Oscillation Theorems for Symplectic Difference Systems

Ondrej Dosly und Werner Kratz

Preprint Series: 2007-05

Fakultät für Mathematik und Wirtschaftswissenschaften
UNIVERSITÄT ULM

OSCILLATION THEOREMS FOR SYMPLECTIC DIFFERENCE SYSTEMS

ONDŘEJ DOŠLÝ AND WERNER KRATZ

Abstract. We consider symplectic difference systems involving a spectral parameter, together with the Dirichlet boundary conditions. The main result of the paper is a discrete version of the so-called oscillation theorem which relates the number of finite eigenvalues less than a given number to the number of focal points of the principal solution of the symplectic system. In two recent papers the same problem was treated and an essential ingredient was to establish the concept of the *multiplicity* of a focal point. But there was still a rather restrictive condition needed, which is eliminated here by using the concept of finite eigenvalues (or zeros) from the theory of matrix pencils.

1. Introduction

We consider the (discrete) *symplectic eigenvalue problem*

$$(\mathrm{E_N}) \qquad \begin{cases} x_{k+1} = & \mathcal{A}_k x_k + \mathcal{B}_k u_k, \\ u_{k+1} = & \mathcal{C}_k x_k + \mathcal{D}_k u_k - \lambda \mathcal{W}_k x_{k+1}, \quad 0 \le k \le N \\ & x_0 = 0 = x_{N+1}, \end{cases}$$

where $N \in \mathbb{N}$, $\lambda \in \mathbb{R}$ is the eigenvalue parameter, and where we assume that $\mathcal{A}_k, \mathcal{B}_k, \mathcal{C}_k, \mathcal{D}_k, \mathcal{W}_k$ are real $n \times n$ matrices for $0 \le k \le N$ such that 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., $\mathcal{S}_k^T \mathcal{J} \mathcal{S}_k = \mathcal{J}$, $\mathcal{J} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$, I being the $n \times n$ identity matrix, and $\mathcal{W}_k$ is nonnegative definite, i.e., $\mathcal{W}_k \ge 0$, and in particular symmetric, i.e. $\mathcal{W}_k^T = \mathcal{W}_k$ for $0 \le k \le N$. Then the above difference system is symplectic for all $\lambda \in \mathbb{R}$, i.e., the matrix

$$\mathcal{S}_k - \lambda \hat{\mathcal{S}}_k, \quad \hat{\mathcal{S}}_k := \begin{pmatrix} 0 & 0 \\ \mathcal{W}_k \mathcal{A}_k & \mathcal{W}_k \mathcal{B}_k \end{pmatrix}$$

1991 *Mathematics Subject Classification.* 39A12, 15A18.

Key words and phrases. Local oscillation theorem, global oscillation theorem, discrete eigenvalue problem, symplectic difference system, focal point, principal solution, matrix pencil, finite eigenvalue.

The research of both authors is supported by the grant 201/04/0580 of the Grant Agency of Czech Republic and by the project MSM0021622409 of the Ministry of Education of the Czech Republic.

is also symplectic for all $\lambda \in \mathbb{R}$ and altogether we have the following identities and inequalities (suppressing the index $k \in \{0, \dots, N\}$), see [4]

$$\begin{cases} \mathcal{A}^T\mathcal{C} = \mathcal{C}^T\mathcal{A},\ \mathcal{B}^T\mathcal{D} = \mathcal{D}^T\mathcal{B},\ \mathcal{A}^T\mathcal{D} - \mathcal{C}^T\mathcal{B} = I,\ \mathcal{W} \geq 0, \\ \mathcal{A}\mathcal{B}^T = \mathcal{B}\mathcal{A}^T,\ \mathcal{C}\mathcal{D}^T = \mathcal{D}\mathcal{C}^T,\ \mathcal{A}\mathcal{D}^T - \mathcal{B}\mathcal{C}^T = I. \end{cases} \tag{1}$$

Now, our symplectic difference system, with $x_k, u_k \in \mathbb{R}^n$,

$$x_{k+1} = \mathcal{A}_k x_k + \mathcal{B}_k u_k, \quad u_{k+1} = \mathcal{C}_k x_k + \mathcal{D}_k u_k - \lambda \mathcal{W}_k x_{k+1} \tag{2}$$

takes the equivalent form

$$\begin{cases} x_k &= (\mathcal{D}_k^T - \lambda \mathcal{B}_k^T \mathcal{W}_k) x_{k+1} - \mathcal{B}_k^T u_{k+1}, \\ u_k &= (-\mathcal{C}_k^T + \lambda \mathcal{A}_k^T \mathcal{W}_k) x_{k+1} + \mathcal{A}_k^T u_{k+1} \end{cases} \tag{3}$$

by using the inverse of $\mathcal{S}_k$, which is contained in the formulas (1).

The results given in our paper can be regarded as a continuation of the research initiated in [4], where the same problem was investigated. However, when that paper was written, the *general* concept of the *multiplicity* of a focal point of a solution of symplectic difference systems did not exist. This general concept was established in [19], but there was still needed an extra (rather restrictive) assumption (we discuss this assumption later in this paper). Now, having at disposal the general concept of multiplicity, and introducing the notion of *finite eigenvalues* (Definition 2 below), we can drop this assumption and treat the problem without any further condition than (1).

Our main result, the so-called Global Oscillation Theorem (Theorem 2), relates the number of finite eigenvalues less than or equal to a given number λ_0 to the number of focal points (counting multiplicity) of the principal solution at 0 of (2) with $\lambda = \lambda_0$. The general idea of the proof of the main results of the paper is similar as in [4] and consists in transforming (2) into a system of a canonical form. The principal solution of this canonical form at 0 has a special structure and it makes possible to derive the local result (Theorem 1). Comparing with [4], on the one hand this construction is more complicated since it has to incorporate certain cases which did not occur in [4] because of the extra assumption (A2) and a certain exceptional set which appeared in that paper. On the other hand, it is even simpler, because the construction is "local", i.e., at one discrete point only.

Recall that the basic elements of the oscillation and transformation theory of symplectic difference systems were established in [2], where the so-called roundabout theorem is proved. This theorem relates oscillatory properties of symplectic difference systems to positivity of the corresponding discrete quadratic functional and to solvability of the associated Riccati matrix difference equation. In the subsequent papers, we mention here [3, 7, 8, 12, 13, 14, 19], this theory was developed and applied in various directions, like the discrete calculus of variations and optimal control, the investigation of nonnegativity of discrete quadratic functionals, etc.

The paper is organized as follows. In the next section we formulate the main results of the paper, we also recall some concepts and statements of the oscillation theory of symplectic systems. The technical part of the paper, the proofs and the description of the transformation of (2) into the canonical form, is postponed to the Sections 3 and 4 of the paper.

2. Notation and main results

Throughout the paper, Ker, ind, Im, †, and def denote the kernel, index (i.e., the number of negative eigenvalues including their multiplicity), the image, the Moore-Penrose generalized inverse, and the defect (i.e., the dimension of the kernel) of a matrix indicated.

Let $X_k(\lambda)$, $U_k(\lambda)$ be the *principal solution* of (2) at $k=0$, i.e.,

$$X_0(\lambda) \equiv 0,\ U_0(\lambda) \equiv I, \quad X_k,\, U_k \in \mathbb{R}^{n\times n}.$$

Note that this matrix-valued solution is a so-called *conjoined basis* of (2), i.e., $X_k, U_k \in \mathbb{R}^{n\times n}$ solve (2) such that

$$X_k^T U_k = U_k^T X_k \quad \text{and} \quad \operatorname{rank}(X_k^T \ \ U_k^T) = n \quad \text{for all } k. \tag{4}$$

According to [19] we define the focal points including their multiplicities of $Z := (X_k(\lambda), U_k(\lambda))$ using the notation

$$\begin{cases} M_k(\lambda) := (I - X_{k+1}(\lambda)X_{k+1}^\dagger(\lambda))\mathcal{B}_k, \\ T_k(\lambda) := I - M_k^\dagger(\lambda)M_k(\lambda), \\ D_k(\lambda) := T_k(\lambda)X_k(\lambda)X_{k+1}^\dagger(\lambda)\mathcal{B}_k T_k(\lambda). \end{cases} \tag{5}$$

Observe that the matrix $T_k(\lambda)$ is symmetric.

Definition 1. The *number of focal points of* Z in the interval $(k, k+1]$ is defined by

$$m(k,\lambda) := m_1(k,\lambda) + m_2(k,\lambda),$$

where $m_1(k,\lambda) = \operatorname{rank} M_k(\lambda)$ is the multiplicity in the point $k+1$ and $m_2(k,\lambda) = \operatorname{ind} D_k(\lambda)$ is the number of focal points in the open interval $(k, k+1)$.

Moreover, we denote by

$$\begin{cases} n_1(\lambda) & \text{the number of focal points of } Z \\ & \text{in the interval } (0, N+1], \end{cases} \tag{6}$$

so that by Definition 1

$$n_1(\lambda) = \sum_{k=0}^{N} m(k,\lambda) = \sum_{k=0}^{N} \{\operatorname{rank} M_k(\lambda) + \operatorname{ind} D_k(\lambda)\}. \tag{7}$$

Since we always have that

$$\operatorname{rank} M_k(\lambda) \le \operatorname{rank}\left(I - X_{k+1}(\lambda)X_{k+1}^\dagger(\lambda)\right) = n - \operatorname{rank} X_{k+1}(\lambda)$$

and $\operatorname{ind} D_k(\lambda) \le \operatorname{rank} D_k(\lambda) \le \operatorname{rank} X_{k+1}(\lambda)$, $D_0(\lambda) \equiv 0$ and $M_0(\lambda) \equiv (I - \mathcal{B}_1\mathcal{B}_1^\dagger)\mathcal{B}_1 \equiv 0$, it follows from Definition (1) and (7) that

$$(8) \qquad m(0,\lambda) \equiv 0, \quad m(k,\lambda) \le n, \quad n_1(\lambda) \le nN \quad \text{for all } \lambda \in \mathbb{R},\ 1 \le k \le N.$$

Our first main theorem is the following local result.

Theorem 1. (Local Oscillation Theorem) *Assume* (1). *Then, for all* $\lambda \in \mathbb{R}$ *and* $0 \le k \le N$, *we have*

$$\begin{aligned} m(k,\lambda+) - m(k,\lambda-) &= \operatorname{ind} D_k(\lambda+) - \operatorname{ind} D_k(\lambda-) \\ &= \operatorname{rank} X_k(\lambda) - \operatorname{rank} X_k(\lambda+) - \operatorname{rank} X_{k+1}(\lambda) + \operatorname{rank} X_{k+1}(\lambda+), \end{aligned}$$

and

$$n_1(\lambda) = n_1(\lambda+).$$

This theorem is a consequence of the following more detailed result.

Proposition 1. *Assume* (1). *Then, for all* $\lambda \in \mathbb{R}$ *and* $0 \le k \le N$*:*

(i) $\operatorname{rank} X_{k+1}(\lambda+) = \operatorname{rank} X_{k+1}(\lambda-)$, $\operatorname{rank} M_k(\lambda+) = \operatorname{rank} M_k(\lambda-)$;

(ii) $\operatorname{ind} D_k(\lambda+) = \operatorname{ind} D_k(\lambda) + \operatorname{rank} M_k(\lambda) - \operatorname{rank} M_k(\lambda+)$,
$\operatorname{ind} D_k(\lambda-) = \operatorname{ind} D_k(\lambda+) + \operatorname{rank} X_{k+1}(\lambda) - \operatorname{rank} X_{k+1}(\lambda+) - \operatorname{rank} X_k(\lambda) + \operatorname{rank} X_k(\lambda+)$.

To avoid the first part of the assumption (A2) of [4], i.e., $\det X_{N+1}(\lambda) \not\equiv 0$ (see Remark 3 (ii) there; we will return to this assumption later in our paper) we need the following notation of "finite eigenvalues" (or "finite zeros") of ($\mathrm{E_N}$) (cf. [4, Remark 2 (i)] and [6, Def. 4.7],[23]).

Definition 2. A number λ is called a *finite eigenvalue* of ($\mathrm{E_N}$), if

$$\operatorname{rank} X_{N+1}(\lambda) < r_{N+1}, \quad \text{where } r_{N+1} := \max_\mu \operatorname{rank} X_{N+1}(\mu),$$

and

$$\theta_N(\lambda) := r_{N+1} - \operatorname{rank} X_{N+1}(\lambda)$$

is the *multiplicity of* λ. By

$$\sigma_N := \{\lambda \in \mathbb{R} :\ \lambda \text{ is a finite eigenvalue of } (\mathrm{E_N})\}$$

we denote the *finite spectrum* of ($\mathrm{E_N}$).

Note that $X_{N+1}(\lambda)$ is a polynomial in λ, so that there are always only *finitely many finite eigenvalues.* Hence

$$r_{N+1} = \operatorname{rank} X_{N+1}(\lambda+) = \operatorname{rank} X_{N+1}(\lambda-) \quad \text{for all } \lambda \in \mathbb{R}$$

and similarly,

$$(9) \qquad r_k := \max_\mu \operatorname{rank} X(\mu) = \operatorname{rank} X_k(\lambda+) = \operatorname{rank} X_k(\lambda+)$$

for all $\lambda \in \mathbb{R}$ and k.

Now we denote by

$$(10)\qquad \begin{cases} n_2(\lambda) & \text{the number of finite eigenvalues including} \\ & \text{multiplicities of (E}_\text{N}\text{) which are less than or equal to } \lambda. \end{cases}$$

Then, we conclude from Theorem 1 and (9), that for all $\lambda \in \mathbb{R}$

$$\begin{aligned} n_1(\lambda+) - n_1(\lambda-) &= \sum_{k=0}^{N}\{m(k,\lambda+) - m(k,\lambda-)\} \\ &= \sum_{k=0}^{N}\{r_{k+1} - \operatorname{rank} X_{k+1}(\lambda) - r_k + \operatorname{rank} X_k(\lambda)\} \\ &= r_{N+1} - \operatorname{rank} X_{N+1}(\lambda) - r_0 + \operatorname{rank} X_0(\lambda) \\ &= n_2(\lambda+) - n_2(\lambda-). \end{aligned}$$

Observe that $r_0 = \operatorname{rank} X_0(\lambda) = 0$. Of course, we have that

$$(11)\qquad n_2(\lambda) = n_2(\lambda+) \quad \text{and} \quad n_1(\lambda) = n_1(\lambda+) \quad \text{for all } \lambda \in \mathbb{R},$$

using assertion (ii) of Proposition 1 and (10).

Altogether, we have derived from our previous results (note that $n_1(\lambda) \le nN$ for all $\lambda \in \mathbb{R}$ and $n_2(\lambda) = 0$ for all $\lambda < \min \sigma_N < \infty$) the main theorem of this paper.

Theorem 2. (Global Oscillation Theorem) *Assume* (1)*. Then there exists an integer* $m \in \{0, 1, \dots, nN\}$*, such that, for all* $\lambda \in \mathbb{R}$*,*

$$(12)\qquad n_1(\lambda) = n_2(\lambda) + m.$$

We conclude this overview of our main results with some simple facts and remarks concerning some further consequences and relations particularly to matrix pencils.

Remark 1. (i) Observe that (8) and (12) imply that

$$n_2(\lambda) \le n_1(\lambda) \le nN \quad \text{for all } \lambda \in \mathbb{R}.$$

(ii) Because $n_2(\lambda) = 0$ for $\lambda < \min \sigma_N < \infty$ by the above results, it follows from (12) that

$$m = n_1(\lambda) \quad \text{for } \lambda < \lambda_{\min} := \min \sigma_N.$$

This number reflects the "positive definiteness" of the associated quadratic functional

$$\mathcal{F}(x,u;\lambda) = \sum_{k=0}^{N}\{x_k^T\mathcal{C}_k^T\mathcal{A}_kx_k + u_k^T\mathcal{D}_k^T\mathcal{B}_ku_k + 2u_k^T\mathcal{B}_k^T\mathcal{C}_kx_k - \lambda x_{k+1}^T\mathcal{W}_kx_{k+1}\}$$

for admissible sequences (x, u), i.e., $x_{k+1} = \mathcal{A}_kx_k + \mathcal{B}_ku_k$ for $0 \le k \le N$ and $x_0 = 0 = x_{N+1}$, if $\lambda < \lambda_{\min}$. More precisely, $m = 0$ if and only if $\mathcal{F}(x,u;\lambda) > 0$ for $\lambda < \lambda_{\min}$ and for every admissible sequence (x, u) with $x \not\equiv 0$ by [2, Th. 2]. Moreover, in general, by the methods of [3] or [8] we have that, for $\lambda < \lambda_{\min}$,

$$m = \dim\{x : \ (x,u) \text{ is admissible such that } \mathcal{F}(x,u;\lambda) \le 0\}.$$

(iii) The eigenvalue problem ($\mathrm{E_N}$) is *self-adjoint* in the following sense. All finite eigenvalues are real. Note that this fact is also a simple consequence of Proposition 2 (i) below, where $\tilde{\lambda}$ and $\tilde{z}$ are the complex conjugates of λ and z, and of assertion (v) of Proposition 2. Also, finite eigenvectors (see Definition 3 below) corresponding to different finite eigenvalues are orthogonal. This orthogonality property and its consequences depend essentially on the assumption that $\mathcal{W}_k \geq 0$ for all k (see also the example given in Remark 2 (i) below).

(iv) As pointed out by the referee, the eigenvalue problem ($\mathrm{E_N}$) is equivalent with a corresponding eigenvalue problem for a $2n(N+1) \times 2n(N+1)$ *matrix pencil* $\mathbf{A} + \lambda\mathbf{B}$ with the eigenvectors $(u_0, x_1, u_1, \dots, x_N, u_N, u_{N+1})$ and the block-diagonal matrix $\mathbf{B} = \operatorname{diag}\{0, \mathcal{W}_0, 0, \mathcal{W}_1, \dots, 0, \mathcal{W}_{N-1}, 0, 0\}$. We omit here to write down $\mathbf{A}$ explicitly, because it is not used here. Then, it follows quite easily from the difference system and the definition of the principal solution at 0 that

$$\det(\mathbf{A} + \lambda\mathbf{B}) = \det X_{N+1}(\lambda).$$

Definition 2 and the references on matrix pencils (cf. [9, Ch. XII] or [5, 22, 23]) imply that r_{N+1} is the *normal rank* of the pencil, and that our notion of *finite eigenvalues* (or zeros) coincides with the corresponding notion for pencils, and particularly, that the first part of assumption (A2) of [4], i.e. $\det X_{N+1} \not\equiv 0$, means that the pencil is *regular*. Hence, by omitting this assumption, we consider the *singular* case.

We also want to mention here that all the *minimal indices* of our special matrix pencil (occurring in the Kronecker canonical form, cf. [9, Ch. XII] or [6, 11, 22]), equal to zero. This is a direct consequence of Proposition 2 (viii) (see also [9, Section XII.6]). This fact does simplify the Kronecker canonical form of the pencil considerably, but it is not used here directly further on.

(v) The concept of eigenvalues and eigenvectors for ($\mathrm{E_N}$) of the main reference [4], i.e. λ is an eigenvalue if and only if $\det X_{N+1}(\lambda) = 0$, stems from the continuous eigenvalue problems for linear Hamiltonian *differential* systems. If $\det X_{N+1}(\lambda) \not\equiv 0$, i.e. if the first part of assumption (A2) of [4] holds, then as mentioned above, the corresponding matrix pencil is regular, and the definitions here and in [4] coincide. But if the pencil is *singular*, then the definition of eigenvalues in [4] is not appropriate any more. Instead, the concept of finite eigenvalues from Definition 2 is the right one for the singular case as one can see, and this concept stems from the theory of matrix pencils. Actually, there exists also the notion of *infinite eigenvalues* (or zeros) in the theory of matrix pencils (cf. [9] or [23]), but it does not play any role here. In particular, the geometric meaning of the concept of finite eigenvalues as formulated in Proposition 2 (v) below depends on the special structure (e.g. selfadjointness) of the corresponding matrix pencil, where the assumption $\mathcal{W}_k \geq 0$ plays an important role.

(vi) We conclude this remark by pointing out some possible applications of formula (12) (see also [4, Remark 3 (iv)]). Let $\lambda_0 \in \mathbb{R}$ be given. If we want to know how many finite eigenvalues of ($\mathrm{E_N}$) are less than or equal to λ_0, we can calculate recursively the principal solution (X, U) at 0 of (2) and determine the number of finite eigenvalues (zeros) of ($\mathrm{E_N}$) resp. $X_{N+1}(\lambda)$ that are less than or equal to λ_0. However, $X_{N+1}(\lambda)$ is a polynomial and it might be difficult to calculate this number. Alternatively, if the number m as discussed above is known, by Theorem 2 we could calculate the principal solution at 0 of (2) for the particular λ_0 in question and count the number of its focal points in the interval $(0, N+1]$. This procedure could possibly lead to a numerical algorithm to treat the algebraic eigenvalue problem ($\mathrm{E_N}$) also in this singular case, although it is well known that singular matrix pencils have in general ill-posed eigenstructure (cf. [22, 23] or [6, p. 180]). But we should point out that such a numerical application is so far only a speculation of the authors rather than a concrete statement. For the numerical treatment of the algebraic eigenvalue problem for symmetric, banded matrices via Sturm-Liouville difference equations (note that this is a very special case of ($\mathrm{E_N}$)!) the theory shows that $\{\det X_k(\lambda)\}$ forms a "Sturmian chain", which may be used similarly as for treating symmetric tridiagonal matrices, see [18] (cf. also [17]).

Concerning *theoretical applications*, our result here without assumption (A2) of [4] can be used to handle also general boundary conditions (rather that Dirichlet conditions) via an augmented "big" $4n \times 4n$ symplectic system (see e.g. [4, Section 4]) without the very restrictive assumption (A5) and the exceptional set $\mathcal{N}$ of [4]. Moreover, the authors are confident, that our general result here can also be used to derive a more general Sturmian separation theorem than in [8] (cf. also [20, Section VII.7] in the continuous case).

3. Auxiliary results

In view of our results below we define as follows.

Definition 3. For any number λ a solution $z = \binom{x_k}{u_k}_{k=0}^{N+1}$ of ($\mathrm{E_N}$) such that $(\mathcal{W}_k x_{k+1})_{k=0}^{N-1} \not\equiv 0$ is called a *finite eigenvector* corresponding to λ, and

$$\tilde{\theta}_N(\lambda) := \dim \left\{ (\mathcal{W}_k x_{k+1})_{k=0}^{N-1} \, : \, z = \begin{pmatrix} x_k \\ u_k \end{pmatrix}_{k=0}^{N+1} \text{ solves } (\mathrm{E_N}) \right\}$$

is the *geometric multiplicity* of λ.

Proposition 2. *Assume* (1). *Then the following statements hold:*

(i) *If* $z = \binom{x_k}{u_k}_{k=0}^{N+1}$, $\tilde{z} = \binom{\tilde{x}_k}{\tilde{u}_k}_{k=0}^{N+1}$ *are finite eigenvectors of* ($\mathrm{E_N}$) *corresponding to eigenvalues* λ *and* $\tilde{\lambda}$, *respectively, then*

$$(\tilde{\lambda} - \lambda)\langle \tilde{z}, z \rangle_{\mathcal{W}} = 0, \quad \langle z, z \rangle_{\mathcal{W}} > 0, \quad \langle \tilde{z}, \tilde{z} \rangle_{\mathcal{W}} > 0,$$

in particular, $\langle \tilde{z}, z \rangle_{\mathcal{W}} = 0$, *if* $\lambda \neq \tilde{\lambda}$, *where the bilinear form* $\langle \cdot, \cdot \rangle$ *is defined by*

$$\langle \tilde{z}, z \rangle_{\mathcal{W}} := \sum_{k=0}^{N-1} \tilde{x}_{k+1}^T \mathcal{W}_k x_{k+1}.$$

(ii) *The total number of* λ*'s, which possess a finite eigenvector, including their geometric multiplicities is* finite, *more precisely, it is* $\le \sum_{k=0}^{N-1} \operatorname{rank} \mathcal{W}_k \le nN$.

(iii) *The spaces*

$$\tilde{\mathcal{V}}_{N+1} = \tilde{\mathcal{V}}_{N+1}(\lambda) := \bigcap_{k=0}^{N-1} \operatorname{Ker}(\mathcal{W}_k X_{k+1}(\lambda)), \tag{13}$$
$$\mathcal{V}_{N+1} = \mathcal{V}_{N+1}(\lambda) := \tilde{V}_{N+1}(\lambda) \cap \operatorname{Ker} X_{N+1}(\lambda).$$

are independent of λ.

(iv) $X_k(\lambda)c$, $U_k(\lambda)c$ *are independent of* λ *for all* $c \in \mathcal{V}_{N+1}$, $0 \le k \le N+1$.

(v) *A number* λ *is a finite eigenvalue of* (E_N) *according to Definition 2 if and only if it possesses a finite eigenvector, and we have that* $\theta_N(\lambda) = \tilde{\theta}(\lambda)$.

(vi) $\operatorname{Ker} X_{N+1}(\lambda) = \mathcal{V}_{N+1}$ *for all* $\lambda \in \mathbb{R} \setminus \sigma_N$.

(vii) *The total number of finite eigenvalues including their multiplicities, denoted by* $|\sigma_N|$, *satisfies*

$$|\sigma_N| \le \sum_{k=0}^{N-1} \mathcal{W}_k \le nN,$$

so that $n_2(\lambda) = 0$ *if* $\lambda < \lambda_{\min} := \min \sigma_N$.

(viii) $\operatorname{Ker} X_{N+1}(\lambda)$, $\operatorname{Ker} X^T_{N+1}(\lambda)$, $X_{N+1}(\lambda)X^\dagger_{N+1}(\lambda)$, $M_N := M_N(\lambda)$, *and* $T_N := T_N(\lambda)$ *are* constant *on* $\mathbb{R} \setminus \sigma_N$.

Note, that we may replace the fixed integer N by another integer, say k, and then we have the corresponding notion, i.e. (E_k), σ_k, $\mathcal{V}_{k+1}$, r_{k+1} etc., and the corresponding results.

Proof. (i) We have

$$z_{k+1} = (\mathcal{S}_k - \lambda \hat{\mathcal{S}}_k) z_k, \quad \tilde{z}_{k+1} = (\mathcal{S}_k - \tilde{\lambda} \hat{\mathcal{S}}_k) \tilde{z}_k,$$

hence

$$\tilde{z}^T_{k+1} \mathcal{J} z_{k+1} = \tilde{z}^T_k (\mathcal{S}_k - \tilde{\lambda} \hat{\mathcal{S}}_k)^T \mathcal{J} (\mathcal{S}_k - \lambda \hat{\mathcal{S}}_k) z_k.$$

Denote $\alpha_k := \tilde{z}^T_k \mathcal{J} z_k = \tilde{x}^T_k u_k - \tilde{u}^T_k x_k$, then using the fact that $\mathcal{S}_k$ is symplectic and that $x_{k+1} = \mathcal{A}_k x_k + \mathcal{B}_k u_k$, we have

$$\begin{aligned} \alpha_{k+1} - \alpha_k &= -\tilde{\lambda} \tilde{z}^T_k \hat{\mathcal{S}}^T_k \mathcal{J} \mathcal{S}_k z_k - \lambda \tilde{z}^T_k \mathcal{S}^T_k \mathcal{J} \hat{\mathcal{S}}_k z_k \\ &= (\tilde{\lambda} - \lambda) \tilde{x}^T_{k+1} \mathcal{W}_k x_{k+1}. \end{aligned}$$

Summing up from $k = 0$ to $k = N$ we obtain a telescope sum, so that

$$0 = (\tilde{\lambda} - \lambda) \langle \tilde{z}, z \rangle_{\mathcal{W}}$$

since $x_0 = x_{N+1} = 0 = \tilde{x}_0 = \tilde{x}_{N+1}$. Finally, since $\mathcal{W}_k \ge 0$ for $0 \le k \le N-1$, we have that $\langle z, z \rangle_{\mathcal{W}} > 0$ if and only if $(\mathcal{W}_k x_{k+1})_{k=0}^{N-1} \not\equiv 0$. As already mentioned in Remark 1, the positive semidefiniteness of $\mathcal{W}_k$ is crucial for our results, in particular for the next assertion.

(ii) Let $z^\nu = \binom{x_k^\nu}{u_k^\nu}$ be finite eigenvectors corresponding to λ_ν for $\nu = 1, \ldots, M$. Then by (i), $\langle z^\nu, z^\nu \rangle_{\mathcal{W}} > 0$ and $\langle z^\mu, z^\nu \rangle_{\mathcal{W}} = 0$ for $\mu \neq \nu$ (use also the Gram-Schmidt theorem if $\lambda_\mu = \lambda_\nu$). Since $\langle \cdot, \cdot \rangle_{\mathcal{W}}$ defines an inner product on the linear space

$$\mathcal{L} := \left\{ \left(\mathcal{W}_k x_{k+1}^\mu\right)_{k=0}^{N-1} : x_{k+1}^\mu \in \mathbb{R}^n \text{ for } \mu = 1, \ldots, M \right\},$$

the vectors $\left(\mathcal{W}_k x_{k+1}^\nu\right)_{k=0}^{N-1}$ are *linearly independent* for $\nu = 1, \ldots, M$. Hence

$$M \leq \dim \mathcal{L} = \sum_{k=0}^{N-1} \operatorname{rank} \mathcal{W}_k \leq nN,$$

which proves assertion (ii).

(iii), (iv) First, let $c \in \tilde{\mathcal{V}}_{N+1}(\lambda)$, and define $z_k = \binom{x_k}{u_k} := \binom{X_k(\lambda)}{U_k(\lambda)} c$, $0 \leq k \leq N+1$. Then $x_0 = 0$, $u_0 = c$ and

$$x_{k+1} = \mathcal{A}_k x_k + \mathcal{B}_k u_k, \quad u_{k+1} = \mathcal{C}_k x_k + \mathcal{D}_k u_k - \lambda \mathcal{W}_k x_{k+1} = \mathcal{C}_k x_k + \mathcal{D}_k u_k$$

for $0 \leq k \leq N-1$ because $c \in \tilde{\mathcal{V}}_{N+1}(\lambda)$, and also for $k = N$ if $c \in \mathcal{V}_{N+1}(\lambda)$. By the *uniqueness* of solutions of initial value problems, we can conclude that $z_k = \binom{X_k(0)}{U_k(0)} c$ for $0 \leq k \leq N$, and also for $k = N+1$ if $c \in \mathcal{V}_{N+1}(\lambda)$. Hence, $c \in \tilde{\mathcal{V}}_{N+1}(0)$ and also $c \in \mathcal{V}_{N+1}(0)$ if $c \in \mathcal{V}_{N+1}(\lambda)$. This proves (iii) and (iv).

(v)–(vii) Next, by the definition of the principal solution at 0, a given λ possesses a finite eigenvector if and only if there exists $c \in \mathbb{R}^n$ such that $c \in \operatorname{Ker} X_{N+1}(\lambda)$ and $(\mathcal{W}_k X_{k+1}(\lambda) c)_{k=0}^{N-1} \not\equiv 0$, and this means that $c \in \operatorname{Ker} X_{N+1}(\lambda)$ and $c \notin \tilde{\mathcal{V}}_{N+1}(\lambda) = \tilde{\mathcal{V}}_{N+1} \supset \mathcal{V}_{N+1}$, which is equivalent with $\mathcal{V}_{N+1} \subsetneqq \operatorname{Ker} X_{N+1}(\lambda)$ by (13). Moreover, we have by Definition 3 that $\tilde{\theta}_N(\lambda) = \dim \operatorname{Ker} X_{N+1}(\lambda) - \dim \mathcal{V}_{N+1}$ for all $\lambda \in \mathbb{R}$. It follows from assertions (ii) and (13) that $\operatorname{Ker} X_{N+1}(\lambda) = \mathcal{V}_{N+1}$ for all $\lambda \in \mathbb{R} \setminus \tilde{\sigma}_N$, where $\tilde{\sigma}_N$ is a certain finite set. Hence, by Definition 2, $r_{N+1} = n - \dim \mathcal{V}_{N+1}$, i.e., $\tilde{\theta}(\lambda) = r_{N+1} - \operatorname{rank} X_{N+1}(\lambda) = \theta_N(\lambda)$. Thus, we have shown the assertions (v) and (vi), and also the first part of (viii). Moreover, (ii) and (v) yield assertion (vii). But, actually, this result is well-known from the theory of matrix pencils (see [5, Lemma 1]). Because, the numbers o_r and o_f occurring there equal zero, since the minimal indices are zero for our matrix pencil, it follows from [5, Lemma 1] that

$$|\sigma_N| = \sum_{k=0}^{N-1} \operatorname{rank} \mathcal{W}_k - o_\infty,$$

where o_∞ denotes the number of *infinite zeros* of the matrix pencil as defined in [9, Ch. XII] (see also [5, 21, 22, 23]).

(viii) Finally, for the proof that $\operatorname{Ker} X_{N+1}^T(\lambda)$ etc. are constant on $\mathbb{R} \setminus \sigma_N$, let $\mathcal{P} = (\mathcal{P}_1 \ \mathcal{P}_2)$ be an orthogonal matrix, such th $\operatorname{Im} \mathcal{P}_2 = \mathcal{V}_{N+1} = \operatorname{Ker} X_{N+1}(\lambda)$, using the Gram-Schmidt theorem. Then, $X_{N+1}(\lambda)\mathcal{P} = (X_{N+1}(\lambda)\mathcal{P}_1 \ 0)$ with $\operatorname{rank} X_{N+1}(\lambda)\mathcal{P}_1 = r_{N+1}$ for $\lambda \in \mathbb{R} \setminus \sigma_N$. Next, let $\lambda_0 \in \mathbb{R} \setminus \sigma_N$, and let $\mathcal{Q} = \binom{\mathcal{Q}_1}{\mathcal{Q}_2}$ be orthogonal with $\operatorname{Im} \mathcal{Q}_2^T = \operatorname{Ker} X_{N+1}^T(\lambda_0)$, so that $\mathcal{Q}_2 X_{N+1}(\lambda_0) = 0$. Then

$$\mathcal{Q}X_{N+1}(\lambda_0)\mathcal{P} = \begin{pmatrix} X_{11} & 0 \\ 0 & 0 \end{pmatrix}$$

with $\det X_{11} \neq 0$. Now, define for all $\lambda \in \mathbb{R}$

$$\begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}(\lambda) := \mathcal{Q}X_{N+1}(\lambda)\mathcal{P}, \quad \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix}(\lambda) := \mathcal{Q}U_{N+1}(\lambda)\mathcal{P},$$

so that, by the definition of $\mathcal{P}$ and $\mathcal{Q}$, $X_{12}(\lambda) = 0$, $X_{22}(\lambda) = 0$ for all $\lambda \in \mathbb{R}$, and $X_{21}(\lambda_0) = 0$, $X_{11}(\lambda_0) = X_{11}$. Next, the matrix $X^T_{N+1}(\lambda)U_{N+1}(\lambda)$ is symmetric and $\operatorname{rank}\bigl(X^T_{N+1}(\lambda)\ U^T_{N+1}(\lambda)\bigr) = n$ for all $\lambda \in \mathbb{R}$, because $Z = \binom{X_k(\lambda)}{U_k(\lambda)}$ is a conjoined basis of (2). Hence,

$$X^T_{N+1}(\lambda_0)U_{N+1}(\lambda_0) = \mathcal{P}\begin{pmatrix} X^T_{11}U_{11}(\lambda_0) & X^T_{11}U_{12}(\lambda_0) \\ 0 & 0 \end{pmatrix}\mathcal{P}^T$$

is symmetric, so that $X^T_{11}U_{12}(\lambda_0) = 0$, $U_{12}(\lambda) = U_{12}(\lambda_0) = 0$ for all $\lambda \in \mathbb{R}$, because $\det X_{11} \neq 0$, and $U_{12}(\lambda)$ does not depend on λ by (iv). Moreover,

$$n = \operatorname{rank}(X^T_{N+1}(\lambda_0)\ U^T_{N+1}(\lambda_0)) = \operatorname{rank}\begin{pmatrix} X^T_{11} & U^T_{11}(\lambda_0) & U^T_{21}(\lambda_0) \\ 0 & 0 & U^T_{22}(\lambda_0) \end{pmatrix},$$

so that $\det U_{22}(\lambda_0) \neq 0$. Observe also that $U_{22}(\lambda) = U_{22}$ is independent of λ again by (iv). Now, for all $\lambda \in \mathbb{R}$

$$X^T_{N+1}(\lambda)U_{N+1}(\lambda) = \mathcal{P}\begin{pmatrix} X^T_{11}(\lambda)U_{11}(\lambda) + X^T_{21}(\lambda)U_{21}(\lambda) & X^T_{21}(\lambda)U_{22} \\ 0 & 0 \end{pmatrix}\mathcal{P}^T,$$

so that $X^T_{21}(\lambda)U_{22} = 0$, i.e., $X_{21}(\lambda) = 0$ since U_{22} is nonsingular. Thus, we have shown that

$$\mathcal{Q}X_{N+1}(\lambda)\mathcal{P} = \begin{pmatrix} X_{11}(\lambda) & 0 \\ 0 & 0 \end{pmatrix} \quad \text{with } \det X_{11}(\lambda) \neq 0$$

for $\lambda \in \mathbb{R} \setminus \sigma_N$. It follows that $X^\dagger_{N+1}(\lambda) = \mathcal{P}\begin{pmatrix} X^{-1}_{11}(\lambda) & 0 \\ 0 & 0 \end{pmatrix}\mathcal{Q}$, so that

$$\begin{aligned} X_{N+1}(\lambda)X^\dagger_{N+1}(\lambda) &= \mathcal{Q}^T\begin{pmatrix} X_{11}(\lambda) & 0 \\ 0 & 0 \end{pmatrix}\mathcal{P}^T\mathcal{P}\begin{pmatrix} X^{-1}_{11}(\lambda) & 0 \\ 0 & 0 \end{pmatrix}\mathcal{Q} \\ &= \mathcal{Q}^T\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}\mathcal{Q}, \end{aligned}$$

and that $\operatorname{Ker} X^T_{N+1}(\lambda) = \operatorname{Ker}\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}\mathcal{Q}$ are constant on $\mathbb{R} \setminus \sigma_N$; and this implies the constancy of $M_N(\lambda)$ and $T_N(\lambda)$ by their definition in (5). □

Remark 2. (i) The following example shows that the statements (i), (ii), and (v) do not hold without the assumption $\mathcal{W}_k \geq 0$. Put $N = 3$, $\mathcal{A}_k = \operatorname{diag}\{1, 0\}$, $\mathcal{B}_k = \operatorname{diag}\{0, 1\}$, $\mathcal{C}_k = \operatorname{diag}\{0, -1\}$, $\mathcal{D}_k = \operatorname{diag}\{1, 0\}$ for $0 \leq k \leq 3$, $\mathcal{W}_0 = \operatorname{diag}\{0, 1\}$, $\mathcal{W}_1 = 0$, and $\mathcal{W}_2 = \operatorname{diag}\{0, -1\}$. Then the system is symplectic, $\mathcal{W}_0 \geq 0$, $\mathcal{W}_1 \geq 0$, but $\mathcal{W}_2 \not\geq 0$. The calculation of the principal solution (X, U) by the difference system yields: $X_0 = 0$, $X_1 = \operatorname{diag}\{0, 1\}$, $X_2 = \operatorname{diag}\{0, -\lambda\}$, $X_3 = \operatorname{diag}\{0, -1\}$, and $X_4 = X_{N+1}(\lambda) \equiv 0$. Hence $r_{N+1} = 0$, so that there are *no* finite eigenvalues by Definition 2, but every λ possesses a finite eigenvector, namely $z = (z_k) = \binom{x_k}{u_k} = \binom{X_k(\lambda)}{U_k(\lambda)}\binom{0}{1}$ so that $x_0 = \binom{0}{0}$, $x_1 = \binom{0}{1}$, $x_2 = \binom{0}{-\lambda}$, $x_3 = \binom{0}{-1}$, and $x_4 = \binom{0}{0}$.

(ii) If $X = X_k$, $U = U_k$ then (4) implies that

$$XX^\dagger U(I - X^\dagger X) = 0 \quad \text{and} \quad \operatorname{rank}(I - XX^\dagger)U(I - X^\dagger X) = \operatorname{rank}(I - X^\dagger X) \tag{14}$$

by using the singular value decomposition of X (cf. [6, Th. 3.2, p. 109]). These formulas (14) can be used in the proof of assertion (viii) above to retrieve directly the special structure of $U = \mathcal{Q}U_{N+1}(\lambda_0)\mathcal{P}$ from the structure of $X = \mathcal{Q}X_{N+1}(\lambda_0)\mathcal{P}$. This alternative argument was pointed out to us by the referee.

4. Proof of Proposition 1

The most technical part of our paper is the proof of Proposition 1. As we have already mentioned below Proposition 2, the results we have presented so far can be reformulated for any $k \in \{1, \dots, N\}$. We proceed similarly as in [4, pp. 1244], but the construction we perform below does not need the technical assumption (A2) of [4] which we assumed in that paper.

4.1. **Construction.** As it is shown in the proof of Proposition 2, there are *orthogonal* matrices $\mathcal{P}, \mathcal{Q} \in \mathbb{R}^{n\times n}$ such that $\mathcal{P} = (\mathcal{P}_1\ \mathcal{P}_2)$ with $\mathcal{P}_1 \in \mathbb{R}^{n\times r}$, $\mathcal{P}_2 \in \mathbb{R}^{n\times(n-r)}$ with $r = r_{N+1}$, $\operatorname{Im}\mathcal{P}_2 = \mathcal{V}_{N+1} = \operatorname{Ker} X_{N+1}(\lambda)$ for $\lambda \in \mathbb{R} \setminus \sigma_N$, and such that

$$\begin{cases} \tilde{X}_{N+1}(\lambda) := \mathcal{Q}X_{N+1}(\lambda)\mathcal{P} = \begin{pmatrix} X_{11}(\lambda) & 0_{r\times(n-r)} \\ 0 & 0_{(n-r)\times(n-r)} \end{pmatrix}, \\ \text{where } r = r_{N+1} = n - \dim \mathcal{V}_{N+1}, \\ \tilde{U}_{N+1}(\lambda) := \mathcal{Q}U_{N+1}(\lambda)\mathcal{P} = \begin{pmatrix} U_{11}(\lambda) & 0_{r\times(n-r)} \\ U_{21}(\lambda) & U_{22} \end{pmatrix} \\ \text{with} \quad U_{11}(\lambda) \in \mathbb{R}^{r\times r} \quad \text{and} \quad U_{22} \in \mathbb{R}^{(n-r)\times(n-r)} \\ \text{for all } \lambda \in \mathbb{R}, \text{ where } U_{22} \text{ does not depend on } \lambda \text{ with } \det U_{22} \neq 0, \\ \text{and that } \det X_{11}(\lambda) \neq 0 \text{ if and only if } \lambda \in \mathbb{R} \setminus \sigma_N, \\ \text{and such that } X_N(\lambda)\mathcal{P}_2 \text{ and } U_N(\lambda)\mathcal{P}_2 \text{ are independent of } \lambda. \end{cases} \tag{15}$$

Here and in the sequel, the index by the zero or identity matrix means it dimension. Also, we do not write this index explicitly, when the dimension is clear from the position of a matrix inside of a larger matrix.

Let $\lambda_1 \in \mathbb{R} \setminus \sigma_{N-1}$ and let $\rho := \operatorname{rank} X_N(\lambda)\mathcal{P}_2$. This quantity ρ is quite important here, and it did not occur in [4] because we had $\operatorname{rank} X_{N+1}(\lambda) \subset \operatorname{rank} X_N(\lambda)$ there. But here, this kernel condition is violated in general, and ρ describes the dimension of the subspace where it is violated. More precisely, we will show that (see (34) below)

$$\rho = \operatorname{rank} M_N(\lambda) = m_1(N,\lambda) \quad \text{for } \lambda \in \mathbb{R} \setminus \sigma_N.$$

We have $\rho \le n - r$ and $\rho \le r_N = \operatorname{rank} X_N(\lambda_1)$ by Proposition 2. Moreover, we define $\tilde r := r_N - \rho$, so that $0 \le \tilde r \le \operatorname{rank} X_N(\lambda_1)\mathcal{P}_1 \le r$. These new integers ρ and $\tilde r$ lead to a *refined block structure* as follows. First, there exist orthogonal matrices $\tilde{\mathcal{Q}}$, $\bar P_1$, $\bar P_2$, $\bar Q_1$ of sizes $n\times n$, $r \times r$, $(n-r)\times(n-r)$, and $(n-\rho)\times(n-\rho)$, respectively, such that the following holds:

$$\tilde{\mathcal{Q}} X_N(\lambda)\mathcal{P}_2 \bar P_2 = \begin{pmatrix} 0_{\tilde r\times(n-\rho-r)} & 0_{\tilde r \times \rho} \\ 0 & 0_{(r-\tilde r)\times\rho} \\ 0_{(n-r-\rho)\times(n-r-\rho)} & 0 \\ 0_{\rho\times(n-r-\rho)} & \tilde X_{44} \end{pmatrix} \quad \text{with } \det \tilde X_{44} \neq 0, \tag{16}$$

and

$$\begin{pmatrix} \bar Q_1 & 0 \\ 0 & I \end{pmatrix} \tilde{\mathcal{Q}} X_N(\lambda_1)\mathcal{P}_1 \bar P_1 = \begin{pmatrix} \tilde X_{11}(\lambda_1) & 0_{\tilde r\times(r-\tilde r)} \\ 0_{(r-\tilde r)\times\tilde r} & 0 \\ 0_{(n-\rho-r)\times\tilde r} & 0 \\ \tilde X_{41}(\lambda_1) & \tilde X_{42}(\lambda_1) \end{pmatrix} \quad \text{with } \det \tilde X_{11}(\lambda_1) \neq 0$$

and $\tilde X_{41}(\lambda_1) \in \mathbb{R}^{\rho\times\tilde r}$. Now, multiplication of $\tilde Q$ from the left by $\begin{pmatrix} \bar Q_1 & 0 \\ 0 & I \end{pmatrix}$ does not change the structure of formula (16), and multiplication of $\mathcal{P}$ from the right by $\begin{pmatrix} \bar P_1 & 0 \\ 0 & \bar P_2 \end{pmatrix}$ does not change the structure of (15) either. Hence, there exists, in addition to $\mathcal{P}$ and $\mathcal{Q}$, another orthogonal matrix $\tilde{\mathcal{Q}} \in \mathbb{R}^{n\times n}$ such that

$$\tilde X_N(\lambda_1) = \begin{pmatrix} \tilde X_{11}(\lambda_1) & 0 & 0 & 0_{\tilde r\times\rho} \\ 0_{(r-\tilde r)\times\tilde r} & 0_{(r-\tilde r)\times(r-\tilde r)} & 0 & 0 \\ 0 & 0 & 0_{(n-r-\rho)(n-r-\rho)} & 0 \\ \tilde X_{41}(\lambda_1) & \tilde X_{42}(\lambda_1) & 0_{\rho\times(n-r-\rho)} & \tilde X_{44} \end{pmatrix},$$

where $\tilde X_N(\lambda) := \tilde Q X_N(\lambda)\mathcal{P}$, with $\det \tilde X_{11}(\lambda_1) \neq 0$ and $\det \tilde X_{44} \neq 0$. We will use this *refinement of the block structure* of (15) further on. It follows that $\operatorname{Ker} \tilde X_N(\lambda_1) = \operatorname{Im} \mathcal{K}$ and $\operatorname{Ker} \tilde X_N^T(\lambda_1) = \operatorname{Im} \tilde{\mathcal{K}}$ with

$$\mathcal{K} = \begin{pmatrix} 0_{\tilde r\times(r-\tilde r)} & 0_{\tilde r\times(n-r-\rho)} \\ I_{r-\tilde r} & 0 \\ 0 & I_{n-r-\rho} \\ -\tilde X_{44}^{-1}\tilde X_{42}(\lambda_1) & 0_{\rho\times(n-r-\rho)} \end{pmatrix}, \quad \tilde{\mathcal{K}} = \begin{pmatrix} 0_{\tilde r\times(r-\tilde r)} & 0_{\tilde r\times(n-r-\rho)} \\ I_{r-\tilde r} & 0 \\ 0 & I_{n-r-\rho} \\ 0_{\rho\times(r-\tilde r)} & 0 \end{pmatrix}.$$

Since, by Proposition 2, the kernels of $X_N(\lambda)$ and $X_N^T(\lambda)$ and then also of $\tilde{X}_N(\lambda)$ and $\tilde{X}_N^T(\lambda)$ are *constant* on $\mathbb{R} \setminus \sigma_{N-1}$, it follows (use continuity for $\lambda \in \sigma_{N-1}$) that

$$\text{(17)} \quad \begin{cases} \tilde{X}_N(\lambda) := \tilde{Q} X_N(\lambda) \mathcal{P} \\ = \begin{pmatrix} \tilde{X}_{11}(\lambda) & 0_{\tilde{r}\times(r-\tilde{r})} & 0 & 0 \\ 0_{(r-\tilde{r})\times\tilde{r}} & 0 & 0 & 0 \\ 0 & 0 & 0_{(n-r-\rho)\times(n-r-\rho)} & 0 \\ \tilde{X}_{41}(\lambda) & \tilde{X}_{42} & 0_{\rho\times(n-r-\rho)} & \tilde{X}_{44} \end{pmatrix} \quad \text{for all } \lambda \in \mathbb{R}, \\ r = r_{N+1},\ \rho = \operatorname{rank} X_N(\lambda)\mathcal{P}_2,\ \tilde{r} + \rho = r_N = n - \dim \mathcal{V}_N, \\ \text{where } \tilde{X}_{42},\ \tilde{X}_{44} \text{ do not depend on } \lambda \text{ with } \det \tilde{X}_{44} \neq 0, \\ \text{and with } \det \tilde{X}_{11}(\lambda) \neq 0 \text{ if and only if } \lambda \in \mathbb{R} \setminus \sigma_{N-1}. \end{cases}$$

Next, let $\tilde{U}_N(\lambda) := \tilde{Q} U_N(\lambda) \mathcal{P} = \big(\tilde{U}_{\mu\nu}\big)$, $1 \leq \mu, \nu \leq 4$, with the above block structure. Then, by (15), $\tilde{U}_{\mu\nu} = \tilde{U}_{\mu\nu}(\lambda)$ do not depend on λ for $\nu = 3$ and $\nu = 4$, $\mu = 1, \dots, 4$. Since $\tilde{X}_N^T(\lambda)\tilde{U}_N(\lambda)$ is *symmetric* for all $\lambda \in \mathbb{R}$, we can conclude from (17) (with $\lambda = \lambda_1$ and within the block structure) that

$$0 = \big(\tilde{X}_N^T(\lambda_1)\tilde{U}_N(\lambda_1)\big)_{34}^T = \big(\tilde{X}_N^T(\lambda_1)\tilde{U}_N(\lambda_1)\big)_{43} = \tilde{X}_{44}^T \tilde{U}_{43}, \text{ so that } \tilde{U}_{43} = 0,$$

and that

$$\begin{aligned} 0 &= \big(\tilde{X}_N^T(\lambda_1)\tilde{U}_N(\lambda_1)\big)_{31}^T = \big(\tilde{X}_N^T(\lambda_1)\tilde{U}_N(\lambda_1)\big)_{13} = \tilde{X}_{11}^T(\lambda_1)\tilde{U}_{13} + \tilde{X}_{41}^T(\lambda_1)\tilde{U}_{43} \\ &= \tilde{X}_{11}^T(\lambda_1)\tilde{U}_{13}, \end{aligned}$$

so that $\tilde{U}_{13} = 0$ as well. Since $\operatorname{rank}\big(\tilde{X}_N^T(\lambda)\ \tilde{U}_N^T(\lambda)\big) = n$ for all $\lambda \in \mathbb{R}$, it follows that the rows are linearly independent, and we conclude for the 3rd block row that

$$\operatorname{rank}\big(\tilde{U}_{23}^T \ \ \tilde{U}_{33}^T\big)_{(n-r-\rho)\times(n-\tilde{r}-\rho)} = n - r - \rho.$$

Hence, there exists an orthogonal matrix $\bar{\mathcal{Q}} \in \mathbb{R}^{(n-\tilde{r}-\rho)\times(n-\tilde{r}-\rho)}$ such that

$$\bar{\mathcal{Q}} \begin{pmatrix} \tilde{U}_{23} \\ \tilde{U}_{33} \end{pmatrix} = \begin{pmatrix} 0 \\ \tilde{\tilde{U}}_{33} \end{pmatrix}$$

with $\det \tilde{\tilde{U}}_{33} \neq 0$. Since multiplication of $\tilde{\mathcal{Q}}$ from the left by the matrix

$$\begin{pmatrix} I_{\tilde{r}} & 0 & 0 \\ 0_{(n-\tilde{r}-\rho)\times\tilde{r}} & \bar{\mathcal{Q}} & 0 \\ 0 & 0 & I_\rho \end{pmatrix}$$

does not change the structure of (17), we can assume, in addition to (17), that

$$(18)\qquad \begin{cases} \tilde U_N(\lambda) := \tilde{\mathcal{Q}} U_N(\lambda)\mathcal{P} = \begin{pmatrix} \tilde U_{11}(\lambda) & \tilde U_{12}(\lambda) & 0 & \tilde U_{14} \\ \tilde U_{21}(\lambda) & \tilde U_{22}(\lambda) & 0 & \tilde U_{24} \\ \tilde U_{31}(\lambda) & \tilde U_{32}(\lambda) & \tilde U_{33} & \tilde U_{34} \\ \tilde U_{41}(\lambda) & \tilde U_{42}(\lambda) & 0 & \tilde U_{44} \end{pmatrix}, \\ \qquad \text{for all } \lambda \in \mathbb{R}, \text{ with } \det \tilde U_{33} \neq 0. \end{cases}$$

Finally, there exists an orthogonal matrix $\hat Q \in \mathbb{R}^{(n-r)\times(n-r)}$ such that $\hat Q \tilde U_{22}$ is upper triangular in the subsequent block structure, see (15). Since the multiplication of $\mathcal{Q}$ from the left by $\begin{pmatrix} I_r & 0 \\ 0_{(n-r)\times r} & \hat Q \end{pmatrix}$ does not change the structure of (15), we can in addition to (15), (17), and (18) assume that

$$(19)\qquad \begin{cases} U_{22} = \begin{pmatrix} \bar U_{33} & \bar U_{34} \\ 0 & \bar U_{44} \end{pmatrix} \quad \text{with } \det \bar U_{33} \det \bar U_{44} \neq 0, \\ \qquad \bar U_{33} \in \mathbb{R}^{(n-r-\rho)\times(n-r-\rho)}, \text{ and } \bar U_{44} \in \mathbb{R}^{\rho\times\rho}. \end{cases}$$

This completes the construction of the orthogonal matrices $\mathcal{P}$, $\mathcal{Q}$, and $\tilde{\mathcal{Q}}$.

4.2. **Consequences of the construction.** With the aid of the orthogonal matrices $\mathcal{Q}$ and $\tilde{\mathcal{Q}}$ we define

$$(20)\qquad \begin{cases} \tilde{\mathcal{A}}_N := \mathcal{Q}\mathcal{A}_N\tilde{\mathcal{Q}}^T &= \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \left(\bar A_{\mu\nu}\right)_{1\le\mu,\nu\le 4}, \quad \text{with} \\ & \qquad A_{11} \in \mathbb{R}^{r\times r},\ A_{22} \in \mathbb{R}^{(n-r)\times(n-r)}, \\ \tilde{\mathcal{B}}_N := \mathcal{Q}\mathcal{B}_N\tilde{\mathcal{Q}}^T &= \left(B_{\mu\nu}\right)_{1\le\mu,\nu\le 2} = \left(\bar B_{\mu\nu}\right)_{1\le\mu,\nu\le 4}, \\ \tilde{\mathcal{C}}_N := \mathcal{Q}\mathcal{C}_N\tilde{\mathcal{Q}}^T &= \left(C_{\mu\nu}\right)_{1\le\mu,\nu\le 2} = \left(\bar C_{\mu\nu}\right)_{1\le\mu,\nu\le 4}, \\ \tilde{\mathcal{D}}_N = \mathcal{Q}\mathcal{D}_N\tilde{\mathcal{Q}}^T &= \left(D_{\mu\nu}\right)_{1\le\mu,\nu\le 2}, \text{ and} \\ \tilde{\mathcal{W}}_N = \mathcal{Q}\mathcal{W}_N\mathcal{Q}^T &= \left(W_{\mu\nu}\right)_{1\le\mu,\nu\le 2} = \left(\bar W_{\mu\nu}\right)_{1\le\mu,\nu\le 4}, \end{cases}$$

where the block structure of the matrices is determined by the formulas (15) and (17), respectively. Then $\tilde{\mathcal{A}}_N$, $\tilde{\mathcal{B}}_N$, $\tilde{\mathcal{C}}_N$, $\tilde{\mathcal{D}}_N$, $\tilde{\mathcal{W}}_N$ satisfy also the conditions (1), because $\mathcal{Q}$ and $\tilde{\mathcal{Q}}$ are orthogonal. Moreover, it follows from (2) and (3) that, for all $\lambda \in \mathbb{R}$,

$$(21)\qquad \begin{cases} \tilde X_{N+1}(\lambda) &= \tilde{\mathcal{A}}_N \tilde X_N(\lambda) + \tilde{\mathcal{B}}_N \tilde U_N(\lambda), \\ \tilde U_{N+1}(\lambda) &= \tilde{\mathcal{C}}_N \tilde X_N(\lambda) + \tilde{\mathcal{D}}_N \tilde U_N(\lambda) - \lambda \tilde{\mathcal{W}}_N \tilde X_{N+1}(\lambda), \\ \tilde X_N(\lambda) &= \left(\tilde{\mathcal{D}}_N^T - \lambda \tilde{\mathcal{B}}_N^T \tilde{\mathcal{W}}_N\right)\tilde X_{N+1}(\lambda) - \tilde{\mathcal{B}}_N^T \tilde U_{N+1}(\lambda), \\ \tilde U_N(\lambda) &= \left(-\tilde{\mathcal{C}}_N + \lambda \tilde{\mathcal{A}}_N^T \tilde{\mathcal{W}}_N\right)\tilde X_{N+1}(\lambda) + \tilde{\mathcal{A}}_N^T \tilde U_{N+1}(\lambda). \end{cases}$$

First, it follows from the second column (within the block structure of (15)) of the *third* equation in (21) and from the formulas (15), (17), and (19) that

$$
\begin{gathered}
0=-B_{21}^T U_{22},\text{ hence } B_{21}=0 \text{ because } \det U_{22}\neq 0,\text{ and that}\\
\begin{pmatrix}0 & 0\\ 0 & \tilde X_{44}\end{pmatrix}=-\begin{pmatrix}\bar B_{33}^T & \bar B_{43}^T\\ \bar B_{34}^T & \bar B_{44}^T\end{pmatrix}\begin{pmatrix}\bar U_{33} & \bar U_{34}\\ 0 & \bar U_{44}\end{pmatrix},\quad\text{hence}\\
\bar B_{33}=0,\ \bar B_{43}=0,\ \bar B_{34}=0,\text{ and } \tilde X_{44}=-\bar B_{44}^T\bar U_{44},\text{ so that } \det\bar B_{44}\neq 0,
\end{gathered}
$$

because $\det\bar U_{33}\det\bar U_{44}\det\tilde X_{44}\neq 0$. Moreover, the third column of the *first* equation of (21), and the formulas (15), (17), and (18) imply that $0=-\tilde{\mathcal B}_N\big(0\ 0\ \tilde U_{33}^T\ 0\big)^T$. Hence, $\bar B_{\mu 3}=0$ for $\mu=1,\dots,4$, because $\det\tilde U_{33}\neq 0$. Altogether we have shown that

$$
\text{(22)}\qquad
\begin{cases}
\tilde{\mathcal B}_N=\begin{pmatrix}B_{11} & B_{12}\\ 0 & B_{22}\end{pmatrix},\ B_{12}=\begin{pmatrix}0 & \bar B_{14}\\ 0 & \bar B_{24}\end{pmatrix},\ B_{22}=\begin{pmatrix}0 & 0\\ 0 & \bar B_{44}\end{pmatrix}\\
\text{with } \det\bar B_{44}\neq 0,\ B_{11}\in\mathbb R^{r\times r},\ B_{22}\in\mathbb R^{(n-r)\times(n-r)},\\
\ \bar B_{14}\in\mathbb R^{\tilde r\times\rho},\ \bar B_{24}\in\mathbb R^{(r-\tilde r)\times\rho},\text{ and } \bar B_{44}\in\mathbb R^{\rho\times\rho}.
\end{cases}
$$

Next, the third column of the *second* equation of (21), and the formulas (15), (17), (18), and (19) imply that $\big(0\quad 0\quad \bar U_{33}^T\quad 0\big)^T=\tilde{\mathcal D}_N\big(0\quad 0\quad \tilde U_{33}^T\quad 0\big)^T$. Hence, $\bar D_{13}=0$, $\bar D_{23}=0$, $\bar D_{43}=0$, and $\det\bar D_{33}\neq 0$, because we have $\det\bar U_{33}\det\tilde U_{33}\neq 0$. By (1), the matrix $\tilde{\mathcal B}_N^T\tilde{\mathcal D}_N$ is symmetric. Hence, by (22), $B_{11}^TD_{11}$ is symmetric as well and $B_{11}^TD_{12}=\big(B_{12}^TD_{11}+B_{22}^TD_{21}\big)^T$, so that, in particular,

$$
B_{11}^T\begin{pmatrix}\bar D_{14}\\ \bar D_{24}\end{pmatrix}=\big\{\big(\bar B_{14}^T\ B_{24}^T\big)D_{11}+\bar B_{44}^T\big(\bar D_{41}\ \bar D_{42}\big)\big\}^T.
$$

Hence, since $\det\bar B_{44}\neq 0$,

$$
\begin{pmatrix}\bar D_{41}^T\\ \bar D_{24}^T\end{pmatrix}=B_{11}^T\begin{pmatrix}\bar D_{14}\\ \bar D_{14}\end{pmatrix}\bar B_{44}^{-1}-D_{11}^T\begin{pmatrix}\bar B_{14}\\ \bar B_{24}\end{pmatrix}\bar B_{44}^{-1}\in\operatorname{Im}\big(B_{11}^T\ D_{11}^T\big).
$$

Moreover, we have that $\tilde{\mathcal D}_N^T\tilde{\mathcal A}_N-\tilde{\mathcal B}_N^T\tilde{\mathcal C}_N=I$ by (1). It follows that

$$\begin{aligned} n &= \operatorname{rank}\bigl(\tilde{\mathcal{B}}_N^T \;\; \tilde{\mathcal{D}}_N^T\bigr) \\ &= \operatorname{rank}\begin{pmatrix} \star & \star & 0 & 0 & \star & \star & \circledast & \star \\ \star & \star & 0 & 0 & \star & \star & \circledast & \star \\ 0 & 0 & 0 & 0 & 0 & 0 & \bar{D}_{33}^T & 0 \\ \circledast & \circledast & 0 & \bar{B}_{44}^T & \circledast & \circledast & \circledast & \circledast \end{pmatrix} \\ &= \rho + n - r - \rho + \operatorname{rank}\begin{pmatrix} B_{11}^T & D_{11}^T & \begin{pmatrix} \bar{D}_{41}^T \\ \bar{D}_{42}^T \end{pmatrix} \end{pmatrix} \\ &= n - r + \operatorname{rank}\bigl(B_{11}^T \;\; D_{11}^T\bigr), \end{aligned}$$

where $\circledast$ blocks can be replaced by zero matrices without changing the rank, and, afterwards, all zero matrices can be erased, the blocks denoted by $\star$ remain.

Altogether we have shown that

$$\begin{cases} B_{11}^T D_{11} = D_{11}^T B_{11} \quad \text{and} \quad \operatorname{rank}\bigl(B_{11}^T \; D_{11}^T\bigr) = r, \\ \text{where } B_{11},\ D_{11} \in \mathbb{R}^{r\times r}. \end{cases} \tag{23}$$

Next, the third equation in (21), and the formulas (15), (17) and (22) imply that for all $\lambda \in \mathbb{R}$

$$Y_{11}(\lambda) = (D_{11}^T - \lambda B_{11}^T W_{11}) X_{11}(\lambda) - B_{11}^T U_{11}(\lambda), \tag{24}$$

where $Y_{11}(\lambda) := \begin{pmatrix} \tilde{X}_{11}(\lambda) & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{r\times r}$. For the final result of this part we use the following statement.

Lemma 1. *Assume* (1)*. Then for all* $\lambda \in \mathbb{R}$*,*

$$\begin{aligned} X_{N+1}^T(\lambda)\frac{d}{d\lambda}U_{N+1}(\lambda) - U_{N+1}^T(\lambda)\frac{d}{d\lambda}X_{N+1}(\lambda) \\ = -\sum_{k=0}^{N} X_{k+1}^T(\lambda)\mathcal{W}_k X_{k+1}(\lambda). \end{aligned}$$

Proof. The differentiability is clear, because all functions are polynomials in λ. Now, let $\lambda \in \mathbb{R}$, $k \in \mathbb{N}$. Then, we obtain from (1) and (2), where we omit the argument and put $\frac{d}{d\lambda} = {}'$, moreover, we suppress the index k, i.e., an index appears only when it is $k+1$:

$$\begin{aligned}
&(X_{k+1}^T U_{k+1}' - U_{k+1}^T X_{k+1}') - (X^T U' - U^T X') \\
=& X_{k+1}^T (\mathcal{C}X' - \mathcal{W}X_{k+1} - \lambda \mathcal{W} X_{k+1}' + \mathcal{D}U') - U_{k+1}^T X_{k+1}' - X^T U' + U^T X' \\
=& - X_{k+1}^T \mathcal{W} X_{k+1} - \lambda X_{k+1}^T \mathcal{W} X_{k+1}' + \lambda X_{k+1}^T \mathcal{W} X_{k+1}' - X^T U' + U^T X' \\
&+ (\mathcal{A}X + \mathcal{B}U)^T (\mathcal{C}X' + \mathcal{D}U') - (\mathcal{C}X + \mathcal{D}U)^T (\mathcal{A}X' + \mathcal{B}U') \\
=& - X_{k+1}^T \mathcal{W} X_{k+1} + X^T (-I + \mathcal{A}^T \mathcal{D} - \mathcal{C}^T \mathcal{B}) U' + U^T (I + \mathcal{B}^T \mathcal{C} - \mathcal{D}^T \mathcal{A}) X' \\
=& - X_{k+1}^T \mathcal{W} X_{k+1}.
\end{aligned}$$

The summation of this formula from $k = 0$ to $k = N$ leads to a telescope sum, and we obtain the assertion, because $\frac{d}{d\lambda} X_0(\lambda) = \frac{d}{d\lambda} U_0(\lambda) \equiv 0$. □

Now, since $\mathcal{W}_k \geq 0$ for $0 \leq k \leq N$, by (1) it follows from Lemma 1 and the formula (15), that for $\lambda \in \mathbb{R} \setminus \sigma_N$ with $\frac{d}{d\lambda} = '$ and by omitting the argument λ as above:

$$\begin{aligned}
\begin{pmatrix} \{U_{11} X_{11}^{-1}\}' & 0 \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} X_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}^T \begin{pmatrix} X_{11}^T U_{11}' - U_{11}^T X_{11}' & \star \\ \star & \star \end{pmatrix} \begin{pmatrix} X_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} X_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}^T \left\{ \begin{pmatrix} X_{11} & 0 \\ 0 & 0 \end{pmatrix}^T \begin{pmatrix} U_{11}' & 0 \\ U_{21}' & U_{22}' \end{pmatrix} \right. \\
&\quad \left. - \begin{pmatrix} U_{11}^T & U_{21}^T \\ 0 & U_{22}^T \end{pmatrix} \begin{pmatrix} X_{11}' & 0 \\ 0 & 0 \end{pmatrix} \right\} \begin{pmatrix} X_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} X_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}^T \mathcal{P}^T \left\{ X_{N+1}^T \mathcal{Q}^T \mathcal{Q} U_{N+1}' \right. \\
&\quad \left. - U_{N+1}^T \mathcal{Q}^T \mathcal{Q} X_{N+1}' \right\} \mathcal{P} \begin{pmatrix} X_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} \\
&\leq \begin{pmatrix} X_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}^T \mathcal{P}^T \left\{ -X_{N+1}^T \mathcal{W}_N X_{N+1} \right\} \mathcal{P} \begin{pmatrix} X_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} \\
&= - \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \tilde{\mathcal{W}}_N \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}.
\end{aligned}$$

Hence,

$$\frac{d}{d\lambda} \left\{ U_{11}(\lambda) X_{11}^{-1}(\lambda) \right\} \leq -W_{11}, \quad \text{for } \lambda \in \mathbb{R} \setminus \sigma_N. \tag{25}$$

4.3. **Discussion of the matrices $M_N(\lambda)$, $T_N(\lambda)$, and $D_N(\lambda)$.** First, we define

$$\tilde{M}_N(\lambda) := \mathcal{Q} M_N(\lambda) \tilde{\mathcal{Q}}^T, \quad \tilde{T}_N(\lambda) := \tilde{\mathcal{Q}} T_N(\lambda) \tilde{\mathcal{Q}}^T, \quad \tilde{D}_N(\lambda) = \tilde{\mathcal{Q}} D_N(\lambda) \tilde{\mathcal{Q}}^T.$$

Then we have that

$$\begin{cases} \tilde{M}_N(\lambda) & = (I - \tilde{X}_{N+1}(\lambda)\tilde{X}^{\dagger}_{N+1}(\lambda))\tilde{\mathcal{B}}_N, \\ \tilde{T}_N(\lambda) & = I - \tilde{M}^{\dagger}_N(\lambda)\tilde{M}_N(\lambda) \end{cases} \tag{26}$$

for $\lambda \in \mathbb{R}$, because $\tilde{X}^{\dagger}_{N+1}(\lambda) = \mathcal{P}^T X^{\dagger}_{N+1}(\lambda)\mathcal{Q}^T$ and $\tilde{M}^{\dagger}(\lambda) = \tilde{\mathcal{Q}} M^{\dagger}_N(\lambda)\mathcal{Q}^T$ by the orthogonality of $\mathcal{P}$, $\mathcal{Q}$, and $\tilde{\mathcal{Q}}$. Moreover, we have that

$$\tilde{D}_N(\lambda) = \tilde{T}_N(\lambda)\tilde{X}_N(\lambda)\tilde{X}^{\dagger}_{N+1}(\lambda)\tilde{\mathcal{B}}_N\tilde{T}_N(\lambda) \tag{27}$$

for all $\lambda \in \mathbb{R}$. It follows from (15) and (22) that

$$\tilde{M}_N(\lambda) = \begin{pmatrix} I - X_{11}(\lambda)X^{\dagger}_{11}(\lambda) & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix} \text{ for all } \lambda \in \mathbb{R}.$$

Since $B_{22} = \begin{pmatrix} 0 & 0 \\ 0 & \bar{B}_{44} \end{pmatrix}$ with $\det \bar{B}_{44} \neq 0$ and with $\bar{B}_{44} \in \mathbb{R}^{\rho\times\rho}$, we obtain that

$$M_N(\lambda) = M_N := \mathcal{Q}^T \begin{pmatrix} 0 & 0 \\ 0 & B_{22} \end{pmatrix} \tilde{\mathcal{Q}}, \quad \rho = \operatorname{rank} M_N(\lambda) = m_1(N,\lambda) \tag{28}$$

for $\lambda \in \mathbb{R} \setminus \sigma_N$. Hence, $\operatorname{rank} M_N(\lambda+) = \operatorname{rank} M_N(\lambda-) = \rho$ for all $\lambda \in \mathbb{R}$. Next, we have that $B_{12} = \begin{pmatrix} 0 & \bar{B}_{14} \\ 0 & \bar{B}_{24} \end{pmatrix}$ by (22), and therefore we get that

$$\operatorname{Ker} \tilde{M}_N(\lambda) = \operatorname{Ker} \begin{pmatrix} I - X_{11}(\lambda)X^{\dagger}_{11}(\lambda) & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} B_{11} & 0 \\ 0 & B_{22} \end{pmatrix} \tag{29}$$

for all $\lambda \in \mathbb{R}$.

We need also the following known result concerning generalized inverses, its proof can be found e.g. in [1, Ch. 2, Theorem 8 and Lemma 3].

Lemma 2. *Let $X, Y \in \mathbb{C}^{m\times n}$ be matrices with* $\operatorname{Ker} X = \operatorname{Ker} Y$. *Then $X^{\dagger}X = Y^{\dagger}Y$.*

Now, let $\bar{M}(\lambda) := (I - X_{11}(\lambda)X^{\dagger}_{11}(\lambda))B_{11}$. Then, we obtain from Lemma 2 and (29) that

$$\tilde{M}^{\dagger}_N(\lambda)\tilde{M}_N(\lambda) = \begin{pmatrix} \bar{M}^{\dagger}(\lambda)\bar{M}(\lambda) & 0 \\ 0 & B^{\dagger}_{22}B_{22} \end{pmatrix}$$

for $\lambda \in \mathbb{R}$, and from (22) we get that $B^{\dagger}_{22}B_{22} = \begin{pmatrix} 0 & 0 \\ 0 & I_\rho \end{pmatrix}$. Hence, we have that for all $\lambda \in \mathbb{R}$,

$$\tilde{T}_N(\lambda) = \begin{pmatrix} \bar{T}(\lambda) & 0 & 0 \\ 0 & I_{n-r-\rho} & 0 \\ 0 & 0 & 0_{\rho\times\rho} \end{pmatrix}, \tag{30}$$

where

$$\text{(31)}\qquad \begin{cases} \bar{T}(\lambda) = I - \bar{M}^{\dagger}(\lambda)\bar{M}(\lambda), \quad \bar{M}(\lambda) = (I - X_{11}(\lambda)X_{11}^{\dagger}(\lambda))B_{11} \in \mathbb{R}^{r\times r}, \\ \text{and, moreover, } \bar{M}(\lambda) = 0,\ \bar{T}(\lambda) = I \quad \text{for } \lambda \in \mathbb{R} \setminus \sigma_N. \end{cases}$$

Finally, we obtain from (27), (30), (31) and the formulas (15), (17), (22), and (24) that

$$\tilde{D}_N(\lambda) = \tilde{T}_N(\lambda) \begin{pmatrix} Y_{11}(\lambda) & 0 & 0 \\ 0 & 0 & 0 \\ \star & 0 & \star \end{pmatrix} \begin{pmatrix} X_{11}^{\dagger}(\lambda) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} B_{11} & 0 & \star \\ 0 & 0 & 0 \\ 0 & 0 & \star \end{pmatrix} \tilde{T}_N(\lambda).$$

Using also (24), we have shown that

$$\text{(32)}\qquad \begin{cases} \operatorname{ind} D_N(\lambda) &= \operatorname{ind} \bar{D}(\lambda) \quad \text{for all } \lambda \in \mathbb{R}, \text{ where} \\ \qquad \bar{D}(\lambda) &:= \bar{T}(\lambda)Y_{11}(\lambda)X_{11}^{\dagger}(\lambda)B_{11}\bar{T}(\lambda) \\ &= \bar{T}(\lambda)\left\{(D_{11}^T - \lambda B_{11}^T W_{11})X_{11}(\lambda) - B_{11}^T U_{11}(\lambda)\right\} X_{11}^{\dagger}(\lambda)B_{11}\bar{T}(\lambda). \end{cases}$$

4.4. **Application of the Index Theorem.** It remains to prove the crucial assertion (ii) of Proposition 1. To do so, we apply the *Index Theorem* [15, Cor. 3.4.2] or [16], and we use the same setting and we proceed in the same way as in [4, p. 1252] or [19, p. 143]. Let

$$D_{11} = B_{11}^T \tilde{S}_1 + S_2, \quad \operatorname{rank}\big(B_{11}^T\ S_2\big) = r = r_{N+1}, \quad \operatorname{Ker} S_2 = \operatorname{Im} B_{11},$$

where $\tilde{S}_1$ is symmetric (see [15, Cor. 3.1.3]), $m := r$, $t := \lambda_0 - \lambda$ with some given $\lambda_0 \in \mathbb{R}$,

$$\begin{gathered} R_1 := D_{11}^T - \lambda_0 B_{11}^T W_{11}, \quad R_2 := B_{11}^T, \quad X := X_{11}(\lambda_0), \quad U := -U_{11}(\lambda_0), \\ X(t) := X_{11}(\lambda_0 - t), \quad U(t) := -U_{11}(\lambda_0 - t), \\ R_1(t) := R_2 S_1(t) + S_2 \quad \text{with} \quad S_1(t) := \tilde{S}_1 + (t - \lambda_0)W_{11}, \\ M(t) := R_1(t)R_2^T + R_2 U(t) X^{-1}(t) R_2^T, \\ \Lambda(t) := R_1(t)X(t) + R_2 U(t), \quad \Lambda := R_1 X + R_2 U. \end{gathered}$$

Then, by (15), (23), and (25) the assumptions of [15, Theorem 3.4.1 and Cor. 3.4.2] or of [16] are satisfied. Use also that $\tilde{X}_{N+1}^T(\lambda)\tilde{U}_{N+1}(\lambda)$ is symmetric and that

$$\operatorname{rank}\big(\tilde{X}_{N+1}^T(\lambda)\ \ \tilde{U}_{N+1}^T(\lambda)\big) = n \quad \text{for } \lambda \in \mathbb{R}.$$

Moreover, it follows from (15), (17), (24), (30), (31), and (32) that for $t \in [-\varepsilon, \varepsilon] \setminus \{0\}$ with $\varepsilon > 0$ sufficiently small:

$$\begin{aligned}\Lambda(t) &= (D_{11}^T - \lambda B_{11}^T W_{11})B_{11} - B_{11}^T U_{11}(\lambda)\\ &= Y_{11}(\lambda) = \begin{pmatrix}\tilde X_{11}(\lambda) & 0_{\tilde r\times(r-\tilde r)}\\ 0 & 0\end{pmatrix},\\ M(t) &= \bar D(\lambda), \quad \bar T(\lambda) = I,\end{aligned}$$

so that

$$\begin{cases}\operatorname{ind} M(0+) = \operatorname{ind} \bar D(\lambda_0-) = \operatorname{ind} D_N(\lambda_0-),\\ \operatorname{ind} M(0-) = \operatorname{ind} \bar D(\lambda_0+) = \operatorname{ind} D_N(\lambda_0+),\\ \operatorname{def} \Lambda = r - \operatorname{rank} \tilde X_{11}(\lambda_0) = r_{N+1} + \rho - \operatorname{rank} X_N(\lambda_0),\\ \operatorname{def} \Lambda(0+) = r - \tilde r = r_{N+1} - r_N + \rho,\\ \operatorname{def} X = r - \operatorname{rank} X_{11}(\lambda_0) = r_{N+1} - \operatorname{rank} X_{N+1}(\lambda_0).\end{cases} \tag{33}$$

Moreover, suppose that S, S^*, T, Q are given as in [15, pp. 75] or [16, p. 118 with $\tilde S$ instead of S^*], namely:

$$R_2^T = XS + S^*, \quad X^T S^* = 0, \quad \operatorname{Im} T = \operatorname{Ker} S^*, \quad Q = T^T \Lambda S T,$$

where we choose (cf. [19, p. 144] or [15, p. 92]):

$$S = X^\dagger R_2^T, \quad S^* = R_2^T - XS = (I - XX^\dagger)R_2^T, \quad T = (S^*)^\dagger S^*.$$

It follows from (28), (29), (30), and (31) that

$$\begin{gathered}S^* = (I - X_{11}(\lambda_0)X_{11}^\dagger(\lambda_0))B_{11} = \bar M(\lambda_0), \quad T = \bar T(\lambda_0),\\ Q = T^T \Lambda S T = \bar T^T(\lambda_0) Y_{11}(\lambda_0) X_{11}^\dagger(\lambda_0) B_{11} \bar T(\lambda_0) = \bar D(\lambda_0),\end{gathered}$$

so that

$$\begin{cases}\operatorname{ind} Q = \operatorname{ind} D_N(\lambda_0),\\ \operatorname{rank} T = r - \operatorname{rank} S^* = r_{N+1} - \operatorname{rank} \bar M(\lambda_0),\\ \operatorname{rank} M_N(\lambda_0) = \rho + \operatorname{rank} \bar M(\lambda_0),\\ \rho = \operatorname{rank} M_N(\lambda_0\pm).\end{cases} \tag{34}$$

Now, [15, Cor. 3.4.2] or [16, Cor. 1], (33) and (34) imply that

$$\begin{aligned}\operatorname{ind} D_N(\lambda_0+) &= \operatorname{ind} M(0-) = \operatorname{ind} Q + m - \operatorname{rank} T\\ &= \operatorname{ind} D_N(\lambda_0) + r_{N+1} - r_{N+1} + \operatorname{rank} \bar M(\lambda_0)\\ &= \operatorname{ind} D_N(\lambda_0) + \operatorname{rank} M_N(\lambda_0) - \operatorname{rank} M_N(\lambda_0+),\end{aligned}$$

so that $\operatorname{rank} M_N(\lambda_0+) + \operatorname{ind} D_N(\lambda_0+) = \operatorname{rank} M_N(\lambda_0) + \operatorname{ind} D_N(\lambda_0)$, and that

$$\begin{aligned}\operatorname{ind} D_N(\lambda_0-) &= \operatorname{ind} M(0+) = \operatorname{ind} M(0-) + \operatorname{def}\Lambda - \operatorname{def}\Lambda(0+) - \operatorname{def} X \\ &= \operatorname{ind} D_N(\lambda_0+) + r_{N+1} + \rho - \operatorname{rank} X_N(\lambda_0) - r_{N+1} + r_N \\ &\quad -\rho - r_{N+1} + \operatorname{rank} X_{N+1}(\lambda_0) \\ &= \operatorname{ind} D_N(\lambda_0+) + \operatorname{rank} X_{N+1}(\lambda_0) - r_{N+1} - \operatorname{rank} X_N(\lambda_0) + r_N\end{aligned}$$

where $r_N = \operatorname{rank} X_N(\lambda_0\pm)$, $r_{N+1} = \operatorname{rank} X_{N+1}(\lambda_0\pm)$. This yields the assertion (ii) of Proposition 1, so this proof is complete. □

Acknowledgement. The authors thank the referee for his/her valuable remarks and suggestions which substantially contributed to the present version of the paper.

References

[1] A. Ben-Israel, E. N. Greville, Generalized Inverses, Theory and Applications, Springer-Verlag, New York, 2003.
[2] M. Bohner, O. Došlý, *Disconjugacy and transfromations for symplectic difference systems*, Rocky Mountain J. Math. **27** (1997), 707–743.
[3] M. Bohner, O. Došlý, W. Kratz, *Positive semidefiniteness of discrete quadratic functionals*, Proc. Edinburg Math. Soc. **46** (2003), 227–236.
[4] M. Bohner, O. Došlý, W. Kratz, *An oscillation theorem for discrete eigenvalue problems*, Rocky Mountain J. Math. **33** (2003), 1233–1260.
[5] D. L. Boley, P. Van Dooren, *Placing zeroes and the Kronecker canonical form*, Circuits Systems Signal Process, **13** (1994), 783–802.
[6] J. W. Demmel, Applied Numerical Linear Algebra, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1997.
[7] O. Došlý, R. Hilscher, V. Zeidan, *Nonnegativity of discrete quadratic functionals corresponding to symplectic difference systems*, Linear Algebra Appl. **375** (2003), 21–44.
[8] O. Došlý, W. Kratz, *A Sturmian separation theorem for symplectic difference systems*, to appear in J. Math. Anal. Appl. (2006).
[9] F. Gantmacher, Theory of Matrices, AMS Chelsea Publishing, Providence, RI, 1998
[10] I. Gohberg, P. Lancester, L. Rodman, Indefinite Linear Algebra and Applications. Birkhäuser Verlag, Basel, 2005.
[11] G. H. Golub, C. F. Van Loan, Matrix Computations, Second edition. Johns Hopkins Series in the Mathematical Sciences, 3. Johns Hopkins University Press, Baltimore, 1989.
[12] R. Hilscher, V. Zeidan, *Symplectic difference systems: variable stepsize discretization and discrete quadratic functionals*, Linear Algebra Appl. **367** (2004), 67–104.
[13] R. Hilscher, V. Zeidan, *Equivalent conditions to the nonnegativity of a quadratic functional in discrete optimal control*, Math. Nachr. **266** (2004), 48–59.
[14] R. Hilscher, V. Zeidan, *Nonnegativity and positivity of quadratic functionals in discrete calculus of variations*, J. Difference Equ. Appl. **11** (2005), 857–875.
[15] W. Kratz, Quadratic Functionals in Variational Analysis and Control Theory, Akademie Verlag, Berlin, 1995.
[16] W. Kratz, *An index theorem for monotone matrix-valued functions*, SIAM J. Matrix. Anal. Appl. **16**, (1995), 113–122.
[17] W. Kratz, *Sturm–Liouville difference equations and banded matrices,* Arch. Math. (Brno), **36** (2000), 499–505.
[18] W. Kratz, *Banded matrices and difference equations.* Linear Algebra Appl. **337** (2001), 1–20.

[19] W. Kratz, *Discrete oscillation*, J. Difference Equ. Appl. **9** (2003), 135–147.
[20] W. T. Reid, Ordinary Differential Equations, John Wiley & Sons, New York, 1971.
[21] G. Verghese, P. Van Dooren, T. Kailath, *Properties of the system matrix of a generalized state-space system*, Int. J. Control, **30** (1979), 235–243.
[22] P. Van Dooren, *The computation of Kronecker's canonical form of a singular pencil*, Linear Algebra Appl. **27** (1979), 103–140.
[23] P. Van Dooren, P. Dewilde, *The eigenstructure of an arbitrary polynomial matrix. Computational aspects*, Lin. Algebra Appl. **50** (1983), 545–579.

Department of Mathematics, Masaryk University, Janáčkovo nám. 2a, CZ-602 00 Brno, Czech Republic
E-mail address: dosly@math.muni.cz

Universität Ulm, Abteilung Angewandte Analysis, D–89069 Ulm, Germany
E-mail address: werner.kratz@uni-ulm.de