}\bar{X}_{m+1}^T M_m = \bar{X}_{m+1}X_{m+1}^T M_m = 0, \\ U_{m+1}S' =&\ U_{m+1}\bar{X}_{m+1}^T M_m = (I + \bar{U}_{m+1}X_{m+1}^T)M_m = M_m, \end{aligned}$$

and thus existence of such a matrix S' is proven. $\square$

Lemma 3.38. *Let $m \in [0, N]$ be fixed. If $x_m \in \operatorname{Im} X_m$ and $x_{m+1} \notin \operatorname{Im} X_{m+1}$, then $M_m^T x_{m+1} \neq 0$.*

Proof. Suppose that $M_m^T x_{m+1} = 0$ and $x_m = X_m \alpha$. Then

$$\begin{aligned} 0 &= \mathcal{B}_m^T(I - X_{m+1}X_{m+1}^\dagger)\, x_{m+1} = \mathcal{B}_m^T(I - X_{m+1}X_{m+1}^\dagger)(\mathcal{A}_m X_m \alpha + \mathcal{B}_m u_m) \\ &= \mathcal{B}_m^T(I - X_{m+1}X_{m+1}^\dagger)X_{m+1}\alpha + M_m^T \mathcal{B}_m(u_m - U_m\alpha) = M_m^T M_m(u_m - U_m\alpha). \end{aligned}$$

From that

$$\begin{aligned} 0 &= M_m(u_m - U_m\alpha) = (I - X_{m+1}X_{m+1}^\dagger)\mathcal{B}_m(u_m - U_m\alpha) \\ &= x_{m+1} - X_{m+1}X_{m+1}^\dagger x_{m+1}, \end{aligned}$$

and this implies $x_{m+1} = X_{m+1}X_{m+1}^\dagger x_{m+1}$, which is a contradiction with $x_{m+1} \notin \operatorname{Im} X_{m+1}$. $\square$

3.2.2 Matrices M^* and T^*

Let $(\bar{X}, \bar{U})$ and (X, U) be normalized conjoined bases of (S) and (X^*, U^*) be the conjoined basis of (S*) defined by (2.18) through $(\bar{X}, \bar{U})$ and (X, U). Then, as in (3.37), we can define the $2n \times 2n$ matrices

$$M_k^* := (I - X_{k+1}^* X_{k+1}^{*\dagger})\, \mathcal{B}_k^*, \quad T_k^* := I - M_k^{*\dagger} M_k^*. \tag{3.44}$$

Lemma 3.39. *Let M_k^* be the matrix defined in* (3.44) *and let M_k be the matrix defined in* (3.37) *via* (X, U). *Then*

$$M_k^* = \begin{pmatrix} 0 & -(I + \bar{X}_{k+1}^T \bar{X}_{k+1})^{-1} \bar{X}_{k+1}^T M_k \\ 0 & [I - \bar{X}_{k+1}(I + \bar{X}_{k+1}^T \bar{X}_{k+1})^{-1} \bar{X}_{k+1}^T]\, M_k \end{pmatrix} \tag{3.45}$$

and

$$M_k^{*\dagger} M_k^* = \begin{pmatrix} 0 & 0 \\ 0 & M_k^\dagger M_k \end{pmatrix}. \tag{3.46}$$

Proof. When we use formula (2.20) for X_{k+1}^* and compute M_k^* from definition (3.44), we get identity (3.45).

Now we prove identity (3.46). Let $M_k = F_k R_k$ be a full rank factorization of M_k, i.e., $F_k \in \mathbb{R}^{n \times r_k}$ and $R_k \in \mathbb{R}^{r_k \times n}$ with $r_k := \operatorname{rank} M_k = \operatorname{rank} F_k = \operatorname{rank} R_k$. We define matrices $F_k^* \in \mathbb{R}^{2n \times r_k}$ and $R_k^* \in \mathbb{R}^{r_k \times 2n}$ by

$$F_k^* := \begin{pmatrix} -(I + \bar{X}_{k+1}^T \bar{X}_{k+1})^{-1} \bar{X}_{k+1}^T F_k \\ [I - \bar{X}_{k+1}(I + \bar{X}_{k+1}^T \bar{X}_{k+1})^{-1} \bar{X}_{k+1}^T]\, F_k \end{pmatrix}, \quad R_k^* := \begin{pmatrix} 0 & R_k \end{pmatrix}.$$

Then $F_k^* R_k^* = M_k^*$, $r_k = \operatorname{rank} R_k^* = \operatorname{rank} M_k^*$, while $\operatorname{rank} F_k^* = r_k$ follows from the invertibility of the matrix on the right-hand side of identity (1.7) with $A = \bar{X}_{k+1}$. Thus, $M_k^* = F_k^* R_k^*$ is a full rank factorization of M_k^* and from identity (1.3), applied first to the matrices M_k^* and R_k^* and then to M_k and R_k, we get

$$M_k^{*\dagger} M_k^* = R_k^{*T} (R_k^* R_k^{*T})^{-1} R_k^* = \begin{pmatrix} 0 & 0 \\ 0 & R_k^T (R_k R_k^T)^{-1} R_k \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & M_k^\dagger M_k \end{pmatrix}.$$

The proof is complete. ☐

Remark 3.40. We can also compute the formula for the Moore-Penrose inverse of M_k^*. When we apply identity (1.2) to the matrices F_k^* and R_k^*, we get

$$M_k^{*\dagger} = \begin{pmatrix} 0 & 0 \\ -H_k \bar{X}_{k+1} (I + \bar{X}_{k+1}^T \bar{X}_{k+1})^{-1} & H_k (I + \bar{X}_{k+1} \bar{X}_{k+1}^T)^{-1} \end{pmatrix},$$

where $H_k := R_k^T (R_k R_k^T)^{-1} [F_k^T (I + \bar{X}_{k+1} \bar{X}_{k+1}^T)^{-1} F_k]^{-1} F_k^T \in \mathbb{R}^{r_k \times n}$.

Lemma 3.41. *The following identities hold*

$$T_k^* = \begin{pmatrix} I & 0 \\ 0 & T_k \end{pmatrix}, \quad T_k^* P_k^* T_k^* = \begin{pmatrix} 0 & 0 \\ 0 & T_k P_k T_k \end{pmatrix}. \tag{3.47}$$

Proof. The first identity is consequence of (3.46), the second identity we get from the first one, Lemma 2.22, and identity (3.43). ☐

3.2.3 Functional with zero endpoints

Recall that the functional $\mathcal{F}_0$ is defined by formula (1.30).

Theorem 3.42. *The following statements are equivalent.*

(i) $\mathcal{F}_0(x,u) \geq 0$ *over* $x_0 = 0$ *and* $x_{N+1} = 0$.

(ii) *The principal solution* $(\hat{X}, \hat{U})$ *of* (S) *satisfies the* P*-condition*

$$T_k P_k T_k \geq 0 \quad \textit{for all } k \in [0, N], \tag{3.48}$$

and the image condition

$$\begin{gathered} x_k \in \operatorname{Im} \hat{X}_k \quad \textit{for all } k \in [0, N+1], \\ \textit{for all admissible } (x,u) \textit{ with } x_0 = 0,\ x_{N+1} = 0. \end{gathered} \tag{3.49}$$

(iii) *The implicit Riccati equation*

$$\mathcal{T}_{k+1}^T G_{k+1}^T R[Q]_k G_k \mathcal{T}_k = 0 \ \textit{on} \ \operatorname{Ker} G_{N+1}, \quad k \in [0, N],$$

has a symmetric solution Q_k *such that* $Q_k \hat{X}_k = \hat{U}_k \hat{X}_k^{\dagger} \hat{X}_k$ *on* $[0, N+1]$, *and*

$$T_k \mathcal{P}_k T_k \geq 0, \quad \textit{for all } k \in [0, N]. \tag{3.50}$$

(iv) *The conjoined basis* (X, U) *of* (S) *given by the initial conditions*

$$X_0 = I - \tilde{X}_{N+1}^{\dagger} \tilde{X}_{N+1}, \quad U_0 = \tilde{X}_{N+1}^{\dagger} \tilde{X}_{N+1}$$

satisfies P*-condition* (3.48) *and the image condition*

$$\begin{gathered} x_k \in \operatorname{Im} X_k \quad \textit{for all } k \in [0, N+1], \\ \textit{for all admissible } (x,u) \textit{ with } \tilde{X}_{N+1} x_0 = 0 = x_{N+1}. \end{gathered}$$

(v) *There exist symmetric matrices* $F_k \leq 0$, $k \in [0, N]$, *and a solution* (X, U) *of the system*

$$\begin{aligned} X_{k+1} &= \mathcal{A}_k X_k + \mathcal{B}_k U_k, \\ F_k X_{k+1} &= U_{k+1} - \mathcal{C}_k X_k - \mathcal{D}_k U_k, \end{aligned}$$

$k \in [0, N]$, satisfying the initial conditions

$$X_0 = I - \underline{\tilde{X}}_{N+1}^{\dagger} \underline{\tilde{X}}_{N+1}, \quad U_0 = \underline{\tilde{X}}_{N+1}^{\dagger} \underline{\tilde{X}}_{N+1},$$

such that $X_k^T U_k$ is symmetric for all $k \in [0, N+1]$, P-condition (3.48) holds, and the image condition

$$\begin{gathered} x_k \in \operatorname{Im} X_k \quad \textit{for all } k \in [0, N+1] \\ \textit{for all admissible } (x, u) \textit{ with } \underline{\tilde{X}}_{N+1} x_0 = 0 = x_{N+1} \end{gathered}$$

holds, where $(\underline{\tilde{X}}, \underline{\tilde{U}})$ is the conjoined basis of the system ($\underline{\mathrm{S}}$), which has the coefficient matrix $\underline{\mathcal{S}}_k$ defined in (3.8), with the initial conditions $(\underline{\tilde{X}}_0, \underline{\tilde{U}}_0) = (I, 0)$.

Proof. The equivalence of statements (i), (ii), and (iii) is a corollary of Theorem 3.43 for functional with separated endpoints, which we prove in the next subsection. For proofs of the equivalence of statements (iv) and (v) with statement (i), see [37]. □

3.2.4 Functional with separated endpoints

In this subsection, let (X, U) be the natural conjoined basis of (S). Recall that the functional $\mathcal{F}$ is defined by formula (1.37).

Theorem 3.43. *The following statements are equivalent.*

(i) *$\mathcal{F}(x, u) \geq 0$ over $\mathcal{M}_0 x_0 = 0$ and $\mathcal{M}_1 x_{N+1} = 0$.*

(ii) *The natural conjoined basis (X, U) of (S) satisfies P-condition (3.48), the image condition*

$$\begin{gathered} x_k \in \operatorname{Im} X_k \quad \textit{for all } k \in [0, N+1], \\ \textit{for all admissible } (x, u) \textit{ with } \mathcal{M}_0 x_0 = 0,\ \mathcal{M}_1 x_{N+1} = 0, \end{gathered} \tag{3.51}$$

and the final endpoint inequality

$$X_{N+1}^T (\Gamma_1 X_{N+1} + U_{N+1}) \geq 0 \quad \textit{on } \operatorname{Ker} \mathcal{M}_1 X_{N+1}. \tag{3.52}$$

(iii) *The implicit Riccati equation*

$$\begin{aligned} &\begin{pmatrix} (I-\mathcal{M}_0)\,\Phi_{k+1,0}^T \\ \mathcal{T}_{k+1}^T G_{k+1}^T \end{pmatrix} R[Q]_k \begin{pmatrix} \Phi_{k,0}(I-\mathcal{M}_0) & G_k\mathcal{T}_k \end{pmatrix} = 0 \\ &\qquad \textit{on } \operatorname{Ker}\mathcal{M}_1 \begin{pmatrix} \Phi_{N+1,0}(I-\mathcal{M}_0) & G_{N+1} \end{pmatrix}, \quad k \in [0,N], \end{aligned} \tag{3.53}$$

has a symmetric solution Q_k *such that* $Q_k X_k = U_k X_k^\dagger X_k$ *on* $[0, N+1]$, $\mathcal{P}$*-condition* (3.50) *holds, and satisfying the final endpoint inequality*

$$Q_{N+1} + \Gamma_1 \geq 0 \quad \textit{on } \operatorname{Ker}\mathcal{M}_1 \cap \operatorname{Im} X_{N+1}. \tag{3.54}$$

First we prove that P-condition (3.48), image condition (3.51), and inequality (3.52) are necessary for the nonnegativity of $\mathcal{F}$. The necessity of the P-condition and of the final inequality is proven similarly as in the case when $\mathcal{F}$ is positive. However, the image condition, which corresponds to the kernel condition in case when $\mathcal{F}$ is positive, is proven rather differently.

Lemma 3.44. *If* $\mathcal{F}(x,u) \geq 0$ *over* $\mathcal{M}_0\, x_0 = 0$ *and* $\mathcal{M}_1 x_{N+1} = 0$*, then* $T_k P_k T_k \geq 0$ *for all* $k \in [0,N]$*, i.e.* P*-condition* (3.48) *holds.*

Proof. Suppose that there exists $m \in [0,N]$ such that $T_m P_m T_m \ngeq 0$. Then there is $c \neq 0$ with $c^T T_m P_m T_m c < 0$ and, therefore, $P_m T_m c \neq 0$. Define $d := X_{m+1}^\dagger \mathcal{B}_m T_m c$. Then $X_m d = P_m T_m c \neq 0$, which follows from identity (3.39) in Lemma 3.33. Now we define (x,u) by

$$\begin{aligned} x_k &:= \begin{cases} X_k d, & \text{for } k \in [0,m], \\ 0, & \text{for } k \in [m+1, N+1], \end{cases} \\ u_k &:= \begin{cases} U_k d, & \text{for } k \in [0, m-1], \\ -\mathcal{A}_m^T(\mathcal{D}_m - Q_{m+1}\mathcal{B}_m)T_m c, & \text{for } k = m, \\ 0, & \text{for } k \in [m+1, N+1], \end{cases} \end{aligned} \tag{3.55}$$

where Q_{m+1} is a symmetric matrix with $Q_{m+1}X_{m+1} = U_{m+1}X_{m+1}^\dagger X_{m+1}$, see e.g. formula (2.3). Such defined (x,u) is admissible and $\mathcal{M}_0\, x_0 = \mathcal{M}_0 X_0\, d =$

$\mathcal{M}_0(I-\mathcal{M}_0)\,d=0$, $\mathcal{M}_1 x_{N+1}=0$. Further,

$$\begin{aligned}
\mathcal{F}(x,u) &= x_0^T\Gamma_0 x_0 + x_{N+1}^T\Gamma_1 x_{N+1} + x_k^T u_k\Big|_0^{N+1} + \sum_{k=0}^{N} x_{k+1}^T(\mathcal{C}_k x_k + \mathcal{D}_k u_k - u_{k+1}) \\
&= x_m^T[\mathcal{C}_{m-1}X_{m-1}d + \mathcal{D}_{m-1}U_{m-1}d + \mathcal{A}_m^T(\mathcal{D}_m - Q_{m+1}\mathcal{B}_m)T_m c] \\
&= d^T X_m^T U_m d + d^T(X_{m+1}^T - U_m^T\mathcal{B}_m^T)(\mathcal{D}_m - Q_{m+1}\mathcal{B}_m)T_m c \\
&= d^T X_m^T U_m d + d^T X_{m+1}^T X_{m+1}^{\dagger T}(X_{m+1}^T\mathcal{D}_m - U_{m+1}^T\mathcal{B}_m)T_m c - U_m^T P_m T_m c \\
&= d^T X_{m+1}^T X_{m+1}^{\dagger T} X_m T_m c = d^T X_m T_m c = c_m^T T_m P_m T_m c < 0,
\end{aligned}$$

which is a contradiction, because $\mathcal{F}$ was supposed to be nonnegative. □

Lemma 3.45. *If $\mathcal{F}(x,u)\geq 0$ over $\mathcal{M}_0\,x_0=0$ and $\mathcal{M}_1 x_{N+1}=0$, then $x_k\in \operatorname{Im} X_k$ for all $k\in[0,N+1]$ for all admissible (x,u) with $\mathcal{M}_0\,x_0=0$ and $\mathcal{M}_1 x_{N+1}=0$, i.e. image condition* (3.51) *holds.*

Proof. Let (x,u) be an admissible pair with $\mathcal{M}_0\,x_0=\mathcal{M}_1 x_{N+1}=0$. Suppose that there exists an index m such that $x_m\in\operatorname{Im} X_m$ but $x_{m+1}\notin \operatorname{Im} X_{m+1}$. Certainly $m>0$, because $\mathcal{M}_0\,x_0=0$ implies $x_0\in\operatorname{Im}(I-\mathcal{M}_0)=\operatorname{Im} X_0$. Let α be such that $X_m\alpha=x_m$. Now we take the matrix S' from Lemma 3.37 and an arbitrary real number t and define $\tilde{\alpha}:=tS'M_m^T x_{m+1}$. Further we define $(\tilde{x},\tilde{u})$ by

$$\begin{aligned}
\tilde{x}_k &:= \begin{cases} X_k(\alpha+\tilde{\alpha}), & \text{for } k\in[0,m], \\ x_k, & \text{for } k\in[m+1,N+1], \end{cases} \\
\tilde{u}_k &:= \begin{cases} U_k(\alpha+\tilde{\alpha}), & \text{for } k\in[0,m-1], \\ u_m + U_m\tilde{\alpha}, & \text{for } k=m, \\ u_k, & \text{for } k\in[m+1,N+1]. \end{cases}
\end{aligned} \tag{3.56}$$

Such defined $(\tilde{x},\tilde{u})$ is admissible, since for $k\in[0,m-1]$ it is a solution of (S), for $k\in[m+1,N]$ it is equal to the admissible (x,u) and for $k=m$ we have

$$\begin{aligned}
\mathcal{A}_m\tilde{x}_m + \mathcal{B}_m\tilde{u}_m &= \mathcal{A}_m X_m(\alpha+\tilde{\alpha}) + \mathcal{B}_m(u_m + U_m\tilde{\alpha}) \\
&= \tilde{x}_{m+1} + X_{m+1}\tilde{\alpha} = \tilde{x}_{m+1} + tX_{m+1}S'M_m^T x_{m+1} = \tilde{x}_{m+1},
\end{aligned}$$

where we used $X_{m+1}S'=0$. Further, $\mathcal{M}_0\tilde{x}_0=\mathcal{M}_0 X_0(\alpha+\tilde{\alpha})=0$, and $\mathcal{M}_1\tilde{x}_{N+1}=\mathcal{M}_1 x_{N+1}=0$, hence by the assumption we have $\mathcal{F}(\tilde{x},\tilde{u})\geq$

0. Now we show that this cannot be true for all $t \in \mathbb{R}$ and thus get the contradiction.

We denote $\mathbb{X}_k := \tilde{x}_{k+1}^T(\mathcal{C}_k\tilde{x}_k + \mathcal{D}_k\tilde{u}_k - \tilde{u}_{k+1})$. Then we have

$$\mathbb{X}_k = \begin{cases} 0, & \text{for } k \in [0, m-2], \\ (\alpha + \tilde{\alpha})^T X_m^T (U_m\alpha - u_m), & \text{for } k = m-1, \\ x_{m+1}^T(\mathcal{C}_m X_m \alpha + \mathcal{D}_m u_m + U_{m+1}\tilde{\alpha} - u_{m+1}), & \text{for } k = m, \\ x_{m+2}^T(\mathcal{C}_{m+1}x_{m+1} + \mathcal{D}_{m+1}u_{m+1} - u_{m+2}), & \text{for } k \in [m+1, N]. \end{cases} \tag{3.57}$$

Now we compute the sum

$$\begin{aligned} \sum_{k=0}^{N} \mathbb{X}_k &= \left\{ \sum_{k=0}^{m-2} + \sum_{k=m+1}^{N} \right\} \mathbb{X}_k + \mathbb{X}_{m-1} + \mathbb{X}_m = (\alpha + \tilde{\alpha})^T X_m^T (U_m\alpha - u_m) \\ &\quad + x_{m+1}^T(\mathcal{C}_m X_m\alpha + \mathcal{D}_m u_m + U_{m+1}\tilde{\alpha} - u_{m+1}) + \sum_{k=m+1}^{N} \mathbb{X}_k \\ &= \alpha^T X_m^T(U_m\alpha - u_m) + x_{m+1}^T(\mathcal{C}_m X_m\alpha + \mathcal{D}_m u_m - u_{m+1}) + \sum_{k=m+1}^{N} \mathbb{X}_k \\ &\quad + [(\alpha^T U_m^T - u_m^T)X_m + x_{m+1}^T U_{m+1}]\tilde{\alpha}, \end{aligned}$$

where further

$$\begin{aligned} &[(\alpha^T U_m^T - u_m^T)X_m + x_{m+1}^T U_{m+1}]\tilde{\alpha} = \\ &\qquad = t[(\alpha^T U_m^T - u_m^T)(\mathcal{D}_m^T X_{m+1} - \mathcal{B}_m^T U_{m+1}) + x_{m+1}^T U_{m+1}] S' M_m^T x_{m+1} \\ &\qquad = t[\mathcal{B}_m(u_m - U_m\alpha) + x_{m+1}]^T M_m M_m^T x_{m+1} = 2t\, x_{m+1}^T M_m M_m^T x_{m+1}. \end{aligned}$$

From that we get

$$\begin{aligned} \mathcal{F}(\tilde{x}, \tilde{u}) &= \tilde{x}_0^T \Gamma_0 \tilde{x}_0 + \tilde{x}_{N+1}^T \Gamma_1 \tilde{x}_{N+1} + \tilde{x}_k^T \tilde{u}_k \big|_0^{N+1} + \sum_{k=0}^{N} \tilde{x}_{k+1}^T(\mathcal{C}_k\tilde{x}_k + \mathcal{D}_k\tilde{u}_k - \tilde{u}_{k+1}) \\ &= x_{N+1}^T(\Gamma_1 x_{N+1} + u_{N+1}) + \alpha^T X_m^T(U_m\alpha - u_m) \\ &\quad + x_{m+1}^T(\mathcal{C}_m X_m\alpha + \mathcal{D}_m u_m - u_{m+1}) + \sum_{k=m+1}^{N} \mathbb{X}_k + 2\,t\|M_m^T x_{m+1}\|^2, \end{aligned}$$

and as $M_m^T x_{m+1} \neq 0$ by Lemma 3.38, there exists t sufficiently large negative such that $\mathcal{F}(\tilde{x}, \tilde{u}) < 0$. $\square$

Lemma 3.46. *If* $\mathcal{F}(x,u) \geq 0$ *over* $\mathcal{M}_0 x_0 = 0$ *and* $\mathcal{M}_1 x_{N+1} = 0$ *then* $X_{N+1}^T(\Gamma_1 X_{N+1} + U_{N+1}) \geq 0$ *on* $\operatorname{Ker} \mathcal{M}_1 X_{N+1}$, *i.e. final inequality* (3.52) *holds.*

Proof. Let $d \in \operatorname{Ker} \mathcal{M}_1 X_{N+1}$ and take the pair (Xd, Ud). It is obviously admissible and $\mathcal{M}_0 x_0 d = 0 = \mathcal{M}_1 X_{N+1} d$. Hence, $\mathcal{F}(Xd, Ud) \geq 0$ and from Lemma 1.31 we have

$$\begin{aligned}\mathcal{F}(Xd, Ud) &= d^T X_0^T \Gamma_0 X_0\, d + d^T X_{N+1}^T \Gamma_1 X_{N+1} d + d^T X_{N+1}^T U_{N+1} d - d^T X_0^T U_0\, d \\ &= d^T X_{N+1}^T (\Gamma_1 X_{N+1} + U_{N+1})\, d.\end{aligned}$$

Thus, the inequality is proven. □

The next lemma says that P-condition (3.48), image condition (3.51), and final endpoint inequality (3.52) are sufficient for the nonnegetivity of $\mathcal{F}$. It is proven again by means of the Picone identity.

Lemma 3.47. *If the natural conjoined basis* (X, U) *of* (S) *satisfies P-condition* (3.48)*, image condition* (3.51)*, and final endpoint inequality* (3.52)*, then* $\mathcal{F}(x,u) \geq 0$ *over* $\mathcal{M}_0 x_0 = 0$ *and* $\mathcal{M}_1 x_{N+1} = 0$.

Proof. Let (x, u) be admissible with $\mathcal{M}_0 x_0 = 0$, $\mathcal{M}_1 x_{N+1} = 0$. As $x_k \in \operatorname{Im} X_k$ for all $k \in [0, N+1]$, we get from Theorem 2.9 (Picone identity), see also formula (3.16), and from Lemma 3.36

$$\mathcal{F}(x,u) = d^T (X_{N+1}^T \Gamma_1 X_{N+1} + X_{N+1}^T U_{N+1})\, d + \sum_{k=0}^{N} w_k^T T_k \mathcal{P}_k T_k w_k,$$

where $d \in \mathbb{R}^n$ is such that $X_{N+1} d = x_{N+1}$, and $\mathcal{P}_k = P_k \geq 0$ by Lemma 2.5. This together with (3.48) and (3.52) imply $\mathcal{F}(x,u) \geq 0$ for all (x,u) with $\mathcal{M}_0 x_0 = 0$, $\mathcal{M}_1 x_{N+1} = 0$. □

Lemma 3.48. *Statements (ii) and (iii) in Theorem 3.43 are equivalent.*

Proof. (ii) ⇒ (iii): We define Q_k by (2.2) or (2.3) via the natural conjoined basis (X, U). Then $T_k \mathcal{P}_k T_k = T_k P_k T_k \geq 0$ on $[0, N]$ and the final endpoint inequality (3.54) holds. Let $\binom{\alpha}{\underline{u}} \in \operatorname{Ker} \mathcal{M}_1 \left(\Phi_{N+1,0}(I - \mathcal{M}_0) \quad G_{N+1}\right)$ be arbitrary with $\underline{u} \in \mathbb{R}^{(N+1)n}$ and define $x_k := G_k \mathcal{T}_k \underline{u}$, $k \in [1, N+1]$ and $x_0 = \alpha$. Then (x, u) is admissible, by Lemma 1.35, and $\mathcal{M}_0 x_0 = 0$ and $\mathcal{M}_1 x_{N+1} = 0$.

Thus, by (ii), we have $x_k \in \operatorname{Im} X_k$ for all $k \in [0, N+1]$, that is, there exists $c_k \in \mathbb{R}^n$ such that $x_k = X_k c_k$ for all $k \in [0, N+1]$. It follows that

$$\begin{aligned}
\begin{pmatrix} \alpha \\ \underline{u} \end{pmatrix}^T \begin{pmatrix} (I - \mathcal{M}_0)\, \Phi_{k+1,0}^T \\ \mathcal{T}_{k+1}^T G_{k+1}^T \end{pmatrix} R[Q]_k \begin{pmatrix} \Phi_{k,0}(I - \mathcal{M}_0) & G_k \mathcal{T}_k \end{pmatrix} \begin{pmatrix} \alpha \\ \underline{u} \end{pmatrix} \\
= x_{k+1}^T R[Q]_k x_k = c_{k+1}^T X_{k+1}^T R[Q]_k X_k c_k = 0,
\end{aligned}$$

where we used $X_{k+1}^T R[Q]_k X_k = 0$ from Lemma 2.2.

(iii) $\Rightarrow$ (ii): Conversely, assume that Q_k satisfies the conditions in (iii). We show that image condition (3.51) holds. Therefore, let us take an admissible (x, u) with $\mathcal{M}_0\, x_0 = 0$ and $\mathcal{M}_1 x_{N+1} = 0$. Let $m \le N$ be any integer such that

$$x_k \in \operatorname{Im} X_k \quad \text{for all } k \le m. \tag{3.58}$$

This is always true for $m = 0$. To prove the image condition for all $k \in [0, N+1]$ it suffices to further show that (3.58) implies $x_{m+1} \in \operatorname{Im} X_{m+1}$.

Let $d \in \mathbb{R}^n$ be arbitrary and define $(\tilde{x}, \tilde{u})$ by

$$(\tilde{x}_k, \tilde{u}_k) := \begin{cases} (X_k, U_k)(I - X_{k+1}^\dagger X_{k+1})\, d, & \text{for } k \le m, \\ (0, 0), & \text{for } k > m. \end{cases}$$

Then $(\tilde{x}, \tilde{u})$ is admissible with $\mathcal{M}_0 \tilde{x}_0 = 0$ and $\mathcal{M}_1 \tilde{x}_{N+1} = 0$ and so is the pair $(x + \tilde{x}, u + \tilde{u})$. From equation (3.53) and Lemma 2.2 used with $(x + \tilde{x}, u + \tilde{u})$ we get $(x_{m+1} + \tilde{x}_{m+1})^T R[Q]_m (x_m + \tilde{x}_m) = 0$. From that and the definition of $\tilde{x}$ we have

$$\begin{aligned}
0 &= x_{m+1}^T R[Q]_m [x_m + X_m (I - X_{m+1}^\dagger X_{m+1})\, d] \\
&= x_{m+1}^T R[Q]_m x_m + x_{m+1}^T R[Q]_m X_m (I - X_{m+1}^\dagger X_{m+1})\, d.
\end{aligned}$$

The first term is zero, by equation (3.53) and Lemma 2.2 again, and we get

$$x_{m+1}^T R[Q]_m X_m (I - X_{m+1}^\dagger X_{m+1})\, d = 0 \quad \text{for all } d \in \mathbb{R}^n.$$

Now we use identity (2.8) and put $d = U_{m+1}^T x_{m+1}$, and get

$$\begin{aligned}
x_{m+1}^T U_{m+1} (I - X_{m+1}^\dagger X_{m+1}) X_m^\dagger X_m (I - X_{m+1}^\dagger X_{m+1}) U_{m+1}^T x_{m+1} &= 0, \\
X_m^\dagger X_m (I - X_{m+1}^\dagger X_{m+1}) U_{m+1}^T x_{m+1} &= 0, \\
X_m (I - X_{m+1}^\dagger X_{m+1}) U_{m+1}^T x_{m+1} &= 0. \qquad (3.59)
\end{aligned}$$

Because of (3.58), there exists c such that $x_m = X_m c$ and

$$\begin{aligned} U_{m+1}^T x_{m+1} &= U_{m+1}^T(\mathcal{A}_m x_m + \mathcal{B}_m u_m) = U_{m+1}^T[X_{m+1}c + \mathcal{B}_m(u_m - U_m c)] \\ &= U_{m+1}^T X_{m+1} c + (U_{m+1}^T \mathcal{B}_m - X_{m+1}^T \mathcal{D}_m + X_{m+1}^T \mathcal{D}_m)(u_m - U_m c) \\ &= X_{m+1}^T[U_{m+1}c + \mathcal{D}_m(u_m - U_m c)] - X_m^T(u_m - U_m c). \end{aligned} \tag{3.60}$$

Now from (3.59) and (3.60) we get

$$X_m(I - X_{m+1}^\dagger X_{m+1}) X_m^T(u_m - U_m c) = 0,$$

which implies

$$(I - X_{m+1}^\dagger X_{m+1}) X_m^T(u_m - U_m c) = 0.$$

Now we use (3.60) again to get x_{m+1} back and we obtain

$$(I - X_{m+1}^\dagger X_{m+1}) U_{m+1}^T x_{m+1} = 0. \tag{3.61}$$

Let $(\bar{X}, \bar{U})$ be such that $(\bar{X}, \bar{U})$, (X, U) are normalized conjoined bases of (S). We multiply equation (3.61) by $\bar{X}_{m+1}$ and get

$$\begin{aligned} 0 &= \bar{X}_{m+1}(I - X_{m+1}^\dagger X_{m+1}) U_{m+1}^T x_{m+1} \\ &= [\bar{X}_{m+1} U_{m+1}^T - X_{m+1} \bar{X}_{m+1}^T X_{m+1}^{\dagger T} U_{m+1}^T] x_{m+1} \\ &= [X_{m+1} \bar{U}_{m+1}^T + I - X_{m+1} \bar{X}_{m+1}^T X_{m+1}^{\dagger T} U_{m+1}^T] x_{m+1} \\ &= X_{m+1}[\bar{U}_{m+1}^T - \bar{X}_{m+1}^T X_{m+1}^{\dagger T} U_{m+1}^T] x_{m+1} + x_{m+1}, \end{aligned}$$

hence $x_{m+1} \in \operatorname{Im} X_{m+1}$ and this lemma is proven. □

Proof of Theorem 3.43. Implication (i) ⇒ (ii) follows from Lemmas 3.44, 3.45, 3.46. Implication (ii) ⇒ (i) follows from Lemma 3.47, while equivalence (ii) ⇔ (iii) is Lemma 3.48. □

3.2.5 Functional with general endpoints

In the statements of the following theorem we again use the conjoined basis $(\hat{X}^*, \hat{U}^*)$ of (S*), defined by (2.19) via the principal solution $(\hat{X}, \hat{U})$ and the associated solution $(\tilde{X}, \tilde{U})$ of (S), and the matrix $\hat{Q}^*$ defined by (2.33) with a symmetric matrix Q. Recall that the functional $\mathcal{F}$ is now defined by formula (1.38).

Theorem 3.49. *The following statements are equivalent.*

(i) $\mathcal{F}(x,u) \geq 0$ *over* $\mathcal{M}\left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right) = 0$.

(ii) *The principal solution* $(\hat{X}, \hat{U})$ *of* (S) *satisfies* P*-condition* (3.48), *the image condition*

$$\begin{gathered} x_k - \tilde{X}_k x_0 \in \operatorname{Im} \hat{X}_k \quad \text{for all } k \in [0, N+1], \\ \text{for all admissible } (x,u) \text{ with } \mathcal{M}\left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right) = 0, \end{gathered} \tag{3.62}$$

and the final endpoint inequality

$$\hat{X}^{*T}_{N+1}(\Gamma\, \hat{X}^*_{N+1} + \hat{U}^*_{N+1}) \geq 0 \quad \text{on } \operatorname{Ker} \mathcal{M} \hat{X}^*_{N+1}. \tag{3.63}$$

(iii) *The implicit Riccati equation*

$$\begin{gathered} \begin{pmatrix} \Phi^T_{k+1,0} - \tilde{X}^T_{k+1} \\ \mathcal{T}^T_{k+1} G^T_{k+1} \end{pmatrix} R[Q]_k \begin{pmatrix} \Phi_{k,0} - \tilde{X}_k & G_k \mathcal{T}_k \end{pmatrix} = 0 \\ \text{on } \operatorname{Ker} \mathcal{M} \begin{pmatrix} I & 0 \\ \Phi_{N+1,0} & G_{N+1} \end{pmatrix}, \quad k \in [0, N], \end{gathered} \tag{3.64}$$

has a symmetric solution Q_k *such that* $Q_k \hat{X}_k = \hat{U}_k \hat{X}^\dagger_k \hat{X}_k$ *on* $[0, N+1]$, $\mathcal{P}$*-condition* (3.50) *holds, and satisfying the final endpoint inequality*

$$\Gamma + \hat{Q}^*_{N+1} \geq 0 \quad \text{on } \operatorname{Ker} \mathcal{M} \cap \operatorname{Im} \hat{X}^*_{N+1}. \tag{3.65}$$

As in the proof of Theorem 3.24, we again use the transformation of the quadratic functional $\mathcal{F}$ and system (S) into the augmented functional $\mathcal{F}^*$ and system (S*).

Proof. The proof of (i) $\Leftrightarrow$ (ii) is similar to the proof of the equivalence of the corresponding conditions for the positivity (Theorem 3.24), we use the last identity in (3.47) and Lemma 2.24. The proof of (i) $\Rightarrow$ (iii) is similar to the proof of Lemma 3.48 in which we replace the natural conjoined basis (X, U) by the principal solution $(\hat{X}, \hat{U})$, and instead of image condition (3.51) we use image condition (3.62). The symmetric matrix Q_k satisfying the conditions in (iii) is defined by equation (2.2) with $(\hat{X}, \hat{U})$ in place of (X, U). Implication (iii) $\Rightarrow$ (i) is again proven similarly as in Lemma 3.48, but with the image condition $x_k - \tilde{X}_k x_0 \in \operatorname{Im} \hat{X}_k$ for all $k \in [0, m]$, and with the vector $d := \hat{U}^T_{m+1}(x_{m+1} - \tilde{X}_{m+1} x_0)$. $\square$

Remark 3.50. Equivalence (i) $\Leftrightarrow$ (ii) in Theorem 3.49 can be again proven similarly as in the alternative proof of equivalence (i) $\Leftrightarrow$ (ii) in Theorem 3.24, see page 62, but here we have to prove the necessity of image condition (3.62) with the use of the augmented functional $\mathcal{F}^*$, because image condition (3.49) for zero enpoints does not imply image condition (3.62).

3.3 Implicit Riccati equations

In this section we collect and compare various forms of implicit Riccati equations. In the case of zero and separated endpoints, we have only one form for the positivity and one form for the nonnegativity. In the case of general endpoints, three equivalent forms of implicit Riccati equations are possible for the positivity and two for the nonnegativity. One of them (for the positivity) involves the augmented Riccati operator $R^*[Q^*]_k$, in the others only the original Riccati operator $R[Q]_k$ in dimension n appears.

For convenience, we denote in this section the $\mathcal{P}$-conditions as follows

$$\mathcal{P}_k = \mathcal{B}_k^T \mathcal{D}_k - \mathcal{B}_k^T Q_{k+1} \mathcal{B}_k \geq 0 \quad \text{for all } k \in [0, N], \tag{$\mathcal{P}$}$$

$$T_k \mathcal{P}_k T_k \geq 0 \quad \text{for all } k \in [0, N], \tag{$T\mathcal{P}T$}$$

where the matrix T is defined in (3.37) through a conjoined basis (X, U) specified in the corresponding statements. Further, $\hat{Q}^*$ is again the matrix defined by (2.33) with a symmetric matrix Q and the associated solution $(\tilde{X}, \tilde{U})$ of (S), and $(\hat{X}^*, \hat{U}^*)$ is the conjoined basis of (S*) defined by (2.19).

First we display implicit Riccati equations for the positivity.

Positivity, zero endpoints. *The implicit Riccati equation*

$$R[Q]_k G_k = 0, \quad k \in [0, N],$$

has a symmetric solution Q_k on $[0, N+1]$ such that condition ($\mathcal{P}$) holds.

Positivity, separated endpoints. *The implicit Riccati equation*

$$R[Q]_k \left(\Phi_{k,0}(I - \mathcal{M}_0) \quad G_k \right) = 0, \quad k \in [0, N],$$

has a symmetric solution Q_k on $[0, N+1]$ such that $Q_0 = \Gamma_0$, condition ($\mathcal{P}$) holds, and satisfying

$$\Gamma_1 + Q_{N+1} > 0 \quad \textit{on } \operatorname{Ker} \mathcal{M}_1 \cap \operatorname{Im} X_{N+1}.$$

Positivity, general endpoints.

(i) *The implicit Riccati equation*

$$R[Q]_k G_k = 0, \quad k \in [0, N], \tag{3.66}$$

has a symmetric solution Q_k *on* $[0, N+1]$ *such that condition* $(\mathcal{P})$ *holds, and* $\hat{Q}^*_{N+1}$ *satisfies the final endpoint inequality*

$$\Gamma + \hat{Q}^*_{N+1} > 0 \quad \text{on } \operatorname{Ker} \mathcal{M} \cap \operatorname{Im} \hat{X}^*_{N+1}. \tag{3.67}$$

(ii) *The implicit Riccati equation*

$$R[Q]_k \begin{pmatrix} \Phi_{k,0} - \tilde{X}_k & G_k \end{pmatrix} = 0, \quad k \in [0, N], \tag{3.68}$$

has a symmetric solution Q_k *on* $[0, N+1]$ *such that* $Q_0 = 0$*, condition* $(\mathcal{P})$ *holds, and* $\hat{Q}^*_{N+1}$ *satisfies final endpoint inequality* (3.67)*.*

(iii) *The implicit Riccati equation*

$$R^*[Q^*]_k \begin{pmatrix} I & 0 \\ \Phi_{k,0} & G_k \end{pmatrix} = 0, \quad k \in [0, N], \tag{3.69}$$

has a symmetric solution $Q^*_k = \left(\begin{smallmatrix} \star & \star \\ \star & Q_k \end{smallmatrix}\right)$ *on* $[0, N+1]$ *such that* $Q^*_0 = 0$ *and* $(\mathcal{P})$ *holds, and satisfying the final endpoint inequality* (3.67) *with* Q^*_{N+1} *instead of* $\hat{Q}^*_{N+1}$*.*

Remark 3.51. The above implicit Riccati equations for the positivity can be replaced by the following weaker forms.

Positivity, zero endpoints.

$$R[Q]_k G_k \mathcal{T}_k = 0 \quad \text{on } \operatorname{Ker} G_{N+1}, \quad k \in [0, N].$$

Positivity, separated endpoints.

$$\begin{aligned} R[Q]_k \begin{pmatrix} \Phi_{k,0}(I - \mathcal{M}_0) & G_k \mathcal{T}_k \end{pmatrix} &= 0 \\ \text{on } \operatorname{Ker} \mathcal{M}_1 \begin{pmatrix} \Phi_{N+1,0}(I - \mathcal{M}_0) & G_{N+1} \end{pmatrix}, &\quad k \in [0, N]. \end{aligned}$$

Positivity, general endpoints.

$$R^*[Q^*]_k \begin{pmatrix} I & 0 \\ \Phi_{k,0} & G_k \mathcal{T}_k \end{pmatrix} = 0 \quad \text{on } \operatorname{Ker} \mathcal{M} \begin{pmatrix} I & 0 \\ \Phi_{N+1,0} & G_{N+1} \end{pmatrix}, \quad k \in [0, N].$$

These forms of implicit Riccati equation appear e.g. in [40].

Next, we display implicit Riccati equations for the nonnegativity.

Nonnegativity, zero endpoints. *The implicit Riccati equation*

$$\mathcal{T}_{k+1}^T G_{k+1}^T R[Q]_k G_k \mathcal{T}_k = 0 \quad \text{on } \operatorname{Ker} G_{N+1}, \quad k \in [0, N], \tag{3.70}$$

has a symmetric solution Q_k *such that* $Q_k \hat{X}_k = \hat{U}_k \hat{X}_k^{\dagger} \hat{X}_k$ *on* $[0, N+1]$, *and* $(T\mathcal{P}T)$ *holds with* T_k *defined via* $(\hat{X}, \hat{U})$.

Nonnegativity, separated endpoints. *The implicit Riccati equation*

$$\begin{pmatrix} (I - \mathcal{M}_0)\, \Phi_{k+1,0}^T \\ \mathcal{T}_{k+1}^T G_{k+1}^T \end{pmatrix} R[Q]_k \begin{pmatrix} \Phi_{k,0}(I - \mathcal{M}_0) & G_k \mathcal{T}_k \end{pmatrix} = 0$$
$$\text{on } \operatorname{Ker} \mathcal{M}_1 \begin{pmatrix} \Phi_{N+1,0}(I - \mathcal{M}_0) & G_{N+1} \end{pmatrix}, \quad k \in [0, N], \tag{3.71}$$

has a symmetric solution Q_k *such that* $Q_k X_k = U_k X_k^{\dagger} X_k$ *on* $[0, N+1]$, *and* $(T\mathcal{P}T)$ *holds with* T_k *defined via the natural conjoined basis* (X, U), *and satisfying the final endpoint inequality*

$$Q_{N+1} + \Gamma_1 \geq 0 \quad \text{on } \operatorname{Ker} \mathcal{M}_1 \cap \operatorname{Im} X_{N+1}. \tag{3.72}$$

Nonnegativity, general endpoints.

(i') *The implicit Riccati equation*

$$\mathcal{T}_{k+1}^T G_{k+1}^T R[Q]_k G_k \mathcal{T}_k = 0$$
$$\text{on } \left\{ \underline{u} \in \mathbb{R}^{(N+1)n} :\ \mathcal{M} \begin{pmatrix} 0 \\ G_{N+1} \end{pmatrix} \underline{u} \in \operatorname{Im} \mathcal{M} \begin{pmatrix} I \\ \tilde{X}_{N+1} \end{pmatrix} \right\}, \quad k \in [0, N], \tag{3.73}$$

has a symmetric solution Q_k *such that* $Q_k \hat{X}_k = \hat{U}_k \hat{X}_k^{\dagger} \hat{X}_k$ *on* $[0, N+1]$, $(T\mathcal{P}T)$ *holds with* T *defined by* $(\hat{X}, \hat{U})$, *and* $\hat{Q}_{N+1}^*$ *satisfies the final endpoint inequality*

$$\Gamma + \hat{Q}_{N+1}^* \geq 0 \quad \text{on } \operatorname{Ker} \mathcal{M} \cap \operatorname{Im} \hat{X}_{N+1}^*. \tag{3.74}$$

(ii') *The implicit Riccati equation*

$$\begin{pmatrix} \Phi_{k+1,0}^T - \tilde{X}_{k+1}^T \\ \mathcal{T}_{k+1}^T G_{k+1}^T \end{pmatrix} R[Q]_k \begin{pmatrix} \Phi_{k,0} - \tilde{X}_k & G_k \mathcal{T}_k \end{pmatrix} = 0$$
$$\text{on } \operatorname{Ker} \mathcal{M} \begin{pmatrix} I & 0 \\ \Phi_{N+1,0} & G_{N+1} \end{pmatrix}, \quad k \in [0, N], \tag{3.75}$$

has a symmetric solution Q_k such that $Q_k\hat{X}_k = \hat{U}_k\hat{X}_k^{\dagger}\hat{X}_k$ on $[0, N+1]$, $(T\mathcal{P}T)$ holds with T defined via $(\hat{X}, \hat{U})$, and $\hat{Q}^*_{N+1}$ satisfies final endpoint inequality (3.74)*.*

Remark 3.52. Similarly as in case of the positivity, we could furthermore formulate a statement analogous to (iii).

(iii') *The implicit Riccati equation*

$$\begin{pmatrix} I & \Phi^T_{k+1,0} \\ 0 & \mathcal{T}^T_{k+1}G^T_{k+1} \end{pmatrix} R^*[Q^*]_k \begin{pmatrix} I & 0 \\ \Phi_{k,0} & G_k\mathcal{T}_k \end{pmatrix} = 0 \quad \text{on } \operatorname{Ker}\mathcal{M}\begin{pmatrix} I & 0 \\ \Phi_{N+1,0} & G_{N+1} \end{pmatrix}, \quad k \in [0, N], \tag{3.76}$$

has a symmetric solution $Q^*_k = \left(\begin{smallmatrix} \star & \star \\ \star & Q_k \end{smallmatrix}\right)$ such that $Q^*_k\hat{X}^*_k = \hat{U}^*_k\hat{X}^{*\dagger}_k\hat{X}^*_k$ on $[0, N+1]$, and $(T\mathcal{P}T)$ holds with T_k defined via $(\hat{X}, \hat{U})$, and satisfying the final endpoint inequality (3.74) *with Q^*_{N+1} instead of $\hat{Q}^*_{N+1}$.*

But, by Remark 2.18, this solution Q^* must have the form (2.22), and thus statement (iii') would say the same as statement (ii').

Remark 3.53. The condition $Q_k\hat{X}_k = \hat{U}_k\hat{X}_k^{\dagger}\hat{X}_k$ (or $Q_kX_k = U_kX_k^{\dagger}X_k$) that does not appear in statements for positivity cannot be removed from statements for nonnegativity. See Example 3.57 at the end of this section. However, we do not know, whether this condition can be replaced by a weaker one, e.g. by the condition $\hat{X}_k^TQ_k\hat{X}_k = \hat{U}_k^T\hat{X}_k$.

In the remaining part of this section, conditions (i)–(iii) and (i')–(iii') refer to implicit Riccati equations displayed on pages 79–80.

Theorem 3.54. *Statements (i) – (iii) are equivalent.*

Theorem 3.55. *Statements (i') and (ii') are equivalent.*

Both theorems are proven independently of the roundabout theorems (Theorem 3.24 and Theorem 3.49), except of the final endpoint inequality in implication (iii) $\Rightarrow$ (i). In proofs we use the next auxiliary lemma.

Lemma 3.56. *If $R[Q]_kG_k\mathcal{T}_k = 0$, then $R[Q]_k\begin{pmatrix}\tilde{X}_k - \Phi_{k,0} & -G_k\mathcal{T}_k\end{pmatrix} = 0$.*

Proof. Let $\binom{\alpha}{\underline{u}}$ be arbitrary $n+(N+1)n$-vector, where $\underline{u} := \begin{pmatrix} u_0 \\ \vdots \\ u_N \end{pmatrix}$. Define a pair $(\tilde{x}, \tilde{u})$ by

$$\begin{aligned} \tilde{x}_k &:= \tilde{X}_k\,\alpha - \Phi_{k,0}\,\alpha - G_k \mathcal{T}_k\,\underline{u}, \qquad && k \in [0, N+1], \\ \tilde{u}_k &:= \tilde{U}_k\,\alpha - u_k, && k \in [0, N]. \end{aligned} \tag{3.77}$$

This $(\tilde{x}, \tilde{u})$ is admissible, because

$$\begin{aligned} \mathcal{A}_k \tilde{x}_k + \mathcal{B}_k \tilde{u}_k &= \mathcal{A}_k \tilde{X}_k\,\alpha - \Phi_{k+1,0}\,\alpha - \mathcal{A}_k G_k \mathcal{T}_k\,\underline{u} - \mathcal{B}_k u_k + \mathcal{B}_k \tilde{U}_k\,\alpha \\ &= \tilde{X}_{k+1}\,\alpha - \Phi_{k+1,0}\,\alpha - G_{k+1}\mathcal{T}_{k+1}\,\underline{u} = \tilde{x}_{k+1}. \end{aligned}$$

Further, $\tilde{x}_0 = \tilde{X}_0 \alpha - \Phi_{0,0}\,\alpha - G_0 \mathcal{T}_0\,\underline{u} = 0$, and Lemma 1.29 implies that $\tilde{x} = G_k \mathcal{T}_k\,\underline{\tilde{u}}$, where $\underline{\tilde{u}} = \begin{pmatrix} \tilde{U}_0\,\alpha - u_0 \\ \vdots \\ \tilde{U}_N\,\alpha - u_N \end{pmatrix} = \underline{\tilde{U}}\,\alpha - \underline{u}$. We have

$$R[Q]_k \begin{pmatrix} \tilde{X}_k - \Phi_{k,0} & -G_k \mathcal{T}_k \end{pmatrix} \begin{pmatrix} \alpha \\ \underline{u} \end{pmatrix} = R[Q]_k \tilde{x}_k = R[Q]_k G_k \mathcal{T}_k\,\underline{\tilde{u}} = 0.$$

Thus, this lemma is proven. $\square$

Proof of Theorem 3.54. Implication (ii) $\Rightarrow$ (i) holds trivially, and (i) $\Rightarrow$ (ii) holds by Lemma 3.56.

(ii)$\Rightarrow$ (iii): From (3.68) we have

$$R[Q]_k \begin{pmatrix} -\tilde{X}_k & I \end{pmatrix} \begin{pmatrix} I & 0 \\ \Phi_{k,0} & G_k \end{pmatrix} = 0, \quad k \in [0, N],$$

which implies

$$\begin{pmatrix} -\tilde{X}_{k+1}^T \\ I \end{pmatrix} R[Q]_k \begin{pmatrix} -\tilde{X}_k & I \end{pmatrix} \begin{pmatrix} I & 0 \\ \Phi_{k,0} & G_k \end{pmatrix} = 0, \quad k \in [0, N],$$

and thus, by identity (2.25), we get $R^*[\hat{Q}^*]_k \left(\begin{smallmatrix} I & 0 \\ \Phi_{k,0} & G_k \end{smallmatrix}\right) = 0$. Further, $Q_0 = 0$ and thus $\hat{Q}_0^* = 0$.

(iii)$\Rightarrow$ (i): Let $Q_k^* = \left(\begin{smallmatrix} \star & \star \\ \star & Q_k \end{smallmatrix}\right)$ be a solution of (3.69) satisfying the given conditions. From the definition of $R^*[Q^*]$ we have

$$\begin{aligned} R^*[Q^*]_k = \begin{pmatrix} \star & \star \\ \star & Q_{k+1} \end{pmatrix} \left[\begin{pmatrix} I & 0 \\ 0 & \mathcal{A}_k \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{B}_k \end{pmatrix} \begin{pmatrix} \star & \star \\ \star & Q_k \end{pmatrix} \right] \\ - \left[\begin{pmatrix} 0 & 0 \\ 0 & \mathcal{C}_k \end{pmatrix} + \begin{pmatrix} I & 0 \\ 0 & \mathcal{D}_k \end{pmatrix} \begin{pmatrix} \star & \star \\ \star & Q_k \end{pmatrix} \right] = \begin{pmatrix} \star & \star \\ \star & R[Q]_k \end{pmatrix}. \end{aligned} \tag{3.78}$$

We multiply equation (3.69) by the matrix $\left(\begin{smallmatrix}0\\ I\end{smallmatrix}\right)$ from the right and get

$$\begin{pmatrix} \star & \star \\ \star & R[Q]_k \end{pmatrix} \begin{pmatrix} 0 \\ G_k \end{pmatrix} = 0,$$

which implies $R[Q]_k G_k = 0$. Now it remains to show the final endpoint inequality. By Theorem 3.24, we have that inequality (3.21) holds, which is equivalent to (3.67). □

Proof of Theorem 3.55. (i')⇒ (ii'): Similarly as in the proof of Lemma 3.56, we take arbitrary $n+(N+1)n$-vector $\left(\begin{smallmatrix}x_0\\ \underline{u}\end{smallmatrix}\right)$ such that $\mathcal{M}\left(\begin{smallmatrix}I & 0\\ \Phi_{N+1,0} & G_{N+1}\end{smallmatrix}\right)\left(\begin{smallmatrix}x_0\\ \underline{u}\end{smallmatrix}\right) = 0$, and define $(\tilde{x}, \tilde{u})$ by (3.77). We get

$$\begin{aligned}\begin{pmatrix} x_0^T & \underline{u}^T \end{pmatrix} \begin{pmatrix} \Phi_{k+1,0}^T - \tilde{X}_{k+1}^T \\ \mathcal{T}_{k+1}^T G_{k+1}^T \end{pmatrix} R[Q]_k \begin{pmatrix} \Phi_{k,0} - \tilde{X}_k & G_k \mathcal{T}_k \end{pmatrix} \begin{pmatrix} x_0 \\ \underline{u} \end{pmatrix} \\ = \underline{\tilde{u}}^T \mathcal{T}_{k+1}^T G_{k+1}^T R[Q]_k G_k \mathcal{T}_k \underline{\tilde{u}}.\end{aligned}$$

Since $\mathcal{M}\left(\begin{smallmatrix}x_0\\ \Phi_{N+1,0}x_0 + G_{N+1}\underline{u}\end{smallmatrix}\right) = 0$, we have $\mathcal{M}\left(\begin{smallmatrix}0\\ G_{N+1}\end{smallmatrix}\right)\underline{\tilde{u}} = \mathcal{M}\begin{pmatrix} I \\ \tilde{X}_{N+1}\end{pmatrix} x_0$, and hence, equation (3.73) implies $\underline{\tilde{u}}^T \mathcal{T}_{k+1}^T G_{k+1}^T R[Q]_k G_k \mathcal{T}_k \underline{\tilde{u}} = 0$.

(ii')⇒ (i'): Let $\underline{u} \in \mathbb{R}^{(N+1)n}$ and $\mathcal{M}\left(\begin{smallmatrix}0\\ G_{N+1}\end{smallmatrix}\right)\underline{u} = \mathcal{M}\begin{pmatrix} I \\ \tilde{X}_{N+1}\end{pmatrix}\alpha$. We define $\underline{\tilde{u}} := \underline{\tilde{U}}\alpha - \underline{u}$, where $\underline{\tilde{U}} := \begin{pmatrix}\tilde{U}_0 \\ \vdots \\ \tilde{U}_N\end{pmatrix}$. Then $\mathcal{M}\left(\begin{smallmatrix}I & 0\\ \Phi_{N+1,0} & G_{N+1}\end{smallmatrix}\right)\left(\begin{smallmatrix}\alpha\\ \underline{\tilde{u}}\end{smallmatrix}\right) = 0$ and $(\Phi_{k,0} - \tilde{X}_k)\alpha + G_k\mathcal{T}_k \underline{tu} = \mathcal{G}_k \mathcal{T}_k \underline{u}$. Note that $\Phi_{k,0}\alpha + G_k \mathcal{T}_k \underline{\tilde{U}}\alpha = \tilde{X}_k\alpha$, by Lemma 1.29. Hence, from equation (3.75) we get equation (3.73). □

Finally, we show an example of $\mathcal{F}_0 \ngeq 0$ such that there is a symmetric solution Q_k on $[0, N+1]$ of equation (3.70) satisfying condition $(T\mathcal{P}T)$ and not satisfying $Q_k \hat{X}_k = \hat{U}_k \hat{X}_k^\dagger \hat{X}_k$ on $[0, N+1]$.

Example 3.57. Let $n = 1$, $N = 3$, and $\mathcal{S}_k \equiv \left(\begin{smallmatrix}0 & 1\\ -1 & 1\end{smallmatrix}\right)$ for $k \in [0,3]$, i.e. $\mathcal{A}_k \equiv 0$ and $\mathcal{B}_k = \mathcal{D}_k = -\mathcal{C}_k \equiv 1$. Assume that both endpoints are zero. Then the functional $\mathcal{F}_0$ takes the form $\mathcal{F}_0(x,u) = \sum_{k=0}^{3}\{u_k^2 - 2\,x_k u_k\}$ over pairs (x,u) satisfying $x_{k+1} = u_k$ for $k \in [0,3]$ and $x_0 = 0 = x_4$. The principal solution $(\hat{X}, \hat{U})$ of (S) is in this case $\hat{X} = \{0,1,1,0,-1\} = \hat{X}^\dagger$ and $\hat{U} = \{1,1,0,-1,-1\}$. Then $\mathcal{F}_0 \ngeq 0$, since $\mathcal{F}(\hat{x},\hat{u}) = -1$ for the admissible pair $(\hat{x},\hat{u})$ defined by $\hat{x} := \{0,1,1,1,0\}$ and $\hat{u} := \{1,1,1,0\}$. Note that $\hat{x}_3 \notin \operatorname{Im}\hat{X}_3$, i.e. image condition (3.49) is violated.

Define $Q := \{2, \frac{1}{2}, -1, 2, 0\}$. Then Q_k satisfies equation (3.70) and condition $(T\mathcal{P}T)$, and does not satisfy that $Q_k \hat{X}_k = \hat{U}_k \hat{X}_k^\dagger \hat{X}_k$ for $k \in [0,4]$. This can be verified when we calculate the sequences $R[Q] = \{0,0,0,-1\}$, $\mathcal{P} = \{\frac{1}{2}, 2, -1, 1\}$, $M = \{0,0,1,0\} = M^\dagger$, $T = \{1,1,0,1\}$, $T\mathcal{P}T = \{\frac{1}{2}, 2, 0, 1\} \geq 0$, $G_3 = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$, and $G_4 = \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix}$.

3.4 Perturbation of quadratic functionals

In this subsection we present theorems that say that the nonnegativity of a functional with zero, separated, and general endpoints, respectively, on admissible pairs (x,u) with corresponding boundary conditions is equivalent to the nonnegativity of a perturbed functional on admissible pairs (x,u) such that $\binom{x_0}{x_{N+1}}$ is restricted to a (larger) subspace. If the (nonperturbed) functional is positive then this restriction holds for all admissible pairs, thus corresponding results regarding the positivity say that the positivity of a functional on admissible pairs (x,u) with given boundary conditions is equivalent to the positivity of a perturbed functional on all admissible pairs (x,u) with $x \not\equiv 0$.

3.4.1 Functional with zero endpoints

Again, recall that $(\hat{X}, \hat{U})$ is the principal solution of (S), i.e. $(\hat{X}_0, \hat{U}_0) = (0, I)$, and $(\tilde{X}, \tilde{U})$ is the associated solution of (S), i.e. $(\tilde{X}_0, \tilde{U}_0) = (I, 0)$. The functional $\mathcal{F}_0$ is defined by formula (1.30).

Theorem 3.58. *The following statements are equivalent.*

(i) $\mathcal{F}_0(x,u) \geq 0$ *over* $x_0 = 0 = x_{N+1}$.

(ii) *There exist* $\alpha > 0$ *and* $\beta > 0$ *such that*

$$\mathcal{F}_0(x,u) + \alpha \|x_0\|^2 + \beta \|x_{N+1}\|^2 \geq 0 \quad \text{over} \quad x_{N+1} - \tilde{X}_{N+1} x_0 \in \operatorname{Im} \hat{X}_{N+1}.$$

(iii) *There exists* $\alpha > 0$ *such that*

$$\mathcal{F}_0(x,u) + \alpha \|x_0\|^2 \geq 0 \quad \text{over} \quad \tilde{X}_{N+1} x_0 \in \operatorname{Im} \hat{X}_{N+1}, \; x_{N+1} = 0.$$

(iv) *There exists* $\beta > 0$ *such that*

$$\mathcal{F}_0(x,u) + \beta \|x_{N+1}\|^2 \geq 0 \quad \text{over} \quad x_{N+1} \in \operatorname{Im} \hat{X}_{N+1}, \; x_0 = 0.$$

(v) *There exists* $\alpha > 0$ *such that*

$$\mathcal{F}_0(x,u) + \alpha \|x_0\|^2 \geq 0 \quad over \quad \tilde{X}_{N+1}x_0 = x_{N+1}.$$

(vi) $\mathcal{F}_0(x,u) \geq 0$ *over* $\tilde{X}_{N+1}x_0 = x_{N+1} = 0$.

Proof. Conditions (ii)–(vi) imply condition (i) trivially, and condition (ii) implies (iii) and (iv). It remains to prove that (i) implies (ii), (v) and (vi).

(i) $\Rightarrow$ (ii): Let (x,u) be an admissible pair with $x_{N+1} - \tilde{X}_{N+1}x_0 \in \operatorname{Im}\hat{X}_{N+1}$. Then, by Theorem 2.31, we have that the identity

$$\mathcal{F}_0(x,u) = \begin{pmatrix} x_0 \\ x_{N+1} \end{pmatrix}^T Q^* \begin{pmatrix} x_0 \\ x_{N+1} \end{pmatrix} + \mathcal{F}_0(\bar{x},\bar{u}), \tag{2.39}$$

holds, where $(\bar{x},\bar{u})$ is an admissible pair with $\bar{x}_0 = \bar{x}_{N+1} = 0$ and Q^* is a symmetric $2n \times 2n$ matrix. Let λ_0 be the smallest eigenvalue of Q^*. Then from (2.39) and inequality (1.8) we get

$$\mathcal{F}_0(x,u) \geq \lambda_0(\|x_0\|^2 + \|x_{N+1}\|^2) \geq -|\lambda_0|(\|x_0^2\| + \|x_{N+1}\|^2).$$

Hence, the inequality in (ii) holds for any $\alpha, \beta \geq |\lambda_0|$.

(i) $\Rightarrow$ (v): If $\tilde{X}_{N+1}x_0 = x_{N+1}$, then, by Corollary 2.33, we get $\mathcal{F}_0(x,u) \geq x_0^T \tilde{X}_{N+1}^T \tilde{U}_{N+1}x_0$. We use again condition (1.8) and get that the inequality in (v) holds for any $\alpha \geq |\lambda_1|$, where λ_1 is the smallest eigenvalue of $\tilde{X}_{N+1}^T \tilde{U}_{N+1}$.

(i) $\Rightarrow$ (vi): As in previous case, we have $\mathcal{F}_0(x,u) \geq x_0^T \tilde{X}_{N+1}^T \tilde{U}_{N+1}x_0$, and $\tilde{X}_{N+1}x_0 = 0$ further implies $\mathcal{F}_0(x,u) \geq 0$. □

Theorem 3.59. *The following statements are equivalent.*

(i') $\mathcal{F}_0(x,u) > 0$ *over* $x_0 = 0 = x_{N+1}$, $x \not\equiv 0$.

(ii') *There exist* $\alpha > 0$ *and* $\beta > 0$ *such that*

$$\mathcal{F}_0(x,u) + \alpha \|x_0\|^2 + \beta \|x_{N+1}\|^2 > 0 \quad over \quad x \not\equiv 0.$$

(iii') *There exists* $\alpha > 0$ *such that*

$$\mathcal{F}_0(x,u) + \alpha \|x_0\|^2 > 0 \quad over \quad x_{N+1} = 0,\ x \not\equiv 0.$$

(iv') *There exists* $\beta > 0$ *such that*

$$\mathcal{F}_0(x,u) + \beta \|x_{N+1}\|^2 > 0 \quad over \quad x_0 = 0,\ x \not\equiv 0.$$

(v') *There exists* $\alpha > 0$ *such that*

$$\mathcal{F}_0(x,u) + \alpha \|x_0\|^2 > 0 \quad over \quad \tilde{X}_{N+1}x_0 = x_{N+1}, \ x \not\equiv 0.$$

(vi') $\mathcal{F}_0(x,u) > 0$ *over* $\tilde{X}_{N+1}x_0 = x_{N+1} = 0, \ x \not\equiv \tilde{X}x_0$.

Proof. Again, conditions (ii')–(vi') imply condition (i') trivially, and condition (ii') implies (iii') and (iv'). It remains to prove that (i') implies (ii'), (v') and (vi'). Conditions (i)–(vi) refer in this proof to Theorem 3.58.

(i') $\Rightarrow$ (ii'): First note that (i') $\Rightarrow$ (i) $\Rightarrow$(ii). From Lemma 3.5 we have that $\operatorname{Ker} \hat{X}_{k+1} \subseteq \operatorname{Ker} \hat{X}_k$ for all $k \in [0,N]$ and hence, by Corollary 2.34, identity (2.39) holds for all admissible (x,u). Thus, we get that there exist $\bar{\alpha} > 0$ and $\bar{\beta} > 0$ such that $\mathcal{F}_0(x,u) + \bar{\alpha}\|x_0\|^2 + \bar{\beta}\|x_{N+1}\|^2 \geq 0$ for all admissible (x,u). Now, for $\alpha := \bar{\alpha}+1$ and $\beta := \bar{\beta}+1$ we have $\mathcal{F}_0(x,u) + \alpha\|x_0\|^2 + \beta\|x_{N+1}\|^2 \geq \|x_0\|^2 + \|x_{N+1}\|^2$, and $\|x_0\|^2 + \|x_{N+1}\|^2 > 0$ except when $x_0 = 0 = x_{N+1}$. But if $x_0 = 0 = x_{N+1}$, then, since $x \not\equiv 0$, condition (i') directly implies that the inequality in (ii') holds for such (x,u) with any α, β.

(i') $\Rightarrow$ (v'): We use (i') $\Rightarrow$ (i) $\Rightarrow$ (v) and get that there exists $\bar{\alpha} > 0$ such that $\mathcal{F}_0(x,u) + \bar{\alpha}\|x_0\|^2 \geq 0$ over $\tilde{X}_{N+1}x_0 = x_{N+1}$. As in the previous case, for $\alpha := \bar{\alpha} + 1$ we get $\mathcal{F}_0(x,u) + \alpha\|x_0\|^2 \geq \|x_0\|^2$, and as $\tilde{X}_{N+1}x_0 = x_{N+1}$, $\|x_0\|^2 > 0$ except when $x_0 = 0 = x_{N+1}$. And if $x_0 = 0 = x_{N+1}$, then, since $x \not\equiv 0$, condition (i') directly implies that the inequality in (ii') holds for such (x,u) with any α.

(i') $\Rightarrow$ (vi') Define $(\bar{x},\bar{u})$ as in (2.42). Then $\bar{x} \not\equiv 0$, because $x \not\equiv \tilde{X}x_0$, and from (2.39) we get $\mathcal{F}_0(x,u) = \mathcal{F}_0(\bar{x},\bar{u}) > 0$. ☐

Remark 3.60. Another way of proving (i') $\Rightarrow$ (ii') is via generalized Picone identity (2.36), where we take $\alpha = x_0$ and the normalized conjoined bases $(\tilde{X},\tilde{U})$ and $(\hat{X},\hat{U})$, and get

$$\mathcal{F}_0(x,u) = \begin{pmatrix} x_0 \\ x_k \end{pmatrix}^T \hat{Q}_k^* \begin{pmatrix} x_0 \\ x_k \end{pmatrix} \Bigg|_0^{N+1} + \sum_{k=0}^{N} \tilde{w}_k^T \mathcal{P}_k \tilde{w}_k,$$

with $\hat{Q}^*$ defined by (2.33), with Q such that $Q_0 = 0$ and $Q_k\hat{X}_k = \hat{U}_k\hat{X}_k^\dagger\hat{X}_k$ on $[0,N+1]$. As $\mathcal{P}_k \geq 0$, by Lemma 3.5, we get $\mathcal{F}_0(x,u) \geq \left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right)^T \hat{Q}_{N+1}^* \left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right)$.

Remark 3.61. Yet another way of proof of (i') $\Rightarrow$ (ii') uses results for the positivity of functional with general endpoints. More specifically, we consider the functional with the matrix $\Gamma := \left(\begin{smallmatrix} \alpha & 0 \\ 0 & \beta \end{smallmatrix}\right)$ and the matrix $\mathcal{M} := 0$ and show that there exist $\alpha, \beta > 0$ such that condition (3.21) holds.

Remark 3.62. We can see that in case of the nonnegativity of $\mathcal{F}_0$, the endpoints x_0 and x_{N+1} cannot be free, but must be restricted to a subspace. This is also shown in the following example where $\mathcal{F}_0(x,u) \geq 0$ over $x_0 = 0 = x_{N+1}$, but there is no $\alpha, \beta > 0$ such that $\mathcal{F}(x,u) = \alpha\,\|x_0\|^2 + \beta\,\|x_{N+1}\|^2 + \mathcal{F}_0(x,u) \geq 0$.

Example 3.63. Consider the coefficients $\mathcal{S}_k \equiv \mathcal{J}$, that is, $\mathcal{A}_k = \mathcal{D}_k \equiv 0$ and $\mathcal{B}_k = -\mathcal{C}_k \equiv I$ for all $k \in [0, N]$. Then the solution $\tilde{X}_k$ is

$$\{\tilde{X}_k\}_{k=0}^{N+1} = \{I, 0, -I, 0, I, 0, -I, 0, \dots\},$$

and the functional $\mathcal{F}_0$ takes the form

$$\mathcal{F}_0(x,u) = -2\left\{x_0^T u_0 + \sum_{k=1}^{N} u_{k-1}^T u_k\right\}$$

for admissible (x,u), i.e. $x_{k+1} = u_k$ on $[0,N]$.

If we take $N = 1$, then

$$\mathcal{F}_0(x,u) = -2\,x_0^T u_0 - 2\,u_0^T u_1$$

for admissible (x,u) and, in particular, $\mathcal{F}_0(x,u) = 0 (\geq 0)$ when $x_0 = x_2 = 0$. Note that in this case $\mathcal{F}_0$ is not positive definite. On the other hand,

$$\mathcal{F}(x,u) = \alpha\,\|x_0\|^2 + \beta\,\|x_{N+1}\|^2 - 2\,x_0^T u_0 - 2\,u_0^T u_1 \not\geq 0$$

over x_0 and x_2 free, which follows for example by choosing $u_0 := (\alpha+\beta)\,x_0 \neq 0$ and $u_1 := x_0$, so that $\mathcal{F}(x,u) = -(\alpha+\beta)\,\|x_0\|^2 < 0$.

Finally, observe that when $N \geq 2$, then $\mathcal{F}_0(x,u) \not\geq 0$ over $x_0 = 0 = x_{N+1}$, which can be shown e.g. by choosing $u_1 := u_0 \neq 0$ and $u_2 = \cdots = u_N := 0$, so that $\mathcal{F}_0(x,u) = -2\,\|u_0\|^2 < 0$.

3.4.2 Functional with separated endpoints

In this subsection, let (X,U) be the normalized conjoined basis of (S), i.e. $X_0 = I - \mathcal{M}_0$ and $U_0 = \Gamma_0 + \mathcal{M}_0$, and let $(\bar{X},\bar{U})$ be its associated solution such that $\bar{X}_0 = \mathcal{M}_0$ and $\bar{U}_0 = \mathcal{M}_0 - I$. Recall that the functional $\mathcal{F}$ is defined by formula (1.37).

Theorem 3.64. *The following statements are equivalent.*

(i) $\mathcal{F}(x,u) \geq 0$ *over* $\mathcal{M}_0 x_0 = 0 = \mathcal{M}_1 x_{N+1}$.

(ii) *There exist* $\alpha > 0$ *and* $\beta > 0$ *such that*

$$\mathcal{F}(x,u) + \alpha \|\mathcal{M}_0 x_0\|^2 + \beta \|\mathcal{M}_1 x_{N+1}\|^2 \geq 0$$

over $\mathcal{M}_1 x_{N+1} - [\bar{X}_{N+1} + (I - \mathcal{M}_1)\bar{U}_{N+1}]\mathcal{M}_0 x_0 \in \operatorname{Im}[X_{N+1} + (I - \mathcal{M}_1)U_{N+1}]$.

(iii) *There exists* $\alpha > 0$ *such that*

$$\mathcal{F}(x,u) + \alpha \|\mathcal{M}_0 x_0\|^2 \geq 0$$

over $[\bar{X}_{N+1} + (I - \mathcal{M}_1)\bar{U}_{N+1}]\mathcal{M}_0 x_0 \in \operatorname{Im}[X_{N+1} + (I - \mathcal{M}_1)U_{N+1}]$, $\mathcal{M}_1 x_{N+1} = 0$.

(iv) *There exists* $\beta > 0$ *such that*

$$\mathcal{F}(x,u) + \beta \|\mathcal{M}_1 x_{N+1}\|^2 \geq 0$$

over $\mathcal{M}_1 x_{N+1} \in \operatorname{Im}[X_{N+1} + (I - \mathcal{M}_1)U_{N+1}]$, $\mathcal{M}_0 x_0 = 0$.

(v) *There exists* $\alpha > 0$ *such that*

$$\mathcal{F}(x,u) + \alpha \|\mathcal{M}_0 x_0\|^2 \geq 0$$

over $\mathcal{M}_1 x_{N+1} = [\bar{X}_{N+1} + (I - \mathcal{M}_1)\bar{U}_{N+1}]\mathcal{M}_0 x_0$.

(vi) $\mathcal{F}(x,u) \geq 0$ *over* $\mathcal{M}_1 x_{N+1} = [\bar{X}_{N+1} + (I - \mathcal{M}_1)\bar{U}_{N+1}]\mathcal{M}_0 x_0 = 0$.

Theorem 3.65. *The following statements are equivalent.*

(i') $\mathcal{F}(x,u) > 0$ *over* $\mathcal{M}_0 x_0 = 0 = \mathcal{M}_1 x_{N+1}$, $x \not\equiv 0$.

(ii') *There exist* $\alpha > 0$ *and* $\beta > 0$ *such that*

$$\mathcal{F}(x,u) + \alpha \|\mathcal{M}_0 x_0\|^2 + \beta \|\mathcal{M}_1 x_{N+1}\|^2 > 0 \quad \textit{over} \quad x \not\equiv 0.$$

(iii') *There exists* $\alpha > 0$ *such that*

$$\mathcal{F}(x,u) + \alpha \|\mathcal{M}_0 x_0\|^2 > 0 \quad \textit{over} \quad \mathcal{M}_1 x_{N+1} = 0, x \not\equiv 0.$$

(iv') *There exists $\beta > 0$ such that*

$$\mathcal{F}(x,u) + \beta \left\| \mathcal{M}_1 x_{N+1} \right\|^2 > 0 \quad \textit{over} \quad \mathcal{M}_0 x_0 = 0, x \not\equiv 0.$$

(v') *There exists $\alpha > 0$ such that*

$$\mathcal{F}(x,u) + \alpha \left\| \mathcal{M}_0 x_0 \right\|^2 > 0$$

over $\mathcal{M}_1 x_{N+1} = [\bar{X}_{N+1} + (I - \mathcal{M}_1)\bar{U}_{N+1}]\mathcal{M}_0 x_0$, $x \not\equiv 0$.

(vi') $\mathcal{F}(x,u) > 0$ *over* $\mathcal{M}_1 x_{N+1} = [\bar{X}_{N+1} + (I - \mathcal{M}_1)\bar{U}_{N+1}]\mathcal{M}_0 x_0 = 0$, $x \not\equiv \bar{X}\mathcal{M}_0 x_0$.

We prove both theorems with the use of the transformed system ($\tilde{\text{S}}$), introduced in Subsection 3.1.2.

First note that if (X,U) and $(\bar{X},\bar{U})$ are the principal and the associated solutions of ($\tilde{\text{S}}$), i.e. if $X_{-1} = 0$, $U_{-1} = I$, and $\bar{X}_{-1} = I$, $\bar{U}_{-1} = 0$ then $X_0 = I - \mathcal{M}_0$, $U_0 = \Gamma_0 + \mathcal{M}_0$, $\bar{X}_0 = \mathcal{M}_0$, $\bar{U}_0 = \mathcal{M}_0 - I$, $X_{N+2} = X_{N+1} + (I - \mathcal{M}_1)U_{N+1}$, and $\bar{X}_{N+2} = \bar{X}_{N+1} + (I - \mathcal{M}_1)\bar{U}_{N+1}$. Thus, the principal solution (X,U) of ($\tilde{\text{S}}$) on $[-1, N+2]$ is in fact the natural conjoined basis of (S) on $[0, N+1]$.

Remark 3.66. The associated solution $(\bar{X},\bar{U})$ given above is not the unique one which will work for this theorem. Another choice is to take the transformed system from Remark 3.17. Then we would get the associated solution with $\bar{X}_0 = [\Gamma_0 + \mathcal{M}_0 - \varepsilon(I - \mathcal{M}_0)]^{-1}$, $\bar{U}_0 = \varepsilon[\Gamma_0 + \mathcal{M}_0 - \varepsilon(I - \mathcal{M}_0)]^{-1}$.

Proof of Theorem 3.64. Conditions (ii)–(vi) imply condition (i) trivially.

Let condition (i) hold. Then, by Lemma 3.20, $\tilde{\mathcal{F}}_0(\tilde{x},\tilde{u}) \geq 0$ over $\tilde{x}_{-1} = 0 = \tilde{x}_{N+2}$, and hence conditions (ii)–(vi) from Theorem 3.58 are satisfied, where we replace the interval $[0, N+1]$ by $[-1, N+2]$, the pair (x,u) by $(\tilde{x},\tilde{u})$, the functional $\mathcal{F}_0(x,u)$ by $\tilde{\mathcal{F}}_0(\tilde{x},\tilde{u})$, and the solutions $(\hat{X},\hat{U})$ and $(\tilde{X},\tilde{U})$ by (X,U) and $(\bar{X},\bar{U})$. Now we will show that (i) implies (ii) in details. The other implications (i) $\Rightarrow$ (iii)–(vi) are proven in a similar way.

Let (x,u) be an admissible pair w.r.t. $(\mathcal{A},\mathcal{B})$ with $\mathcal{M}_1 x_{N+1} - [\bar{X}_{N+1} + (I - \mathcal{M}_1)\bar{U}_{N+1}]\mathcal{M}_0 x_0 \in \operatorname{Im}[X_{N+1} + (I - \mathcal{M}_1)U_{N+1}]$. Then the pair $(\tilde{x},\tilde{u})$ defined by (3.18) is admissible w.r.t. $(\tilde{\mathcal{A}},\tilde{\mathcal{B}})$ on $[-1, N+2]$, $\tilde{x}_{-1} = \mathcal{M}_0 x_0$, $\tilde{x}_{N+2} = \mathcal{M}_1 x_{N+1}$, and $\tilde{x}_{N+2} - \bar{X}_{N+2}\tilde{x}_{-1} \in \operatorname{Im} X_{N+2}$. Furthermore, by Lemma 3.18, $\mathcal{F}(x,u) = \tilde{\mathcal{F}}_0(\tilde{x},\tilde{u})$. Hence, by condition (ii) from Theorem 3.58, we have

$$0 \leq \tilde{\mathcal{F}}_0(\tilde{x},\tilde{u}) + \alpha \left\| \tilde{x}_{-1} \right\|^2 + \beta \left\| \tilde{x}_{N+2} \right\|^2 = \mathcal{F}(x,u) + \alpha \left\| \mathcal{M}_0 x_0 \right\|^2 + \beta \left\| \mathcal{M}_1 x_{N+1} \right\|^2,$$

where α and β are the positive numbers that exist by condition (ii) from Theorem 3.58. □

Proof of Theorem 3.65. The proof is same as the proof of Theorem 3.64, only we use Lemma 3.21 and Theorem 3.59 instead of Lemma 3.20 and Theorem 3.58. □

Other perturbation type conditions are possible, formulated via the principal solution $(\hat{X}, \hat{U})$ and the associated solution $(\tilde{X}, \tilde{U})$ of (S) instead of the natural conjoined basis (X, U) and the solution $(\bar{X}, \bar{U})$ of (S).

Theorem 3.67. *The following statements are equivalent.*

(i) $\mathcal{F}(x,u) \geq 0$ *over* $\mathcal{M}_0 x_0 = 0 = \mathcal{M}_1 x_{N+1}$.

(ii) *There exist* $\alpha > 0$ *and* $\beta > 0$, *and a* $2n \times 2n$ *matrix* O^* *such that*

$$\mathcal{F}(x,u) + \alpha \|\mathcal{M}_0 x_0\|^2 + \beta \|\mathcal{M}_1 x_{N+1}\|^2 - 2 \begin{pmatrix} x_0 \\ x_{N+1} \end{pmatrix}^T O^* \begin{pmatrix} \mathcal{M}_0 x_0 \\ \mathcal{M}_1 x_{N+1} \end{pmatrix} \geq 0$$

over $\mathcal{M}_1 x_{N+1} - \tilde{X}_{N+1} \mathcal{M}_0 x_0 \in \operatorname{Im} \hat{X}_{N+1}$.

(iii) *There exists* $\alpha > 0$ *such that*

$$\mathcal{F}(x,u) + \alpha \|\mathcal{M}_0 x_0\|^2 - 2\, x_{N+1}^T \tilde{U}_{N+1} \mathcal{M}_0 x_0 \geq 0$$

over $\tilde{X}_{N+1} \mathcal{M}_0 x_0 = \mathcal{M}_1 x_{N+1}$.

Proof. Conditions (ii) and (iii) imply condition (i) trivially.

(i) ⇒ (ii): Let (x,u) be an admissible pair with $\mathcal{M}_1 x_{N+1} - \tilde{X}_{N+1} \mathcal{M}_0 x_0 \in \operatorname{Im} \hat{X}_{N+1}$. Then, by Theorem 2.38, we have that the identity

$$\begin{aligned} \mathcal{F}_0(x,u) = \mathcal{F}_0(\bar{x},\bar{u}) &- \begin{pmatrix} \mathcal{M}_0 x_0 \\ \mathcal{M}_1 x_{N+1} \end{pmatrix}^T Q^* \begin{pmatrix} \mathcal{M}_0 x_0 \\ \mathcal{M}_1 x_{N+1} \end{pmatrix} \\ &+ 2 \begin{pmatrix} x_0 \\ x_{N+1} \end{pmatrix}^T \hat{U}_{N+1}^* \hat{X}_{N+1}^{*\dagger} \begin{pmatrix} \mathcal{M}_0 x_0 \\ \mathcal{M}_1 x_{N+1} \end{pmatrix}, \end{aligned}$$

holds, where Q^* is a symmetric $2n \times 2n$ matrix, and $(\bar{x},\bar{u})$ is an admissible pair with $\bar{x}_0 = (I - \mathcal{M}_0) x_0$ and $\bar{x}_{N+1} = (I - \mathcal{M}_1) x_{N+1}$. Thus, $x_0^T \Gamma_0 x_0 + x_{N+1}^T \Gamma_1 x_{N+1} = \bar{x}_0^T \Gamma_0 \bar{x}_0 + \bar{x}_{N+1}^T \Gamma_1 \bar{x}_{N+1}$.

Let λ_0 be the smallest eigenvalue of $-Q^*$. Then from inequality (1.8) we get

$$-\begin{pmatrix} \mathcal{M}_0 x_0 \\ \mathcal{M}_1 x_{N+1} \end{pmatrix}^T Q^* \begin{pmatrix} \mathcal{M}_0 x_0 \\ \mathcal{M}_1 x_{N+1} \end{pmatrix} \geq \lambda_0(\|\mathcal{M}_0 x_0\|^2 + \|\mathcal{M}_1 x_{N+1}\|^2).$$

Hence, the inequality in (ii) holds for the matrix $O^* := \hat{U}^*_{N+1} \hat{X}^{*\dagger}_{N+1}$ and any $\alpha, \beta \geq |\lambda_0|$.

(i) $\Rightarrow$ (iii): If $\tilde{X}_{N+1} \mathcal{M}_0 x_0 = \mathcal{M}_1 x_{N+1}$, then, by Corollary 2.39, we get

$$\mathcal{F}(x,u) \geq (\mathcal{M}_0 x_0)^T \tilde{X}^T_{N+1} \tilde{U}_{N+1} \mathcal{M}_0 x_0 + 2\, x^T_{N+1} \tilde{U}_{N+1} \mathcal{M}_0 x_0.$$

We use condition (1.8) and get that the inequality in (iii) holds for any $\alpha \geq |\lambda_1|$, where λ_1 is the smallest eigenvalue of $\tilde{X}^T_{N+1} \tilde{U}_{N+1}$. □

Remark 3.68. We can see that when the conditions are formulated via the solutions $(\hat{X}, \hat{U})$ and $(\tilde{X}, \tilde{U})$ of (S), an extra term with a $2n \times 2n$ matrix O^* appears in the perturbed functional. This term cannot be removed, as is shown in the following example, where $\mathcal{F} \geq 0$ over $\mathcal{M}_0 x_0 = 0 = \mathcal{M}_1 x_{N+1}$, but there is no $\alpha, \beta > 0$ such that $\alpha\, \|\mathcal{M}_0 x_0\|^2 + \beta\, \|\mathcal{M}_1 x_{N+1}\|^2 + \mathcal{F}(x,u) \geq 0$ over $\mathcal{M}_1 x_{N+1} - \tilde{X}_{N+1} \mathcal{M}_0 x_0 \in \operatorname{Im} \hat{X}_{N+1}$.

Example 3.69. Consider the coefficients $\mathcal{B}_0 = -\mathcal{C}_0 = \mathcal{D}_0 = I$, $\mathcal{A}_0 = 0$, and $\mathcal{A}_k = \mathcal{D}_k = I$, $\mathcal{B}_k = \mathcal{C}_k = 0$ for $k \in [1, N]$, and the matrices $\mathcal{M}_0 = 0, \mathcal{M}_1 = I$, and $\Gamma_0 = \Gamma_1 = 0$. Then the principal solution $\hat{X}_k$ is

$$\{\hat{X}_k\}_{k=0}^{N+1} = \{0, I, I, I, \dots, I\}.$$

Admissible pairs (x, u) are

$$\{(x_k, u_k)\}_{k=0}^{N+1} = \big\{(x_0, u_0), (u_0, u_1), (u_0, u_2), \dots, (u_0, u_{N+1})\big\},$$

and admissible pairs (x, u) with $\mathcal{M}_0 x_0 = 0 = \mathcal{M}_1 x_{N+1}$ are

$$\{(x_k, u_k)\}_{k=0}^{N+1} = \big\{(x_0, 0), (0, u_1), (0, u_2), \dots, (0, u_{N+1})\big\}.$$

The functional $\mathcal{F}$ takes the form

$$\mathcal{F}(x,u) = -2 x_0^T u_0 + u_0^T u_0,$$

and $\mathcal{F}(x,u) = 0 (\geq 0)$ when $\mathcal{M}_0 x_0 = 0 = \mathcal{M}_1 x_{N+1}$. On the other hand,

$$\mathcal{F}(x,u) + \alpha\, \|\mathcal{M}_0 x_0\|^2 + \beta\, \|\mathcal{M}_1 x_{N+1}\|^2 = -2 x_0^T u_0 + u_0^T u_0 + \beta\, \|u_0\|^2 \ngeq 0$$

over $\mathcal{M}_1 x_{N+1} - \tilde{X}_{N+1} \mathcal{M}_0 x_0 \in \operatorname{Im} \hat{X}_{N+1}$, i.e. over all (x, u) admissible, which follows for example by choosing $x_0 := (1 + \beta)\, u_0 \neq 0$, so that $\mathcal{F}(x,u) = -(1+\beta)\, \|u_0\|^2 < 0$.

3.4.3 Functional with general endpoints

In this subsection we again use the conjoined basis $(\hat{X}^*, \hat{U}^*)$ of (S^*), defined by (2.19) via the principal solution $(\hat{X}, \hat{U})$ and the associated solution $(\tilde{X}, \tilde{U})$ of (S). Recall that the functional $\mathcal{F}$ is now defined by formula (1.38).

Theorem 3.70. *The following statements are equivalent.*

(i) $\mathcal{F}(x,u) \geq 0$ *over* $\mathcal{M}\left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right) = 0$.

(ii) *There exists* $\alpha > 0$ *such that*

$$\mathcal{F}(x,u) + \alpha \left\| \mathcal{M}\left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right) \right\|^2 \geq 0$$

over $\mathcal{M}\left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right) \in \operatorname{Im}[\hat{X}^*_{N+1} + (I - \mathcal{M})\hat{U}^*_{N+1}]$.

Theorem 3.71. *The following statements are equivalent.*

(i') $\mathcal{F}(x,u) > 0$ *over* $\mathcal{M}\left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right) = 0$, $x \not\equiv 0$.

(ii') *There exists* $\alpha > 0$ *such that*

$$\mathcal{F}(x,u) + \alpha \left\| \mathcal{M}\left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right) \right\|^2 > 0 \quad \textit{over} \quad x \not\equiv 0.$$

The proofs of Theorems 3.70, 3.71 are again based on transforming system (S) and the quadratic functional $\mathcal{F}$ into augmented system (S^*) and the augmented quadratic functional

$$\mathcal{F}^*(x^*, u^*) := x_0^{*T} \Gamma_0^* \, x_0^* + x_{N+1}^{*T} \Gamma_1^* \, x_{N+1}^* + \mathcal{F}_0^*(x^*, u^*),$$

which has separated endpoints, where $\mathcal{M}_0^* := \frac{1}{2}\left(\begin{smallmatrix} I & -I \\ -I & I \end{smallmatrix}\right)$, $\mathcal{M}_1^* := \mathcal{M}$, $\Gamma_0^* := 0$, $\Gamma_1^* := \Gamma$.

Proof of Theorem 3.70. Condition (ii) implies condition (i) trivially.

Let condition (i) hold. Then, by Lemma 3.26 and Lemma 3.27, we have $\mathcal{F}^*(x^*, u^*) \geq 0$ over $\mathcal{M}_0^* x_0^* = 0 = \mathcal{M}_1^* x_{N+1}^*$, and hence condition (iv) from Theorem 3.64 holds for (x^*, u^*) and $\mathcal{F}^*(x^*, u^*)$ in place of (x,u) and $\mathcal{F}(x,u)$, and with the natural conjoined basis (X^*, U^*) of $(\tilde{\mathrm{S}})$ in place of the natural conjoined basis (X, U) of (S). This is equivalent to condition (ii), because $\mathcal{M}_0^* x_0^* = 0$ implies that (x^*, u^*) has the form from Lemma 3.26. Hence, $\mathcal{M}_1^* x_{N+1}^* = \mathcal{M}\left(\begin{smallmatrix} x_0 \\ x_{N+1} \end{smallmatrix}\right)$, $\mathcal{F}^*(x^*, u^*) = \mathcal{F}(x,u)$, and $(\hat{X}^*, \hat{U}^*) = (X^*, U^*)\left(\begin{smallmatrix} -I & I \\ I & I \end{smallmatrix}\right)$ and $\operatorname{Im}[X^*_{N+1} + (I - \mathcal{M})U^*_{N+1}] = \operatorname{Im}\left\{[X^*_{N+1} + (I - \mathcal{M})U^*_{N+1}]\left(\begin{smallmatrix} -I & I \\ I & I \end{smallmatrix}\right)\right\}$. □

Proof of Theorem 3.71. The proof is same as the proof of Theorem 3.70, we only use Theorem 3.65 instead of Theorem 3.64. □

3.5 Notes

The focal point definition (Definition 3.1) was for discrete symplectic systems first introduced in [13], and earlier for Hamiltonian difference systems in [9]. Regarding the positivity for zero endpoints in Theorem 3.4, the equivalence of (i)–(iii) is from [13], the equivalence of (i), (iv), and (v) is from [40], and the equivalence of (i), (vi), and (vii) is new and is contained in [37]. Regarding the positivity for separated endpoints in Theorem 3.14, tje equivalence of (i)–(v) is from [40], while the equivalence of (i), (vi), and (vii) is new and is contained in [37]. The tranformation of separated endpoints into zero endpoints in (3.17) is a modification of the one in [40]. Regarding the positivity for general endpoints in Theorem 3.24, (i)$\Leftrightarrow$(ii) is in [11] under the assumption $\operatorname{Ker}\mathcal{M} \subseteq \operatorname{Im}\hat{X}^*_{N+1}$. This assumption was eliminated (by taking $\operatorname{Ker}\mathcal{M} \cap \operatorname{Im}\hat{X}^*_{N+1}$) in [40] where this result was extended by the equivalence of (iii)–(v). Another transformation as it is described in Remark 3.25 can be found in [40].

The basic results on the matrices M_k, T_k in Subsection 3.2.1 are from [46]. Lemmas 3.36–3.38 are from [18]. Subsection 3.2.2 is new and it is contained in [37]. Regarding the nonnegativity for zero endpoints in Theorem 3.42, (i)$\Leftrightarrow$(ii) is from [18], (i)$\Leftrightarrow$(iii) is new and it is in [39], and the equivalence of (i), (iv), and (v) is new and is contained in [37]. The necessity of P-condition (3.48) (see also Lemma 3.44) is established in [24]. Regarding the nonnegativity for separated endpoints in Theorem 3.43, (i)$\Leftrightarrow$(ii) is from [16], and (i)$\Leftrightarrow$(iii) is new and it is contained in [39]. Regarding the nonnegativity for general endpoints, Theorem 3.49 is new. More precisely, the equivalence of (i) and (ii) is in [37] and the equivalence of (i) and (iii) is in [39].

The comparisons of implicit Riccati equations in Theorems 3.54, 3.55 are new, some parts are in [39]. The equivalences for perturbed quadratic functionals in Theorem 3.58 as well as Theorems 3.59, 3.64, 3.65, 3.70, 3.71 are new.

3.6 Perspectives

Some of our new results for discrete symplectic systems have already been generalized to time scales. For example, Riccati inequality (3.4) and some of the perturbations in Theorem 3.58 are derived in [44] and [38] for time scale symplectic systems, respectively. We believe that also the other new results can be extended to such systems, which would lead to new results even for continuous time linear Hamiltonian systems.

Bibliography

[1] C. D. Ahlbrandt, Equivalence of discrete Euler equations and discrete Hamiltonian systems, *J. Math. Anal. Appl.* **180** (1993), 498–517.

[2] C. D. Ahlbrandt, M. Heifetz, J. W. Hooker, W. T. Patula, Asymptotics of discrete time Riccati equations, robust control, and discrete linear Hamiltonian systems, *PanAmer. Math. J.* **5** (1996), no. 2, 1–39.

[3] C. D. Ahlbrandt, A. Peterson, The (n, n)-disconjugacy of a $2n^{\text{th}}$ order linear difference equation, In: "Advances in Differences Equations", R. P. Agarwal, editor. *Comput. Math. Appl.* **28** (1994), no. 1-3, 1–9.

[4] C. D. Ahlbrandt, A. C. Peterson, *Discrete Hamiltonian Systems: Difference Equations, Continued Fractions, and Riccati Equations,* Kluwer Academic Publishers, Boston, 1996.

[5] D. R. Anderson, Discrete trigonometric matrix functions, *PanAmer. Math. J.* **7** (1997), no. 1, 39–54.

[6] A. Ben-Israel, T. N. E. Greville, *Generalized Inverses: Theory and Applications,* Second Edition. Springer-Verlag, New York, NY, 2003.

[7] D. S. Bernstein, *Matrix Mathematics. Theory, Facts, and Formulas with Application to Linear Systems Theory,* Princeton University Press, Princeton, 2005.

[8] M. Bohner, *On Positivity of Discrete Quadratic Functionals,* PhD dissertation. University of Ulm, 1995.

[9] M. Bohner, Linear Hamiltonian difference systems: disconjugacy and Jacobi-type conditions, *J. Math. Anal. Appl.* **199** (1996), 804–826.

[10] M. Bohner, Riccati matrix difference equations and linear Hamiltonian difference systems, *Dynam. Contin. Discrete Impuls. Systems* **2** (1996), no. 2, 147–159.

[11] M. Bohner, Symplectic systems and related discrete quadratic functionals, *Facta Univ. Ser. Math. Inform.* **12** (1997), 143–156.

[12] M. Bohner, Discrete Sturmian theory, *Math. Inequal. Appl.* **1** (1998), 375–383.

[13] M. Bohner, O. Došlý, Disconjugacy and transformations for symplectic systems, *Rocky Mountain J. Math.* **27** (1997), no. 3, 707–743.

[14] M. Bohner, O. Došlý, Positivity of block tridiagonal matrices, *SIAM J. Matrix Anal.* **20** (1998), 182–195.

[15] M. Bohner, O. Došlý, The discrete Prüfer transformation, *Proc. Amer. Math. Soc.* **129** (2001), no. 9, 2715–2726.

[16] M. Bohner, O. Došlý, R. Hilscher, W. Kratz, Diagonalization approach to discrete quadratic functionals, *Arch. Inequal. Appl.* **1** (2003), no. 2, 261–274.

[17] M. Bohner, O. Došlý, W. Kratz, Discrete Reid roundabout theorems, *Dynam. Systems Appl.* **8** (1999), no. 3-4, 345–352.

[18] M. Bohner, O. Došlý, W. Kratz, Positive semidefiniteness of discrete quadratic functionals, *Proc. Edinburgh Math. Soc.* **46** (2003), 627-636.

[19] M. Bohner, O. Došlý, W. Kratz, An oscillation theorem for discrete eigenvalue problems, *Rocky Mountain J. Math.* **33** (2003), no. 4, 1233–1260.

[20] W. A. Coppel, *Disconjugacy,* Lecture Notes in Mathematics, Vol. 220, Springer-Verlag, Berlin - Heidelberg, 1971.

[21] O. Došlý, Transformations of linear Hamiltonian difference systems and some of their applications, *J. Math. Anal. Appl.* **191** (1995), 250–265.

[22] O. Došlý, S. Hilger, R. Hilscher, Symplectic dynamic systems, In: "Advances in Dynamic Equations on Time Scales", M. Bohner and A. Peterson, editors, pp. 293–334. Birkhäuser, Boston, 2003.

[23] O. Došlý, R. Hilscher, Disconjugacy, transformations and quadratic functionals for symplectic dynamic systems on time scales, *J. Differ. Equations Appl.* **7** (2001), 265–295.

[24] O. Došlý, R. Hilscher, V. Zeidan, Nonnegativity of discrete quadratic functionals corresponding to symplectic difference systems, *Linear Algebra Appl.* **375** (2003), 21–44.

[25] O. Došlý, W. Kratz, A Sturmian separation theorem for symplectic difference systems, *J. Math. Anal. Appl.*, To appear.

[26] O. Došlý, Z. Pospíšil, Hyperbolic transformation and hyperbolic difference systems, *Fasc. Math.* **32** (2001), 26–48.

[27] L. Erbe, P. Yan, Disconjugacy for linear Hamiltonian difference systems, *J. Math. Anal. Appl.* **167** (1992), 355–367.

[28] L. Erbe, P. Yan, Qualitative properties of Hamiltonian difference systems, *J. Math. Anal. Appl.* **171** (1992), 334–345.

[29] L. Erbe, P. Yan, Oscillation criteria for Hamiltonian matrix difference systems, *Proc. Amer. Math. Soc.* **119** (1993), no. 2, 525–533.

[30] L. Erbe, P. Yan, On the discrete Riccati equation and its applications to discrete Hamiltonian systems, *Rocky Mountain J. Math.* **25** (1995), no. 1, 167–178.

[31] K. Feng, The Hamiltonian way for computing Hamiltonian dynamics, In: "Applied and Industrial Mathematics", (Venice, 1989), pp. 17–35. *Math. Appl.*, Vol. 56. Kluwer, Dordrecht, 1991.

[32] K. Feng, D. Wang, A note on conservation laws of symplectic difference schemes for Hamiltonian systems, *J. Comput. Math.* **9** (1991), no. 3, 229–237.

[33] K. Feng, H. Wu, M. Qin, Symplectic difference schemes for linear Hamiltonian canonical systems, *J. Comput. Math.* **8** (1990), no. 4, 371–380.

[34] T. E. Fortmann, A matrix inversion identity, *IEEE Trans. Automat. Control* **15** (1970), 599–599.

[35] R. Hilscher, Disconjugacy of symplectic systems and positivity of block tridiagonal matrices, *Rocky Mountain J. Math.* **29** (1999), no. 4, 1301–1319.

[36] R. Hilscher, Reid roundabout theorem for symplectic dynamic systems on time scales, *Appl. Math. Optim.* **43** (2001), no. 2, 129–146.

[37] R. Hilscher, V. Růžičková, Riccati inequality and other results for discrete symplectic systems, *J. Math. Anal. Appl.*, To appear.

[38] R. Hilscher, V. Růžičková, Perturbation of nonnegative time scale quadratic functionals, Submitted.

[39] R. Hilscher, V. Růžičková, Implicit Riccati equations and quadratic functionals for discrete symplectic systems, *Internat. J. Difference Equ. Dynam. Systems*, To appear.

[40] R. Hilscher, V. Zeidan, Symplectic difference systems: variable stepsize discretization and discrete quadratic functionals, *Linear Algebra Appl.* **367** (2003), 67–104.

[41] R. Hilscher, V. Zeidan, A remark on discrete quadratic functionals with separable endpoints, *Rocky Mountain J. Math.* **33** (2003), no. 4, 1337–1351.

[42] R. Hilscher, V. Zeidan, Discrete quadratic functionals with jointly varying endpoints via separable endpoints, In: "New Progress in Difference Equations", Proceedings of the Sixth International Conference on Difference Equations (Augsburg, 2001), B. Aulbach, S. Elaydi, and G. Ladas, editors, pp. 461–470. Chapman & Hall/CRC, Boca Raton, FL, 2004.

[43] R. Hilscher, V. Zeidan, Coupled intervals for discrete symplectic systems, Submitted.

[44] R. Hilscher, V. Zeidan, Applications of time scale symplectic systems without normality, Submitted.

[45] W. Kratz, *Quadratic Functionals in Variational Analysis and Control Theory,* Akademie Verlag, Berlin, 1995.

[46] W. Kratz, Discrete oscillation, *J. Difference Equ. Appl.* **9** (2003), no. 1, 135–147.

[47] W. Kratz, Definiteness of quadratic functional, *Analysis* **23** (2003), 163–183.

[48] P. D. Lax, *Linear Algebra,* John Wiley & Sons, Inc., New York, 1997.

[49] T. Peil, A. Peterson, Criteria for C-disfocality of a self-adjoint vector difference equation, *J. Math. Anal. Appl.* **179** (1993), no. 2, 512–524.

[50] A. Peterson, C-disfocality for linear Hamiltonian difference systems, *J. Differential Equations* **110** (1994), no. 1, 53–66.

[51] A. Peterson, J. Ridenhour, The $(2,2)$-disconjugacy of a fourth order differnce equation, *J. Differ. Equations Appl.* **1** (1995), 87–93.

[52] M. Picone, Sulle autosoluzioni e sulle formule di maggiorazione per gli integral delle equazioni differenziali lineari ordinarie autoaggiunte, *Math. Z.* **28** (1928), 519–555.

[53] W. T. Reid, *Ordinary differential equations,* John Wiley & Sons, Inc., New York, 1971.

[54] W. T. Reid, *Riccati Differential Equations,* Academic Press, New York - London, 1972.

[55] W. T. Reid, *Sturmian theory for ordinary differential equations,* Springer-Verlag, New York, 1980.

[56] Y. Shi, Symplectic structure of discrete Hamiltonian systems, *J. Math. Anal. Appl.* **266** (2002), no. 2, 472–478.

[57] A. Xiao, Y. Tang, Symplectic properties of multistep Runge-Kutta methods, *Comput. Math. Appl.* **44** (2002), no. 10-11, 1329–1338.