ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 33, Number 4, Winter 2003

AN OSCILLATION THEOREM FOR DISCRETE EIGENVALUE PROBLEMS

MARTIN BOHNER, ONDŘEJ DOŠLÝ AND WERNER KRATZ

ABSTRACT. In this paper we consider problems that consist of symplectic difference systems depending on an eigenvalue parameter, together with self-adjoint boundary conditions. Such symplectic difference systems contain as important cases linear Hamiltonian difference systems and also Sturm-Liouville difference equations of second and of higher order. The main result of this paper is an oscillation theorem that relates the number of eigenvalues to the number of generalized zeros of solutions.

1. Introduction. Consider the symplectic difference system

$$z_{k+1} = \mathcal{S}_k z_k, \quad k \in \mathbf{Z}, \tag{S}$$

where the $2n \times 2n$ matrices $\mathcal{S}_k$ are symplectic, i.e.,

$$\mathcal{S}_k^T \mathcal{J} \mathcal{S}_k = \mathcal{J} \quad \text{with} \quad \mathcal{J} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}.$$

Symplectic difference systems (S) cover a large variety of difference equations and systems, among them also linear Hamiltonian difference systems

$$\Delta x_k = A_k x_{k+1} + B_k u_k, \quad \Delta u_k = C_k x_{k+1} - A_k^T u_k,$$

where the $n \times n$ matrices B_k and C_k are symmetric and $I - A_k$ is nonsingular, as discussed, e.g., in the monograph by Ahlbrandt and Peterson [**2**]. This means, in turn, that systems (S) also cover higher order Sturm-Liouville difference equations

$$\sum_{\mu=0}^{n} (-\Delta)^\mu \{r_\mu(k) \Delta^\mu y_{k+1-\mu}\} = 0 \quad \text{with} \quad r_n(k) \neq 0,$$

1991 AMS *Mathematics Subject Classification.* 39A12, 39A13, 15A18.
Key words and phrases. Oscillation, symplectic, Hamiltonian, discrete systems, eigenvalue problem.
The second author is supported by the grant 201/01/0079 of the Grant Agency of Czech Republic.
Received by the editors on March 6, 2001, and in revised form on June 25, 2001.

in particular its special case, Sturm-Liouville second order difference equations

$$\Delta(r_k \Delta x_k) + p_k x_{k+1} = 0 \quad \text{with} \quad r_k \neq 0,$$

which are well studied in the recent literature, see [**1, 9**].

The principal aim of our paper is to investigate an eigenvalue problem where various boundary conditions are associated with the system

$$\text{(SE)} \qquad z_{k+1} = (\mathcal{S}_k - \lambda \hat{\mathcal{S}}_k) z_k,$$

where λ is a real parameter and

$$\mathcal{S}_k = \begin{pmatrix} \mathcal{A}_k & \mathcal{B}_k \\ \mathcal{C}_k & \mathcal{D}_k \end{pmatrix} \quad \text{and} \quad \hat{\mathcal{S}}_k = \begin{pmatrix} 0 & 0 \\ \mathcal{W}_k \mathcal{A}_k & \mathcal{W}_k \mathcal{B}_k \end{pmatrix},$$

$\mathcal{W}_k$ being nonnegative definite $n \times n$ matrices. Observe that (SE) is still a symplectic system of the form (S) for every $\lambda \in \mathbf{R}$, as can be verified by a direct computation. Our investigation can be viewed as a discrete counterpart of some results from the monograph by Kratz [**14**]. There the eigenvalue problem for linear Hamiltonian differential systems

$$\text{(H)} \qquad x' = A(t)x + B(t)u, \quad u' = (C(t) - \lambda \hat{C}(t))x - A^T(t)u,$$

where $B(t)$ and $C(t)$ are symmetric $n \times n$ matrices for $t \in \mathbf{R}$ with the boundary condition

$$\text{(B)} \qquad R_1 \begin{pmatrix} -x(a) \\ x(b) \end{pmatrix} + R_2 \begin{pmatrix} u(a) \\ u(b) \end{pmatrix} = 0,$$

R_1 and R_2 being $2n \times 2n$ matrices, was investigated (in a more general setting than presented here). A formula is proved there, which relates the number of focal points of a conjoined basis of (H) to the number of eigenvalues of (H), (B), which are less than a given λ, and the index (i.e., the number of negative eigenvalues) of a certain symmetric matrix associated with the boundary condition (B). For more details, see [**14**, Chapter 7].

In our paper we derive results which are in a certain sense discrete versions of this investigation, but under more restrictive assumptions on the dependence of the matrices in the investigated system on the parameter λ than those in [**14**]. The reason for the more restrictive assumptions is that some of the phenomena connected with oscillation theory of the discrete system are considerably more complicated than those associated with the continuous system (although often the theory in the discrete case is "easier" than in the continuous case). One just needs to compare the complicated definition of a focal point in the discrete case (see Definition 1 (iii) and (iv)) which has first been introduced in [**4, 5**], with its continuous counterpart (see [**14**, Definition 1.1.1 (ii)]) which is simply explained in terms of invertibility of the first part of the solution matrix. Because of the discrepancies between the continuous and the discrete, it will be of interest to eventually unify our results by using the concept of time scales (see [**8, 11**]) but this will be a topic of future research.

The paper is organized as follows. In the next section we recall some results from oscillation theory of (S), and we also present some basic facts of matrix theory (the Moore-Penrose generalized inverse) needed in our investigation. In this section we also state the main result of this paper, Theorem 1, the so-called *oscillation theorem*, which states that the number of focal points (i.e., "generalized" zeros) of a conjoined basis of (SE) (i.e., a matrix-valued solution) is equal to the number of eigenvalues less than λ of (SE) with $x_0 = x_{N+1} = 0$. The most technical part of this paper is contained in Section 3, where the proof of our main result is presented, via the so-called *local oscillation theorem*. Finally, in Section 4, we consider an eigenvalue problem consisting of (SE) together with more general boundary conditions. First a result corresponding to Theorem 1 is proved for separated boundary conditions, and in fact Theorem 1 is utilized to prove this more general oscillation theorem. Finally we use this theorem to derive the oscillation theorem for the case of (SE) together with arbitrary self-adjoint boundary conditions.

2. Notion and main result.

We consider the $2n$-dimensional vector *symplectic difference system*

$$z_{k+1} = \mathcal{S}_k(\lambda) z_k \quad \text{with} \quad \mathcal{S}_k(\lambda) = \mathcal{S}_k - \lambda \hat{\mathcal{S}}_k \quad \text{for} \quad k \in \mathbf{Z}, \tag{1}$$

where λ is a real parameter. Here

$$\mathcal{S}_k = \begin{pmatrix} \mathcal{A}_k & \mathcal{B}_k \\ \mathcal{C}_k & \mathcal{D}_k \end{pmatrix}, \quad \hat{\mathcal{S}}_k = \begin{pmatrix} 0 & 0 \\ \mathcal{W}_k\mathcal{A}_k & \mathcal{W}_k\mathcal{B}_k \end{pmatrix}, \quad \mathcal{J} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix},$$

where $\mathcal{A}_k, \mathcal{B}_k, \mathcal{C}_k, \mathcal{D}_k, \mathcal{W}_k \in \mathbf{R}^{n\times n}$ (i.e., they are real $n \times n$ matrices for each $k \in \mathbf{Z}$), I denotes the $n \times n$ identity matrix, and we put

$$z_k = (x_k, u_k) \quad \text{with vectors} \quad x_k, u_k \in \mathbf{R}^n.$$

With this notation our difference system reads as follows:

$$(1') \quad x_{k+1} = \mathcal{A}_k x_k + \mathcal{B}_k u_k, u_{k+1} = \mathcal{C}_k x_k - \lambda \mathcal{W}_k x_{k+1} + \mathcal{D}_k u_k \text{ for } k \in \mathbf{Z}.$$

Note that the first equation of $(1')$, the so-called *equation of motion*, does not depend on the parameter λ based on the special form of $\hat{\mathcal{S}}_k$. Because of analogies to the calculus of variations, we call the second equation of $(1')$ the *Euler equation*. Throughout we will assume that $\mathcal{S}_k$ is *symplectic*, i.e.,

$$\mathcal{S}_k^T \mathcal{J} \mathcal{S}_k = \mathcal{J} \quad \text{for } k \in \mathbf{Z},$$

and that $\mathcal{W}_k$ is symmetric and nonnegative definite, i.e., $\mathcal{W}_k \geq 0$. In summary our assumptions in terms of the matrices $\mathcal{A}_k, \mathcal{B}_k, \mathcal{C}_k, \mathcal{D}_k, \mathcal{W}_k$ read as follows (cf. [**6**, Remark 1]):

$$\text{(A1)} \qquad \begin{gathered} \mathcal{A}_k^T \mathcal{C}_k = \mathcal{C}_k^T \mathcal{A}_k, \quad \mathcal{B}_k^T \mathcal{D}_k = \mathcal{D}_k^T \mathcal{B}_k, \quad \mathcal{A}_k^T \mathcal{D}_k - \mathcal{C}_k^T \mathcal{B}_k = I, \\ \mathcal{W}_k \geq 0 \quad \text{for } k \in \mathbf{Z}. \end{gathered}$$

A simple calculation shows that, under these assumptions, the matrix $\mathcal{S}_k(\lambda)$ is symplectic for all $\lambda \in \mathbf{R}$.

Next we want to introduce the main notion where we use the following.

Notation. Let M be any (real) matrix. By $\operatorname{Ker} M$, $\operatorname{Im} M$, $\operatorname{rank} M$, $\operatorname{def} M$, $\operatorname{ind} M$, $\det M$ and $M^\dagger$, respectively, we denote the kernel of M, the image of M, the rank of M, the defect of M (i.e., the dimension of $\operatorname{Ker} M$), the index of M provided M is symmetric (i.e., the number of negative eigenvalues of M), the determinant of M provided M is a square matrix, and the Moore-Penrose inverse of M, cf. [**3**]. We write $M \geq 0$, as already above, $M > 0$ if the (real) matrix M is symmetric and nonnegative definite, positive definite, respectively.

Assume (A1) and let $\lambda \in \mathbf{R}$ be fixed. As above we denote vector-valued solutions $z = (z_k)_{k\in\mathbf{Z}} = (x, u) = (x_k, u_k)_{k\in\mathbf{Z}}$ of (1) or $(1')$ by small letters and we use capital letters for $2n\times n$ matrix-valued solutions $Z = (X,U) = (X_k,U_k)_{k\in\mathbf{Z}}$ of (1) or (1′) so that $X_k, U_k \in \mathbf{R}^{n\times n}$ for $k \in \mathbf{Z}$. For the symplectic system (1) the *Wronskian identity* (cf. [**6**]) holds, i.e., if $z = (x,u)$ and $\tilde{z} = (\tilde{x},\tilde{u})$ solve (1), then $z_k^T \mathcal{J} \tilde{z}_k = x_k^T \tilde{u}_k - u_k^T \tilde{x}_k$ is constant, in particular it equals zero for all $k \in \mathbf{Z}$ if it is zero for one $k \in \mathbf{Z}$.

Definition 1. Assume (A1) and let $\lambda \in \mathbf{R}$ be fixed.

(i) A $2n \times n$ matrix-valued solution $Z = (X,U) = (X_k,U_k)_{k\in\mathbf{Z}}$ is called a *conjoined basis* of (1) or (1′) if

$$X_k^T U_k - U_k^T X_k = 0, \quad \operatorname{rank} \begin{pmatrix} X_k \\ U_k \end{pmatrix} = n \quad \text{for } k \in \mathbf{Z}.$$

(ii) The conjoined basis $Z = (X,U)$ of (1) or (1′) with $X_0 = 0$, $U_0 = I$ is called the *principal solution* of (1) at 0, while the solution $\tilde{Z} = (\tilde{X},\tilde{U})$ of (1) with $\tilde{X}_0 = -I$, $\tilde{U}_0 = 0$ is called the *associated solution* of (1) at 0.

(iii) A conjoined basis $Z = (X,U)$ of (1) has *no focal point* in the interval $(k,k+1]$ for some $k \in \mathbf{Z}$ if

$$\operatorname{Ker} X_{k+1} \subset \operatorname{Ker} X_k \quad \text{and} \quad X_k X_{k+1}^\dagger \mathcal{B}_k \geq 0.$$

(iv) If a conjoined basis $Z = (X,U)$ of (1) has a focal point in the interval $(k,k+1)$ for some $k \in \mathbf{Z}$ and if $k+1$ is not a focal point of Z, i.e., if

$$\operatorname{Ker} X_{k+1} \subset \operatorname{Ker} X_k \quad \text{but} \quad X_k X_{k+1}^\dagger \mathcal{B}_k \ngeq 0,$$

then $\operatorname{ind} X_k X_{k+1}^\dagger \mathcal{B}_k$ is called the *multiplicity* of the focal point.

Remark 1. Assume (A1), let $\lambda \in \mathbf{R}$ be fixed, and let $Z = (X,U)$ be a conjoined basis of (1). We shall use the notation

$$Q_k := X_k X_k^\dagger U_k X_k^\dagger \quad \text{and} \quad D_k := X_k X_{k+1}^\dagger \mathcal{B}_k \tag{2}$$

for $k \in \mathbf{Z}$. Note that D_k is the same as $P_k[Q]$ in the notation of [**6**].

(i) First we repeat some facts from [**6**] and note some formulas. Using Definition 1 (i) and that $X_k X_k^\dagger$ is symmetric because of the properties of Moore-Penrose inverses, it follows that Q_k is symmetric.

From the difference equation (1) and our assumption (A1) we easily obtain the formula

$$X_k = (\mathcal{D}_k^T - \lambda \mathcal{B}_k^T \mathcal{W}_k) X_{k+1} - \mathcal{B}_k^T U_{k+1}. \tag{3}$$

This identity leads to the next formula (cf., [**6**, Lemma 3])

$$D_k = \mathcal{B}_k^T (\mathcal{D}_k - \lambda \mathcal{W}_k \mathcal{B}_k) - \mathcal{B}_k^T Q_{k+1} \mathcal{B}_k \quad \text{if} \quad \operatorname{Ker} X_{k+1} \subset \operatorname{Ker} X_k, \tag{4}$$

so that, by (A1),

$$D_k \quad \text{is symmetric if} \quad \operatorname{Ker} X_{k+1} \subset \operatorname{Ker} X_k;$$

because, by [**6**, Remark 1], $\mathcal{B}_k = X_{k+1} X_{k+1}^\dagger \mathcal{B}_k$ in this case.

(ii) Our Definition 1 (i), (ii) and (iii) is the same as in [**6**], while part (iv), the definition of the multiplicity of focal points, is new. But note that this is defined only if the "kernel condition" (i.e., $\operatorname{Ker} X_{k+1} \subset \operatorname{Ker} X_k$) is satisfied. Otherwise the problem of defining the multiplicity remains an open problem. Actually this lack of the definition leads to the "exceptional finite set" in our results below. We shall count the number of focal points in some interval always, as usual, including multiplicities.

(iii) Of course, the conjoined basis (X, U) depends in general on λ. If, for example, as for the principal and associated solutions at 0, the "initial" matrices X_0, U_0 do not depend on λ, then the matrix elements of $X_k = X_k(\lambda)$, $U_k = U_k(\lambda)$ are *polynomials* in λ for $k \in \mathbf{Z}$. Thus, as can easily be seen via suitable representations of Moore-Penrose inverses (cf. [**14**, Remark 3.3.2]) the matrix elements of the corresponding matrices $Q_k = Q_k(\lambda)$, $D_k = D_k(\lambda)$ are *rational functions* in λ.

We shall study the oscillatory behavior of the following eigenvalue problem (E), where $N \in \mathbf{N}$ is a given fixed integer:

$$\text{(E)} \qquad \begin{cases} x_{k+1} = \mathcal{A}_k x_k + \mathcal{B}_k u_k,\ u_{k+1} = \mathcal{C}_k x_k - \lambda \mathcal{W}_k x_{k+1} + \mathcal{D}_k u_k \\ \qquad\qquad \text{for } 0 \le k \le N \\ \text{with the boundary conditions } x_0 = x_{N+1} = 0. \end{cases}$$

As usual, λ is an eigenvalue of (E) if a nontrivial solution $z = (x, u) = (x_k, u_k)_{k=0}^{N+1}$ exists, a corresponding eigenvector of (E), i.e., z solves (E) and $k \in \{0, \dots, N+1\}$ exists with $(x_k, u_k) \neq (0, 0)$.

Remark 2. Let us make here some comments on the eigenvalue problem (E). To do this, assume (A1) and let $Z = (X, U)$ be the principal solution of (1) at 0 according to Definition 1 (ii).

(i) As can easily be seen, a number λ is an eigenvalue of (E) if and only if

$$\det X_{N+1}(\lambda) = 0,$$

and then the dimension of the kernel of $X_{N+1}(\lambda)$ (i.e., $\operatorname{def} X_{N+1}(\lambda)$) is its multiplicity. Similarly, as for focal points, we shall count the number of eigenvalues always *including multiplicities*.

(ii) Let $z = (x, u)$ and $\tilde{z} = (\tilde{x}, \tilde{u})$ solve (1) or (1′) for reals $\lambda = \lambda_0$ and $\lambda = \lambda_1$, respectively. Then using the assumption (A1), a simple computation leads to the formula

$$(\lambda_0 - \lambda_1) x_{k+1}^T \mathcal{W}_k \tilde{x}_{k+1} = \alpha_{k+1} - \alpha_k,$$

where $\alpha_k := x_k^T \tilde{u}_k - u_k^T \tilde{x}_k$ for $k \in \mathbf{Z}$. Hence, by the formula for a telescope sum,

$$(\lambda_0 - \lambda_1)\langle z, \tilde{z}\rangle = \alpha_{N+1} - \alpha_0$$

for a given $N \in \mathbf{N}$, where the product $\langle \cdot, \cdot \rangle$ is defined by

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

Therefore, if λ_0 and λ_1 are eigenvalues of (E) with corresponding eigenvectors z and $\tilde{z}$, it follows that

$$(\lambda_0 - \lambda_1)\langle z, \tilde{z}\rangle = 0,$$

because $x_0 = x_{N+1} = \tilde{x}_0 = \tilde{x}_{N+1} = 0$, so that $\alpha_0 = \alpha_{N+1} = 0$. Thus we have shown that *eigenvectors of* (E) *belonging to distinct eigenvalues are orthogonal.*

(iii) Assume additionally that $\det X_{N+1}(\lambda) \not\equiv 0$ (i.e., not every λ is an eigenvalue of (E) by part (i)); see assumption (A2) of Theorem 1 and Remark 3 (i) below. We prove that *all eigenvalues of* (E) *are real.* In view of this statement and of part (ii), the eigenvalue problem (E) is *self-adjoint.* Now let $\lambda_0 \in \mathbf{C}$ (of course, we have to deal with complex eigenvalues and eigenvectors here in contrast to the rest of this paper) be an eigenvalue of (E) with corresponding eigenvector $z = (x, u) = (x_k, u_k)_{k=0}^{N+1} \neq 0$, where $x_k, u_k \in \mathbf{C}$. Define

$$\alpha_k := \bar{x}_k^T u_k - \bar{u}_k^T x_k, \quad \beta_k := \bar{x}_{k+1}^T \mathcal{W}_k x_{k+1}.$$

Since $x_0 = x_{N+1} = 0$ by (E), we have that $\alpha_0 = \alpha_{N+1} = 0$. It follows from the difference equation $(1')$ of (E) and the assumption (A1) by a simple calculation that

$$(\bar{\lambda}_0 - \lambda_0)\beta_k = \alpha_{k+1} - \alpha_k \quad \text{for} \quad 0 \leq k \leq N.$$

Since $\alpha_0 = \alpha_{N+1} = 0$, we obtain that

$$(\bar{\lambda}_0 - \lambda_0) \sum_{k=0}^{N} \beta_k = 0.$$

If $\sum_{k=0}^{N} \beta_k = 0$, then $\beta_k = 0$ for $0 \leq k \leq N$ because $\beta_k \geq 0$ by (A1) for all k. Hence, $\mathcal{W}_k x_{k+1} = 0$ for $0 \leq k \leq N$, and therefore $z \neq 0$ satisfies (E) for all $\lambda \in \mathbf{C}$, so that every λ is an eigenvalue, which contradicts our additional assumption $\det X_{N+1}(\lambda) \not\equiv 0$. Thus $\sum_{k=0}^{N} \beta_k \neq 0$, so that $\bar{\lambda}_0 - \lambda_0 = 0$ (i.e., λ_0 is real) which is what we wanted to show. Note finally that we have also proven the following: if $z = (x_k, u_k)_{k=0}^{N+1}$ solves (E) for some number λ and if $\mathcal{W}_k x_{k+1} = 0$ for $0 \leq k \leq N$, then $z = 0$.

The main result of this paper reads as follows.

Theorem 1 (Oscillation theorem). *Assume* (A1) *and let* $Z = (X, U) = (X_k(\lambda), U_k(\lambda))_{k \in \mathbf{Z}}$ *be the principal solution at* 0 *of* (1). *Moreover, suppose that*

(A2) $$\lim_{\lambda \to -\infty} n_1(\lambda) = 0 \quad \textit{and} \quad \lim_{\lambda \to -\infty} n_2(\lambda) = 0$$

holds, where

$n_1(\lambda)$ *denotes the number of focal points of* (X, U) *in the interval* $(0, N+1]$,

$n_2(\lambda)$ *denotes the number of eigenvalues of* (E), *which are less than or equal to* λ.

Then

(5) $$n_1(\lambda) = n_2(\lambda) \quad \textit{for all} \quad \lambda \in \mathbf{R} \setminus \mathcal{N},$$

where the "exceptional" set

(6)

$$\mathcal{N} := \mathbf{R} \setminus \{\mu \in \mathbf{R} : \operatorname{rank} X_k(\mu) = \max_{\lambda \in \mathbf{R}} \operatorname{rank} X_k(\lambda) \textit{for } 0 \leq k \leq N+1\}$$

is finite.

Remark 3. First let us comment on the assumption (A2). To do so, assume (A1), and let (X, U) be the principal solution at 0 of (1), i.e., $X_0 = 0$, $U_0 = I$.

(i) The second part of (A2) (i.e., $\lim_{\lambda \to -\infty} n_2(\lambda) = 0$) simply means that $\lambda_0 \in \mathbf{R}$ exists such that $n_2(\lambda) = 0$ for $\lambda \leq \lambda_0$ and, by Remark 2, this is equivalent with

$$\det X_{N+1}(\lambda) \neq 0 \quad \text{for all} \quad \lambda \leq \lambda_0.$$

By Remark 1 (iii), $\det X_{N+1}(\lambda)$ is a polynomial in λ, and this is in turn equivalent with $\det X_{N+1}(\lambda) \not\equiv 0$ so that the eigenvalue problem (E) is *nondegenerate*, i.e., not every $\lambda \in \mathbf{R}$ is an eigenvalue of (E).

(ii) The first part of (A2) (i.e., $\lim_{\lambda \to -\infty} n_1(\lambda) = 0$) means that $\lambda_0 \in \mathbf{R}$ exists such that $n_1(\lambda) = 0$ for $\lambda \leq \lambda_0$, and by Definition 1 (iii) this means that

(7) $$\operatorname{Ker} X_{k+1}(\lambda) \subset \operatorname{Ker} X_k(\lambda) \quad \text{and} \quad D_k(\lambda) = X_k(\lambda) X_{k+1}^{\dagger}(\lambda) \mathcal{B}_k \geq 0$$

holds for all $0 \leq k \leq N$ and all $\lambda \leq \lambda_0$. Moreover, we shall see in the next section that the kernel condition holds for all $0 \leq k \leq N$ and $\lambda \notin \mathcal{N}$, so that $n_1(\lambda)$ is well defined by Definition 1 (iv) for $\lambda \notin \mathcal{N}$. The following representations of $n_1(\lambda)$ and $n_2(\lambda)$ follow directly from Definition 1 (ii), (iv) and Remark 2:

(8)

$$n_1(\lambda) = \sum_{k=0}^{N} \operatorname{ind} D_k(\lambda),\ n_2(\lambda) = \sum_{\substack{\mu \in \mathcal{N} \\ \mu \leq \lambda}} \operatorname{def} X_{N+1}(\mu) \quad \text{for all} \quad \lambda \in \mathbf{R} \setminus \mathcal{N}.$$

Finally it follows from the Reid roundabout theorem for symplectic systems [**6**, Theorem 1] that the second part of (A2), which is the same as assertion (7), is equivalent to

$$(9)\quad \begin{cases} \mathcal{F}_0(z,\lambda) := \sum_{k=0}^{N}\{x_k^T\mathcal{C}_k^T\mathcal{A}_kx_k + u_k^T\mathcal{D}_k^T\mathcal{B}_ku_k + 2u_k^T\mathcal{B}_k^T\mathcal{C}_kx_k\} \\ \qquad\qquad -\lambda\sum_{k=0}^{N}x_{k+1}^T\mathcal{W}_kx_{k+1} > 0 \\ \text{for all } \lambda \le \lambda_0 \text{ and for all } \textit{admissible } z = (x_k,u_k)_{k=0}^{N+1}, \text{ i.e.,} \\ x_{k+1} = \mathcal{A}_kx_k + \mathcal{B}_ku_k \text{ for } 0 \le k \le N \\ \text{and } x_0 = x_{N+1} = 0 \text{ with } x = (x_k)_{k=1}^N \ne 0. \end{cases}$$

Note that this positivity of $\mathcal{F}_0(z,\lambda)$ holds for all $\lambda \le \lambda_0$ if it is true just for $\lambda = \lambda_0$, because $\mathcal{W}_k \ge 0$ for all $k \in \mathbf{Z}$ by (A1).

(iii) Next we discuss the form of $\hat{\mathcal{S}}_k$ in our difference system (1). We shortly prove that the assumptions on $\hat{\mathcal{S}}_k$ are *necessary* in the following sense. First we require that the "equation of motion," that is the equation for x_{k+1} resulting from (1), does not depend on λ (which is important when considering the quadratic form in (ii)). Hence $\hat{\mathcal{S}}_k$ must be of the form

$$\hat{\mathcal{S}}_k = \begin{pmatrix} 0 & 0 \\ \hat{\mathcal{C}}_k & \hat{\mathcal{D}}_k \end{pmatrix} \quad \text{with certain } n \times n \text{ matrices } \hat{\mathcal{C}}_k, \hat{\mathcal{D}}_k.$$

Next we impose that $\mathcal{S}_k(\lambda)$ is symplectic, i.e., $\mathcal{S}_k^T(\lambda)\mathcal{J}\mathcal{S}_k(\lambda) = \mathcal{J}$ for all $\lambda \in \mathbf{R}$. Since $-\mathcal{J}\mathcal{J}^T = \mathcal{J}^2 = -I$, we obtain that $(\mathcal{S}_k^T(\lambda)\mathcal{J})^{-1} = \mathcal{S}_k(\lambda)\mathcal{J}^T$ and therefore $\mathcal{S}_k(\lambda)\mathcal{J}\mathcal{S}_k^T(\lambda) = \mathcal{J}$ for all $\lambda \in \mathbf{R}$. Altogether we obtain the following formulas, see also (A1), for all $k \in \mathbf{Z}$:

$$\begin{aligned} &\mathcal{A}_k^T\mathcal{C}_k = \mathcal{C}_k^T\mathcal{A}_k, \quad \mathcal{B}_k^T\mathcal{D}_k = \mathcal{D}_k^T\mathcal{B}_k, \quad \mathcal{A}_k^T\mathcal{D}_k - \mathcal{C}_k^T\mathcal{B}_k = I, \\ &\mathcal{A}_k^T\hat{\mathcal{C}}_k = \hat{\mathcal{C}}_k^T\mathcal{A}_k, \quad \mathcal{A}_k^T\hat{\mathcal{D}}_k = \hat{\mathcal{C}}_k^T\mathcal{B}_k, \quad \mathcal{B}_k^T\hat{\mathcal{D}}_k = \hat{\mathcal{D}}_k^T\mathcal{B}_k; \\ &\mathcal{A}_k\mathcal{B}_k^T = \mathcal{B}_k\mathcal{A}_k^T, \quad \mathcal{C}_k\mathcal{D}_k^T = \mathcal{D}_k\mathcal{C}_k^T, \quad \mathcal{A}_k\mathcal{D}_k^T - \mathcal{B}_k\mathcal{C}_k^T = I, \\ &\mathcal{A}_k\hat{\mathcal{D}}_k^T = \mathcal{B}_k\hat{\mathcal{C}}_k^T, \quad \mathcal{C}_k\hat{\mathcal{D}}_k^T = \mathcal{D}_k\hat{\mathcal{C}}_k^T, \quad \hat{\mathcal{C}}_k\hat{\mathcal{D}}_k^T = \hat{\mathcal{D}}_k\hat{\mathcal{C}}_k^T. \end{aligned}$$

Hence $\mathcal{A}_k\mathcal{D}_k^T - \mathcal{B}_k\mathcal{C}_k^T = I$ and $\operatorname{rank}(\mathcal{A}_k,\mathcal{B}_k) = n$ so that the matrix

$$K_k := \mathcal{A}_k\mathcal{A}_k^T + \mathcal{B}_k\mathcal{B}_k^T \quad \text{is invertible.}$$

Next $\operatorname{def}(\mathcal{A}_k,\mathcal{B}_k) = n = \operatorname{rank}\begin{pmatrix} \mathcal{B}_k^T \\ -\mathcal{A}_k^T \end{pmatrix}$ and $\mathcal{A}_k\mathcal{B}_k^T = \mathcal{B}_k\mathcal{A}_k^T$ so that $\operatorname{Ker}(\mathcal{A}_k,\mathcal{B}_k) = \operatorname{Im}\begin{pmatrix} \mathcal{B}_k^T \\ -\mathcal{A}_k^T \end{pmatrix}$. Since $\mathcal{A}_k\hat{\mathcal{D}}_k^T = \mathcal{B}_k\hat{\mathcal{C}}_k^T$, a matrix $\mathcal{W}_k \in \mathbf{R}^{n\times n}$ exists such that

$$\hat{\mathcal{D}}_k^T = \mathcal{B}_k^T\mathcal{W}_k \quad \text{and} \quad \hat{\mathcal{C}}_k^T = \mathcal{A}_k^T\mathcal{W}_k.$$

Finally we obtain from the above formulas that

$$K_k^T(\mathcal{W}_k^T - \mathcal{W}_k)K_k$$
$$= (\mathcal{A}_k, \mathcal{B}_k)\begin{pmatrix} \mathcal{A}_k^T\hat{\mathcal{C}}_k - \hat{\mathcal{C}}_k^T\mathcal{A}_k & \mathcal{A}_k^T\hat{\mathcal{D}}_k - \hat{\mathcal{C}}_k^T\mathcal{B}_k \\ \hat{\mathcal{D}}_k^T\mathcal{A}_k - \mathcal{B}_k^T\hat{\mathcal{C}}_k & \mathcal{B}_k^T\hat{\mathcal{D}}_k - \hat{\mathcal{D}}_k^T\mathcal{B}_k \end{pmatrix}\begin{pmatrix} \mathcal{A}_k^T \\ \mathcal{B}_k^T \end{pmatrix} = 0.$$

Thus $\mathcal{W}_k^T = \mathcal{W}_k$ because K_k is invertible, and therefore

$$\mathcal{W}_k \quad \text{is symmetric for all} \quad k \in \mathbf{Z}.$$

As we shall see in the next section, the nonnegativity of $\mathcal{W}_k$ is needed for the monotonicity of $D_k(\lambda)$, which is crucial for the whole theory, and which corresponds to the comment after statement (9) in Remark 3 (ii) above.

(iv) We conclude this remark with pointing out one of the applications of formula (5) (a similar comment also applies to the statements of Theorems 3 and 4 below; see formula (12)). Let $\lambda_0 \in \mathbf{R}$ be given. If we want to know how many eigenvalues of (E) are less than or equal to λ_0, we could calculate the principal solution (X, U) at 0 of (1) and determine the number of zeros of $\det X_{N+1}(\lambda)$ that are less than or equal to λ_0 (observe part (i) of this remark). However, $\det X_{N+1}(\lambda)$ is a polynomial in λ, and hence it might be difficult to calculate the number of its zeros that are less than or equal to λ_0. Alternatively, if the assumptions of Theorem 1 are satisfied, then we just need to calculate the principal solution of (1) at 0 for the particular λ_0 in question and count the number of its focal points in the interval $(0, N+1]$. Both calculating the principal solution recursively and counting the number of its focal points are easy tasks and can be done numerically. Moreover, this procedure may be used to treat numerically the algebraic eigenvalue problem for symmetric, banded matrices via Sturm-Liouville difference equations as discussed in Section 1 of [**16**], cf. also [**15**].

3. Proof of the main result. Assume (A1), let $N \in \mathbf{N}$ be fixed, and suppose that $(X, U) = (X_k(\lambda), U_k(\lambda))_{k\in\mathbf{Z}}$ is a conjoined basis of (1) such that $X_0(\lambda) \equiv X_0$ and $U_0(\lambda) \equiv U_0$ do not depend on λ. Moreover, we assume that a $\lambda_1 \in \mathbf{R}$ exists such that, compare (6),

(A3)
$$\begin{cases} r_k := \max_{\lambda\in\mathbf{R}} \operatorname{rank} X_k(\lambda) = \operatorname{rank} X_k(\lambda_1) \text{ for } 0 \le k \le N+1 \\ \text{and } \operatorname{Ker} X_{k+1}(\lambda_1) \subset \operatorname{Ker} X_k(\lambda_1) \text{ for } 0 \le k \le N. \end{cases}$$

In the sequel we derive a number of conclusions, which will lead to the "local oscillation theorem," below, and our main result, Theorem 1, is more or less a consequence of this local result.

(C1)

$\mathcal{N} := \mathbf{R}\backslash\{\lambda \in \mathbf{R} : \operatorname{rank} X_k(\lambda) = r_k \text{ for } 0 \leq k \leq N+1\}$ is a finite set.

Proof. By definition, $\lambda_0 \in \mathbf{R}$ exists such that $\operatorname{rank} X_k(\lambda_0) = r_k$. Hence there is a submatrix of $X_k(\lambda_0)$ of size $r_k \times r_k$ whose determinant is not zero. Since this subdeterminant of $X_k(\lambda)$ as a function of λ is a polynomial, it has finitely many zeros. Hence, $\mathcal{N}$ is a finite set. □

(C2) $\operatorname{Ker} X_k(\lambda) = \mathcal{V}_k := \operatorname{Ker} X_k(\lambda_1)$ for all $\lambda \in \mathbf{R} \setminus \mathcal{N}$ and $\mathcal{V}_k \subset \operatorname{Ker} X_k(\lambda)$ for all $\lambda \in \mathbf{R}$ and all $0 \leq k \leq N+1$.

Proof. Let $k \in \{0, \dots, N+1\}$, $c \in \mathcal{V}_k$ and put $x_\mu(\lambda) = X_\mu(\lambda)c$, $u_\mu(\lambda) = U_\mu(\lambda)c$, $x_\mu = x_\mu(\lambda_1)$, $u_\mu = u_\mu(\lambda_1)$ for $0 \leq \mu \leq k$. Then, by (A3), $x_0 = \cdots = x_k = 0$. We prove by induction that

$$x_\mu(\lambda) = x_\mu = 0,\ u_\mu(\lambda) = u_\mu \quad \text{for} \quad 0 \leq \mu \leq k.$$

This is clear for $\mu = 0$, because X_0 and U_0 do not depend on λ. It follows inductively for $0 \leq \mu < k$, using $(1')$, that

$$\begin{aligned} x_{\mu+1}(\lambda) &= \mathcal{A}_\mu x_\mu(\lambda) + \mathcal{B}_\mu u_\mu(\lambda) = \mathcal{A}_\mu x_\mu + \mathcal{B}_\mu u_\mu = x_{\mu+1} = 0, \\ u_{\mu+1}(\lambda) &= \mathcal{C}_\mu x_\mu(\lambda) - \lambda \mathcal{W}_\mu x_{\mu+1}(\lambda) + \mathcal{D}_\mu u_\mu(\lambda) = \mathcal{D}_\mu u_\mu = u_{\mu+1}. \end{aligned}$$

Hence, $c \in \operatorname{Ker} X_k(\lambda)$ for all $\lambda \in \mathbf{R}$ so that $\mathcal{V}_k \subset \operatorname{Ker} X_k(\lambda)$. Moreover, $\mathcal{V}_k = \operatorname{Ker} X_k(\lambda)$ if $\lambda \notin \mathcal{N}$ because

$$\dim \mathcal{V}_k = n - r_k = \dim \operatorname{Ker} X_k(\lambda) \quad \text{for} \quad \lambda \notin \mathcal{N}$$

by the definition of r_k in (A3). □

Based on statement (C2), we can undertake the following.

Construction. Starting with an orthonormal basis of $\mathcal{V}_{N+1} = \operatorname{Ker} X_{N+1}(\lambda_1)$ we successively supplement an orthonormal basis of $\mathcal{V}_{k+1}$

to such a basis of $\mathcal{V}_k$ for $k = N, \dots, 0$. This is possible because $\mathcal{V}_{k+1} \subset \mathcal{V}_k$ by (A3). Using (C2) we can conclude that an *orthogonal matrix* $\mathcal{P} \in \mathbf{R}^{n\times n}$ exists such that

$$X_k(\lambda)\mathcal{P} = (\underbrace{*}_{r_k} 0) \quad \text{for all} \quad \lambda \in \mathbf{R} \text{ and } 0 \le k \le N+1.$$

Note that, by (A3), $0 \le r_0 \le r_1 \le \cdots \le r_{N+1} \le n$.

Next, using Gram-Schmidt orthogonalization, we may choose *orthogonal matrices* $\mathcal{Q}_k \in \mathbf{R}^{n\times n}$ for $0 \le k \le N+1$ such that, use (A3),

$$\tilde{X}_k := \mathcal{Q}_k X_k(\lambda_1)\mathcal{P} = \begin{pmatrix} X_{11} & 0 \\ 0 & 0 \end{pmatrix},$$

where $X_{11} = X_{11}(k, \lambda_1) \in \mathbf{R}^{r_k \times r_k}$ is invertible.

Let

$$\tilde{U}_k := \mathcal{Q}_k U_k(\lambda_1)\mathcal{P} = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix}, \quad \text{where} \quad U_{11} \in \mathbf{R}^{r_k \times r_k}.$$

Then, by Definition 1 (i),

$$\tilde{X}_k^T \tilde{U}_k = \begin{pmatrix} X_{11}^T U_{11} & X_{11}^T U_{12} \\ 0 & 0 \end{pmatrix} = \mathcal{P}^T X_k^T(\lambda_1) U_k(\lambda_1)\mathcal{P} \text{ is symmetric.}$$

Hence, $U_{12} = 0$ because X_{11} is invertible. Moreover,

$$n = \operatorname{rank}(X_k^T(\lambda_1), U_k^T(\lambda_1)) = \operatorname{rank}\begin{pmatrix} X_{11}^T & U_{11}^T & U_{21}^T \\ 0 & 0 & U_{22}^T \end{pmatrix},$$

and therefore U_{22} is invertible. By [**12**, page 114, Exercise on QR factorization], an orthogonal matrix $\mathcal{Q}$ exists such that $U_{22}^{-1}\mathcal{Q}$ is *lower triangular*, and then, of course, $\mathcal{Q}^T U_{22}$ is also. Hence, we may choose the orthogonal matrices $\mathcal{Q}_k$ in such a way that U_{22} is lower triangular. This completes our construction.

Now we define new matrices $\tilde{X}_k(\lambda), \tilde{U}_k(\lambda)$, etc., and we arrange a block structure with the agreement that certain blocks do not occur if $r_k = 0$ or $r_k = n$, which was already presupposed above. For $0 \le k \le N+1$, respectively, $\le N$ and $\lambda \in \mathbf{R}$, we put

$$\tilde{X}_k(\lambda) := \mathcal{Q}_k X_k(\lambda)\mathcal{P} = \begin{pmatrix} X_{11}(k,\lambda) & X_{12}(k,\lambda) \\ X_{21}(k,\lambda) & X_{22}(k,\lambda) \end{pmatrix} = \begin{pmatrix} \tilde{X}_{11}(k,\lambda) & \tilde{X}_{12}(k,\lambda) \\ \tilde{X}_{21}(k,\lambda) & \tilde{X}_{22}(k,\lambda) \end{pmatrix}$$

where

$$X_{11}(k,\lambda) \in \mathbf{R}^{r_k \times r_k} \quad \text{and} \quad \tilde{X}_{11}(k,\lambda) \in \mathbf{R}^{r_{k+1} \times r_{k+1}}$$

and $\tilde{U}_k(\lambda) := \mathcal{Q}_k U_k(\lambda)\mathcal{P} = (U_{\mu\nu}(k,\lambda)) = (\tilde{U}_{\mu\nu}(k,\lambda))$ with the same block structure. Moreover, we define

$$\tilde{\mathcal{A}}_k := \mathcal{Q}_{k+1}\mathcal{A}_k\mathcal{Q}_k^T = \begin{pmatrix} A_{11}(k) & A_{12}(k) \\ A_{21}(k) & A_{22}(k) \end{pmatrix} \text{ where } A_{11}(k) \in \mathbf{R}^{r_{k+1} \times r_{k+1}},$$

and with the same block structure

$$\tilde{\mathcal{B}}_k := \mathcal{Q}_{k+1}\mathcal{B}_k\mathcal{Q}_k^T = (B_{\mu\nu}(k)), \quad \tilde{\mathcal{C}}_k := \mathcal{Q}_{k+1}\mathcal{C}_k\mathcal{Q}_k^T = (C_{\mu\nu}(k)),$$
$$\tilde{\mathcal{D}}_k := \mathcal{Q}_{k+1}\mathcal{D}_k\mathcal{Q}_k^T = (D_{\mu\nu}(k)), \quad \tilde{\mathcal{W}}_k := \mathcal{Q}_{k+1}\mathcal{W}_k\mathcal{Q}_{k+1}^T = (W_{\mu\nu}(k)).$$

We continue with our conclusions and with the understanding that they are valid for all $0 \le k \le N+1$, respectively $\le N$.

(C3) $(\tilde{X}, \tilde{U}) = (\tilde{X}_k, \tilde{U}_k)_{k=0}^{N+1}$ is a conjoined basis of the *symplectic* difference system

$$\begin{cases} \tilde{X}_{k+1}(\lambda) = \tilde{\mathcal{A}}_k\tilde{X}_k(\lambda) + \tilde{\mathcal{B}}_k\tilde{U}_k(\lambda), \\ \tilde{U}_{k+1}(\lambda) = \tilde{\mathcal{C}}_k\tilde{X}_k(\lambda) - \lambda\tilde{\mathcal{W}}_k\tilde{X}_{k+1}(\lambda) + \tilde{\mathcal{D}}_k\tilde{U}_k(\lambda), \end{cases}$$

and $\tilde{\mathcal{W}}_k \ge 0$.

Proof. $\tilde{X}_k^T(\lambda)\tilde{U}_k(\lambda) = \mathcal{P}^T X_k^T(\lambda)U_k(\lambda)\mathcal{P}$ is symmetric,

$$\operatorname{rank}(\tilde{X}_k^T(\lambda), \tilde{U}_k^T(\lambda)) = \operatorname{rank}(X_k^T(\lambda), U_k^T(\lambda)) = n$$

by Definition 1 (i). The system is symplectic, because by (A1),

$$\tilde{\mathcal{A}}_k^T\tilde{\mathcal{C}}_k = \mathcal{Q}_k\mathcal{A}_k^T\mathcal{C}_k\mathcal{Q}_k^T, \tilde{\mathcal{B}}_k^T\tilde{\mathcal{D}}_k = \mathcal{Q}_k\mathcal{B}_k^T\mathcal{D}_k\mathcal{Q}_k^T \quad \text{are symmetric,}$$

and $\tilde{\mathcal{A}}_k^T\tilde{\mathcal{D}}_k - \tilde{\mathcal{C}}_k^T\tilde{\mathcal{B}}_k = \mathcal{Q}_k(\mathcal{A}_k^T\mathcal{D}_k - \mathcal{C}_k^T\mathcal{B}_k)\mathcal{Q}_k^T = \mathcal{Q}_k\mathcal{Q}_k^T = I$. $\square$

(C4) $X_{12}(k,\lambda) = 0$, $X_{22}(k,\lambda) = 0$ for all $\lambda \in \mathbf{R}$.

Proof. This holds by the construction of $\mathcal{P}$. $\square$

(C5) $X_{11}(k) := X_{11}(k,\lambda_1)$ is invertible, $X_{21}(k,\lambda_1) = 0$, $U_{12}(k,\lambda_1) = 0$, and $U_{22}(k) := U_{22}(k,\lambda_1)$ is an invertible, lower triangular matrix.

Proof. This is true by the construction of $\mathcal{Q}_k$ and the corresponding calculations. $\square$

(C6) $\tilde{X}_{11}(k) := \tilde{X}_{11}(k,\lambda_1) = \left(\begin{smallmatrix} X_{11}(k) & 0 \\ 0 & 0 \end{smallmatrix}\right)$, $\tilde{X}_{21}(k,\lambda_1) = 0$, $\tilde{X}_{12}(k,\lambda) = 0$ and $\tilde{X}_{22}(k,\lambda) = 0$ for all $\lambda \in \mathbf{R}$, $\tilde{U}_{12}(k,\lambda_1) = 0$, $\tilde{U}_{22}(k) := \tilde{U}_{22}(k,\lambda_1)$ is an invertible, lower triangular matrix. Moreover,

$$\tilde{U}_k(\lambda_1) = \begin{pmatrix} U_{11}(k,\lambda_1) & 0 \\ U_{21}(k,\lambda_1) & U_{22}(k) \end{pmatrix} \text{ where } U_{22}(k) \in \mathbf{R}^{(n-r_k)\times(n-r_k)},$$
$$U_{22}(k) = \begin{pmatrix} \tilde{\tilde{U}}_{22}(k) & 0 \\ * & \tilde{U}_{22}(k) \end{pmatrix} \text{ where } \tilde{\tilde{U}}_{22}(k) \in \mathbf{R}^{(r_{k+1}-r_k)\times(r_{k+1}-r_k)},$$

such that $\tilde{\tilde{U}}_{22}(k)$ is also an invertible, lower triangular matrix.

Proof. This follows directly from (C5) and the arranged block structure. $\square$

(C7) $B_{12}(k) = D_{12}(k) = 0$, $A_{21}(k) = B_{21}(k) = 0$, $B_{22}(k) = 0$, $A_{22}^T(k)D_{22}(k) = I$ so that $A_{22}(k)$ and $D_{22}(k)$ are invertible.

Proof. It follows from (C3), (C4), (C5) and (C6) that

$$\begin{aligned}
&0 = X_{12}(k+1,\lambda_1) = B_{12}(k)\tilde{U}_{22}(k) \\
&\qquad \text{so that } B_{12}(k) = 0 \text{ (since } \tilde{U}_{22}(k) \text{ is invertible)}, \\
&0 = X_{22}(k+1,\lambda_1) = B_{22}(k)\tilde{U}_{22}(k) \text{ so that } B_{22}(k) = 0, \\
&0 = U_{12}(k+1,\lambda_1) = D_{12}(k)\tilde{U}_{22}(k) \text{ so that } D_{12}(k) = 0.
\end{aligned}$$

Hence by (C3), in particular simplecticity, and from what we have shown, we have

$$\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} A_{11}^T & A_{21}^T \\ A_{12}^T & A_{22}^T \end{pmatrix}(k) \begin{pmatrix} D_{11} & 0 \\ D_{21} & D_{22} \end{pmatrix}(k) \\ - \begin{pmatrix} C_{11}^T & C_{21}^T \\ C_{12}^T & C_{22}^T \end{pmatrix}(k) \begin{pmatrix} B_{11} & 0 \\ B_{21} & 0 \end{pmatrix}(k),$$

and this implies that

$$I = A_{22}^T(k)D_{22}(k),\ 0 = A_{21}^T(k)D_{22}(k) \text{ so that } A_{21}(k) = 0,$$

because $D_{22}(k)$ is invertible by the first equation. Moreover,

$$\begin{pmatrix} B_{11}^T & B_{21}^T \\ 0 & 0 \end{pmatrix}(k) \begin{pmatrix} D_{11} & 0 \\ D_{21} & D_{22} \end{pmatrix}(k) \text{ is symmetric,}$$

and therefore $B_{21}^T(k)D_{22}(k) = 0$ so that $B_{21}(k) = 0$. □

(C8) The matrix $\begin{pmatrix} A_{11}(k) & B_{11}(k) \\ C_{11}(k) & D_{11}(k) \end{pmatrix} - \lambda \begin{pmatrix} 0 & 0 \\ W_{11}(k)A_{11}(k) & W_{11}(k)B_{11}(k) \end{pmatrix}$ is symplectic for all $\lambda \in \mathbf{R}$ and $W_{11}(k) \geq 0$.

Proof. We have to prove (A1) for the corresponding matrices. First $W_{11}(k) \geq 0$ by (C3) and its definition. By (A1) and using (C3) and (C7), we have that

$$\tilde{\mathcal{A}}_k^T \tilde{\mathcal{C}}_k = \begin{pmatrix} A_{11}^T(k) & 0 \\ A_{12}^T(k) & A_{22}^T(k) \end{pmatrix} \begin{pmatrix} C_{11}(k) & C_{12}(k) \\ C_{12}(k) & C_{22}(k) \end{pmatrix} \quad \text{and}$$
$$\tilde{\mathcal{B}}_k^T \tilde{\mathcal{D}}_k = \begin{pmatrix} B_{11}^T(k) & 0 \\ 0 & 0cr \end{pmatrix} \begin{pmatrix} D_{11}(k) & 0 \\ D_{12}(k) & D_{22}(k) \end{pmatrix} \quad \text{are symmetric,}$$

and $\tilde{\mathcal{A}}_k^T \tilde{\mathcal{D}}_k - \tilde{\mathcal{C}}_k^T \tilde{\mathcal{B}}_k = I$. Hence $A_{11}^T(k)C_{11}(k)$ and $B_{11}^T(k)D_{11}(k)$ are symmetric, and $A_{11}^T(k)D_{11}(k) - C_{11}^T(k)B_{11}(k) = I$. □

(C9) $U_{12}(k,\lambda)$ and $U_{22}(k,\lambda)$ do not depend on λ so that $U_{12}(k,\lambda) = U_{12}(k,\lambda_1) = 0$ and $U_{22}(k,\lambda) = U_{22}(k)$ is an invertible, lower triangular matrix for all $\lambda \in \mathbf{R}$.

Proof. This assertion is true for $k = 0$ by our construction, see (C5). It follows inductively, using (C3), (C6) and (C7) that

$$\begin{aligned} U_{12}(k+1,\lambda) &= D_{11}(k)\tilde{U}_{12}(k,\lambda) + D_{12}(k)\tilde{U}_{22}(k,\lambda) \\ &= D_{11}(k)\tilde{U}_{12}(k,\lambda) = 0, \\ U_{22}(k+1,\lambda) &= D_{21}(k)\tilde{U}_{12}(k,\lambda) + D_{22}(k)\tilde{U}_{22}(k,\lambda) \\ &= D_{22}(k)\tilde{U}_{22}(k,\lambda) = U_{22}(k+1,\lambda_1) = U_{22}(k+1) \end{aligned}$$

for all $\lambda \in \mathbf{R}$. $\square$

(C10) $X_{21}(k,\lambda) = 0$ for all $\lambda \in \mathbf{R}$ and $X_{11}(k,\lambda)$ is invertible for all $\lambda \in \mathbf{R} \setminus \mathcal{N}$.

Proof. It follows from (C3), (C4) and (C9) that

$$\tilde{X}_k^T(\lambda)\tilde{U}_k(\lambda) = \begin{pmatrix} * & X_{21}^T(k,\lambda)U_{22}(k) \\ 0 & 0 \end{pmatrix} \text{ is symmetric.}$$

Hence $X_{21}^T(k,\lambda)U_{22}(k) = 0$ so that $X_{21}(k,\lambda) = 0$ for all $\lambda \in \mathbf{R}$ because $U_{22}(k)$ is invertible. This last conclusion, (A3), and (C2) imply that

$$\operatorname{rank} X_{11}(k,\lambda) = \operatorname{rank} \tilde{X}_k(\lambda) = \operatorname{rank} X_k(\lambda) = \operatorname{rank} X_k(\lambda_1) = r_k$$

so that $X_{11}(k,\lambda)$ is invertible for all $\lambda \in \mathbf{R} \setminus \mathcal{N}$. $\square$

(C11) We have for all $\lambda \in \mathbf{R}$:

$$\tilde{X}_k(\lambda) = \begin{pmatrix} X_{11}(k,\lambda) & 0 \\ 0 & 0 \end{pmatrix}, \qquad \tilde{U}_k(\lambda) = \begin{pmatrix} U_{11}(k,\lambda) & 0 \\ U_{21}(k,\lambda) & U_{22}(k) \end{pmatrix},$$

$$\tilde{X}_{11}(k,\lambda) = \begin{pmatrix} X_{11}(k,\lambda) & 0 \\ 0 & 0 \end{pmatrix}, \qquad \tilde{U}_{11}(k,\lambda) = \begin{pmatrix} U_{11}(k,\lambda) & 0 \\ * & \tilde{\tilde{U}}_{22}(k) \end{pmatrix},$$

$$\begin{aligned} X_{11}(k+1,\lambda) &= A_{11}(k)\tilde{X}_{11}(k,\lambda) + B_{11}(k)\tilde{U}_{11}(k,\lambda), \\ U_{11}(k+1,\lambda) &= C_{11}(k)\tilde{X}_{11}(k,\lambda) - \lambda W_{11}(k)X_{11}(k+1,\lambda) \\ &\quad + D_{11}(k)\tilde{U}_{11}(k,\lambda), \end{aligned}$$

$$X_{11}^T(k,\lambda)U_{11}(k,\lambda) \text{ and } \tilde{X}_{11}^T(k,\lambda)\tilde{U}_{11}(k,\lambda) \text{ are symmetric,}$$

$$\operatorname{rank}\,(X_{11}^T(k,\lambda), U_{11}^T(k,\lambda)) = r_k.$$

Proof. Besides the last assertion, the other statements follow from what we have shown so far, more precisely from (C4), (C10), (C9), (C6), (C7) and (C3). The last assertion follows by induction: $\operatorname{rank}(X_{11}^T(0), U_{11}^T(0)) = \operatorname{rank} X_{11}(0) = r_0$, and

$$\begin{aligned}\operatorname{rank}&(X_{11}^T(k+1,\lambda), U_{11}^T(k+1,\lambda))\\ &= \operatorname{rank}(\tilde{X}_{11}^T(k,\lambda), \tilde{U}_{11}^T(k,\lambda))\\ &= \operatorname{rank}\begin{pmatrix} X_{11}^T(k,\lambda) & U_{11}^T(k,\lambda) & * \\ 0 & 0 & \tilde{\tilde{U}}_{22}(k)\end{pmatrix}\\ &= \operatorname{rank}(X_{11}^T(k,\lambda), U_{11}^T(k,\lambda)) + \operatorname{rank}\tilde{\tilde{U}}_{22}(k)\\ &= r_k + (r_{k+1} - r_k) = r_{k+1},\end{aligned}$$

because $\tilde{\tilde{U}}_{22}(k)$ is invertible by (C6). □

(C12) $Q_{11}(k+1,\lambda) := U_{11}(k+1,\lambda)X_{11}^{-1}(k+1,\lambda)$ is symmetric and

$$\frac{d}{d\lambda} Q_{11}(k+1,\lambda) \le -W_{11}(k) \le 0 \quad \text{for all} \quad \lambda \in \mathbf{R} \setminus \mathcal{N}.$$

Proof. First $Q_{11}(k+1,\lambda)$ is symmetric by (C11). In the following calculation we omit the arguments (k), (k,λ), $(k+1,\lambda)$, respectively, and we put $\prime = (d/d\lambda)$. Using (C11) we obtain that

$$\begin{aligned}\frac{d}{d\lambda} Q_{11}(k+1,\lambda) &= U_{11}' X_{11}^{-1} - U_{11}X_{11}^{-1}X_{11}'X_{11}^{-1}\\ &= (X_{11}^T)^{-1}(X_{11}^T U_{11}' - U_{11}^T X_{11}')X_{11}^{-1}\\ &= (X_{11}^T)^{-1}(X_{11}^T\{C_{11}\tilde{X}_{11}' - \lambda W_{11}X_{11}'\\ &\quad - W_{11}X_{11} + D_{11}\tilde{U}_{11}'\} - U_{11}^T X_{11}')X_{11}^{-1}\\ &= -W_{11} + (X_{11}^T)^{-1}[*]X_{11}^{-1},\end{aligned}$$

where, using (C11), (C8) and (A1) for the matrix in (C8),

$$\begin{aligned}
[*] &= \{\tilde{X}_{11}^T A_{11}^T + \tilde{U}_{11}^T B_{11}^T\}\{C_{11}\tilde{X}'_{11} + D_{11}\tilde{U}'_{11}\} \\
&\quad - \lambda X_{11}^T W_{11} X'_{11} - \{\tilde{X}_{11}^T C_{11}^T + \tilde{U}_{11}^T D_{11}^T\}\{A_{11}\tilde{X}'_{11} + B_{11}\tilde{U}'_{11}\} \\
&\quad + \lambda X_{11}^T W_{11} X'_{11} \\
&= \tilde{X}_{11}^T (A_{11}^T C_{11} - C_{11}^T A_{11})\tilde{X}'_{11} \\
&\quad + \tilde{X}_{11}^T (A_{11}^T D_{11} - C_{11}^T B_{11})\tilde{U}'_{11} + \tilde{U}_{11}^T (B_{11}^T D_{11} - D_{11}^T B_{11})\tilde{U}'_{11} \\
&\quad + \tilde{U}_{11}^T (B_{11}^T C_{11} - D_{11}^T A_{11})\tilde{X}'_{11} \\
&= \tilde{X}_{11}^T \tilde{U}'_{11} - \tilde{U}_{11}^T \tilde{X}'_{11} \\
&= \begin{pmatrix} X_{11}^T & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} U'_{11} & 0 \\ * & * \end{pmatrix} - \begin{pmatrix} U_{11}^T & * \\ 0 & * \end{pmatrix}\begin{pmatrix} X'_{11} & 0 \\ 0 & 0 \end{pmatrix} \\
&= \tilde{X}_{11}^T \begin{pmatrix} U'_{11}X_{11}^{-1} - U_{11}X_{11}^{-1}X'_{11}X_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}\tilde{X}_{11} \\
&= \tilde{X}_{11}^T \begin{pmatrix} Q'_{11}(k,\lambda) & 0 \\ 0 & 0 \end{pmatrix}\tilde{X}_{11} \le 0
\end{aligned}$$

by induction on k because $(d/d\lambda)Q_{11}(0,\lambda) \equiv 0$. Hence, (C12) holds. □

(C13) $D_k(\lambda) := X_k(\lambda)X_{k+1}^\dagger(\lambda)\mathcal{B}_k = \mathcal{Q}_k^T \left(\begin{smallmatrix} \tilde{D}_k(\lambda) & 0 \\ 0 & 0 \end{smallmatrix}\right) \mathcal{Q}_k$ with

$$\tilde{D}_k(\lambda) := \tilde{X}_{11}(k,\lambda)X_{11}^{-1}(k+1,\lambda)B_{11}(k),$$

$\operatorname{ind} D_k(\lambda) = \operatorname{ind} \tilde{D}_k(\lambda)$ and

$\tilde{D}_k(\lambda) = B_{11}^T(k)(D_{11}(k) - \lambda W_{11}(k)B_{11}(k)) - B_{11}^T(k)Q_{11}(k+1,\lambda)B_{11}(k)$

for all $\lambda \in \mathbf{R} \setminus \mathcal{N}$.

Proof. It follows from easy properties of the Moore-Penrose inverse as, for example, the behavior under orthogonal transformations (note that $\mathcal{P}$ and $\mathcal{Q}_k$ are orthogonal matrices), and from (C11) and (C7) that

$$\begin{aligned}
D_k(\lambda) &= \mathcal{Q}_k^T \tilde{X}_k(\lambda)\mathcal{P}^T(\mathcal{Q}_{k+1}^T \tilde{X}_{k+1}(\lambda)\mathcal{P}^T)^\dagger \mathcal{Q}_{k+1}^T \tilde{\mathcal{B}}_k \mathcal{Q}_k \\
&= \mathcal{Q}_k^T \tilde{X}_k(\lambda)\mathcal{P}^T\mathcal{P}\tilde{X}_{k+1}^\dagger(\lambda)\mathcal{Q}_{k+1}\mathcal{Q}_{k+1}^T \tilde{\mathcal{B}}_k \mathcal{Q}_k \\
&= \mathcal{Q}_k^T \begin{pmatrix} \tilde{X}_{11}(k,\lambda) & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} X_{11}^{-1}(k+1,\lambda)B_{11}(k) & 0 \\ 0 & 0 \end{pmatrix}\mathcal{Q}_k \\
&= \mathcal{Q}_k^T \begin{pmatrix} \tilde{D}_k(\lambda) & 0 \\ 0 & 0 \end{pmatrix}\mathcal{Q}_k.
\end{aligned}$$

Hence, $\operatorname{ind} D_k(\lambda) = \operatorname{ind} \tilde{D}_k(\lambda)$. Moreover, the last formula for $\tilde{D}_k(\lambda)$ follows from (3), (C8) and (C11), because $X_{11}(k+1,\lambda)$ is invertible for all $\lambda \in \mathbf{R} \setminus \mathcal{N}$. □

A similar calculation as in the proof above shows that

$$Q_k(\lambda) := X_k(\lambda)X_k^\dagger(\lambda)U_k(\lambda)X_k^\dagger(\lambda) = \mathcal{Q}_k^T \begin{pmatrix} Q_{11}(k,\lambda) & 0 \\ 0 & 0 \end{pmatrix} \mathcal{Q}_k$$

for all $\lambda \in \mathbf{R} \setminus \mathcal{N}$, see (2) and (C12).

Altogether, the above conclusions lead to the following *local* result.

Theorem 2 (Local oscillation theorem). *Assume* (A1)*, let* $N \in \mathbf{N}$ *and suppose that* $(X,U) = (X_k(\lambda), U_k(\lambda))_{k \in \mathbf{Z}}$ *is a conjoined basis of* (1) *such that* $X_0(\lambda) \equiv X_0$ *and* $U_0(\lambda) \equiv U_0$ *do not depend on* λ*. Moreover assume that* $\lambda_1 \in \mathbf{R}$ *exists such that* (A3) *holds. Then for all* $\lambda_0 \in \mathbf{R}$ *and* $0 \le k \le N$,

$$\operatorname{ind} D_k(\lambda_0+) - \operatorname{ind} D_k(\lambda_0-) = \operatorname{def} X_{k+1}(\lambda_0) - \operatorname{def} X_k(\lambda_0) + r_{k+1} - r_k,$$

where $D_k(\lambda) := X_k(\lambda)X_{k+1}^\dagger(\lambda)\mathcal{B}_k$*, as in* (2) *or* (C13).

Proof. Let $k \in \{0, \dots, N\}$ and $\lambda_0 \in \mathbf{R}$. By (C8) and (A1), the matrix $B_{11}^T(k)D_{11}(k)$ is symmetric and $\operatorname{rank}(B_{11}^T(k), D_{11}^T(k)) = r_{k+1}$. Hence, by [**14**, Corollary 3.1.3], a symmetric matrix $\tilde{S}_1$ and a matrix S_2 exist such that

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

We apply [**14**, Theorem 3.4.1, *Index Theorem*] (cf. also [**13**]) with the same notation. To do so, we put $m = r_{k+1}$, $t = \lambda_0 - \lambda$,

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

By (C1) and (C10), there exists $\varepsilon > 0$ such that $[\lambda_0 - \varepsilon, \lambda_0 + \varepsilon] \setminus \{\lambda_0\} \subset \mathbf{R} \setminus \mathcal{N}$ so that

$$X(t) \text{ is invertible for } t \in [-\varepsilon, \varepsilon] \setminus \{0\}.$$

Moreover, by (C11), $X^T(t)U(t) = U^T(t)X(t)$ for $t \in [-\varepsilon, \varepsilon]$ and, of course, $X(t) \to X$ and $U(t) \to U$ as $t \to 0$. Finally by (C12),

$$S_1(t) + U(t)X^{-1}(t) = \tilde{S}_1 - \lambda W_{11}(k) - U_{11}(k+1, \lambda)X_{11}^{-1}(k+1, \lambda)$$

decreases for $t = \lambda_0 - \lambda \in [-\varepsilon, 0)$ and for $t \in (0, \varepsilon]$. Hence the assumptions of [**14**, Theorem 3.4.1] are satisfied. If we denote

$$M(t) = R_1(t)R_2^T + R_2U(t)X^{-1}(t)R_2^T,$$
$$\Lambda(t) = R_1(t)X(t) + R_2U(t), \quad \Lambda = R_1X + R_2U,$$

then $M(t) = \tilde{D}_k(\lambda_0 - t)$ by (C13), $\Lambda(t) = \tilde{X}_{11}(k, \lambda_0 - t)$ by (C11), (C8) and (3). It follows from (C10) and (C11) that

$$\operatorname{def} \Lambda(0+) = r_{k+1} - r_k \quad \text{and} \quad \operatorname{def} \Lambda = r_{k+1} - r_k + \operatorname{def} X_{11}(k, \lambda_0),$$

i.e., $\operatorname{ind} \tilde{D}_k(\lambda_0-) - \operatorname{ind} \tilde{D}_k(\lambda_0+) = \operatorname{def} X_{11}(k, \lambda_0) - \operatorname{def} X_{11}(k+1, \lambda_0)$, so that by (C13) and (C11),

$$\operatorname{ind} D_k(\lambda_0+) - \operatorname{ind} D_k(\lambda_0-) = \operatorname{def} \tilde{X}_{k+1}(\lambda_0) - n + r_{k+1} - \operatorname{def} \tilde{X}_k(\lambda_0) + n - r_k,$$

which yields our assertion by the definitions of $\tilde{X}_k(\lambda)$ and $\tilde{X}_{k+1}(\lambda)$. □

Remark 4. Note that, by (C1), for $0 \leq k \leq N$,

$$\operatorname{ind} D_k(\lambda_0+) = \operatorname{ind} D_k(\lambda_0-) \quad \text{for all} \quad \lambda_0 \in \mathbf{R} \setminus \mathcal{N}.$$

Proof of Theorem 1. First the assumptions (A1) and (A2) imply via the conclusions (7) of Remark 3 and (C1) that (A3) holds. Since $n_1(\lambda) = n_2(\lambda)$ for sufficiently small λ by (A1), we have to show that (see Remark 4)

$$n_1(\lambda_0+) - n_1(\lambda_0-) = n_2(\lambda_0+) - n_2(\lambda_0-) \quad \text{for all} \quad \lambda_0 \in \mathcal{N},$$

where $\mathcal{N}$ is a finite set by (C1). Since $r_0 = 0$ by Definition 1 (ii), i.e., $X_0 = 0$ and $r_{N+1} = n$ by Remark 3 (i), it follows from (8) and Theorem 2 that

$$\begin{aligned} n_1(\lambda_0+) - n_1(\lambda_0-) &= \sum_{k=0}^{N} \{\operatorname{ind} D_k(\lambda_0+) - \operatorname{ind} D_k(\lambda_0-)\} \\ &= \sum_{k=0}^{N} \{\operatorname{def} X_{k+1}(\lambda_0) - \operatorname{def} X_k(\lambda_0) + r_{k+1} - r_k\} \\ &= \operatorname{def} X_{N+1}(\lambda_0) - \operatorname{def} X_0(\lambda_0) + r_{N+1} - r_0 \\ &= \operatorname{def} X_{N+1}(\lambda_0) - n + n - 0 \\ &= \operatorname{def} X_{N+1}(\lambda_0) \\ &= n_2(\lambda_0+) - n_2(\lambda_0-), \end{aligned}$$

which completes the proof. □

Remark 5. In view of the results in [**16**] (see also [**15**]), in particular because of [**16**, Remark 11, Lemma 12 and Theorem 16], Theorem 1 can be considered as a generalization of an old result of Jacobi [**10**, Section 3].

4. General boundary conditions.

4.1 Separated boundary conditions. Now we consider discrete eigenvalue problems with more general boundary conditions. First we deal with so-called separated boundary conditions. This leads to the following eigenvalue problem (E_s), where $N \in \mathbf{N}$ is a given fixed integer as before.

(E_s)

$$\begin{cases} x_{k+1} = \mathcal{A}_k x_k + \mathcal{B}_k u_k,\, u_{k+1} = \mathcal{C}_k x_k - \lambda \mathcal{W}_k x_{k+1} + \mathcal{D}_k u_k \text{ for } 0 \le k \le N \\ \text{with} \quad R_0^* x_0 + R_0 u_0 = 0, \quad R_{N+1}^* x_{N+1} + R_{N+1} u_{N+1} = 0, \end{cases}$$

where R_0^*, R_0, R_{N+1}^* and R_{N+1} are real $n \times n$ matrices such that

$$\text{(A4)} \qquad \begin{cases} \operatorname{rank}(R_0^*, R_0) = \operatorname{rank}(R_{N+1}^*, R_{N+1}) = n, \\ R_0^* R_0^T = R_0 R_0^{*T}, \quad R_{N+1}^* R_{N+1}^T = R_{N+1} R_{N+1}^{*T} \end{cases}$$

holds. Note that (E_s) is the same as (E) if

$$R_0^* = R_{N+1}^* = I \quad \text{and} \quad R_0 = R_{N+1} = 0.$$

By [**14**, Theorem 3.1.2], a matrix $S_{N+1} \in \mathbf{R}^{n\times n}$ exists such that

$$R_{N+1}^* R_{N+1}^T = R_{N+1} S_{N+1} R_{N+1}^T \quad \text{and} \quad S_{N+1} \text{ is symmetric.} \tag{10}$$

Theorem 3.

Assume (A1), (A4), *and let* $Z = (X, U) = (X_k(\lambda), U_k(\lambda))_{k\in\mathbf{Z}}$ *be the conjoined basis of* (1) *with*

$$X_0 = X_0(\lambda) \equiv -R_0^T \quad \textit{and} \quad U_0 = U_0(\lambda) \equiv R_0^{*T}.$$

Moreover, suppose that

$$\text{(A5)} \qquad \lim_{\lambda\to-\infty} n_1(\lambda) = 0, \qquad \lim_{\lambda\to-\infty} n_2(\lambda) = 0 \quad \textit{and} \quad \lim_{\lambda\to-\infty} n_3(\lambda) = 0$$

holds, where $n_1(\lambda)$ *denotes the number of focal points of* (X, U) *in the interval* $(0, N+1]$, $n_2(\lambda)$ *denotes the number of eigenvalues of* (E_s) *which are less than* λ *and* $n_3(\lambda) = \textit{ind}\, D_{N+1}(\lambda)$ *with* $D_{N+1}(\lambda) := X_{N+1}(\lambda) X_{N+2}^\dagger(\lambda) \mathcal{B}_{N+1}$ *with*

(11)

$$\begin{cases} X_{N+2}(\lambda) := (\mathcal{B}_{N+1} S_{N+1} + I - \mathcal{B}_{N+1} \mathcal{B}_{N+1}^\dagger) X_{N+1}(\lambda) + \mathcal{B}_{N+1} U_{N+1}(\lambda), \\ \mathcal{B}_{N+1} := R_{N+1}^T R_{N+1} \textit{and } S_{N+1} \textit{ as in } (10). \end{cases}$$

Then

$$n_1(\lambda) + n_3(\lambda) = n_2(\lambda) \quad \textit{for all} \quad \lambda \in \mathbf{R} \setminus \mathcal{N}, \tag{12}$$

where the exceptional set

(13)

$$\mathcal{N} := \mathbf{R} \setminus \{\mu \in \mathbf{R} : \operatorname{rank} X_k(\mu) = \max_{\lambda\in\mathbf{R}} \operatorname{rank} X_k(\lambda) \textit{ for } 0 \le k \le N+1 \\ \textit{and } \det \Lambda(\mu) \neq 0\} \textit{ is finite,}$$

with

$$\Lambda(\lambda) := R_{N+1}^* X_{N+1}(\lambda) + R_{N+1} U_{N+1}(\lambda). \tag{14}$$

Proof. Since X_0 and U_0 satisfy the first boundary condition of (E_s), i.e., $R_0^* X_0 + R_0 U_0 = 0$, it follows easily that λ is an eigenvalue of (E_s) if and only if

$$\det \Lambda(\lambda) = 0,$$

and then $\operatorname{def} \Lambda(\lambda)$ is its multiplicity.

It follows from [**14**, Corollary 3.1.3] that $R_{N+1}^* d_1 + R_{N+1} d_2 = 0$ if and only if $d_1 \in \operatorname{Im} R_{N+1}^T$ and $d_2 + S_{N+1} d_1 \in \operatorname{Ker} R_{N+1}$ if and only if $\tilde{R}_1 d_1 + \tilde{R}_2 d_2 = 0$ for $\tilde{R}_1 := \mathcal{B}_{N+1} S_{N+1} + I - \mathcal{B}_{N+1} \mathcal{B}_{N+1}^\dagger$, $\tilde{R}_2 := \mathcal{B}_{N+1}$, because $\operatorname{Ker} \tilde{R}_2 = \operatorname{Ker} R_{N+1}$, $\operatorname{Im} \tilde{R}_2^T = \operatorname{Im} R_{N+1}^T$ and $\operatorname{rank}(\tilde{R}_1, \tilde{R}_2) = n$, $\tilde{R}_1 \tilde{R}_2^T = \mathcal{B}_{N+1} S_{N+1} \mathcal{B}_{N+1}^T$ is symmetric. Hence $(\tilde{R}_1, \tilde{R}_2) = C(R_1, R_2)$ where C is an invertible matrix, so that for $\lambda \in \mathbf{R}$,

$$X_{N+1}(\lambda) = C \Lambda(\lambda) \quad \text{with an invertible matrix} \quad C \in \mathbf{R}^{n \times n}. \tag{15}$$

Now we *construct* an equivalent eigenvalue problem $(\tilde{E})$ to which Theorem 1 applies (cf. [**7**]):

$(\tilde{E})$

$$\begin{pmatrix} x_{k+1} \\ u_{k+1} \end{pmatrix} = \tilde{\mathcal{S}}_k(\lambda) \begin{pmatrix} x_k \\ u_k \end{pmatrix} \quad \text{for } -1 \le k \le N+1 \text{ with } x_{-1} = x_{N+2} = 0.$$

We define for all $\lambda \in \mathbf{R}$,

$$\begin{aligned}
\tilde{\mathcal{S}}_k(\lambda) &:= \mathcal{S}_k(\lambda) \text{ for } 0 \le k \le N-1 \text{ (see Section 2)}, \\
\tilde{\mathcal{S}}_{-1}(\lambda) &:= \begin{pmatrix} R_0^{*T} K & -R_0^T \\ R_0^T K & R_0^{*T} \end{pmatrix} \quad \text{with} \quad K := (R_0^* R_0^{*T} + R_0 R_0^T)^{-1}, \\
\tilde{\mathcal{S}}_N(\lambda) &:= \mathcal{S}_N(\lambda) + \begin{pmatrix} 0 & 0 \\ (S_{N+1} - \mathcal{B}_{N+1}^\dagger) \mathcal{A}_N & (S_{N+1} - \mathcal{B}_{N+1}^\dagger) \mathcal{B}_N \end{pmatrix}, \\
\tilde{\mathcal{S}}_{N+1}(\lambda) &:= \begin{pmatrix} I & \mathcal{B}_{N+1} \\ 0 & I \end{pmatrix}.
\end{aligned}$$

By (A4) and (A1), K exists and the matrices $\tilde{\mathcal{S}}_k(\lambda)$ are symplectic for $-1 \le k \le N+2$ and $\lambda \in \mathbf{R}$. Let $\tilde{Z} = (\tilde{X}_k, \tilde{U}_k)_{k=-1}^{N+2}$ denote the principal solution at -1 of this symplectic difference system, i.e., $\tilde{X}_{-1} = 0$ and $\tilde{U}_{-1} = I$. Then $\tilde{X}_0 = -R_0^T = X_0$ and $\tilde{U}_0 = R_0^{*T} = U_0$, so that $\tilde{X}_k = X_k$ and $\tilde{U}_k = U_k$ for $0 \le k \le N$. Moreover,

$$\tilde{X}_{N+1} = X_{N+1} \quad \text{and} \quad \tilde{U}_{N+1} = U_{N+1} + (S_{N+1} - \mathcal{B}_{N+1}^{\dagger})X_{N+1},$$

and $\tilde{X}_{N+2} = X_{N+2}$ as defined in (11). If $\tilde{n}_1(\lambda)$ and $\tilde{n}_2(\lambda)$ are defined according to Theorem 1 for the eigenvalue problem $(\tilde{E})$, then $\tilde{n}_1(\lambda) = \tilde{n}_2(\lambda)$ for all $\lambda \in \mathbf{R} \setminus \mathcal{N}$ by Theorem 1 using (15), because the definitions of $\mathcal{N}$ by (6) and (13) coincide. Moreover, again by (11), Definition 1 (iv) and our notation,

$$n_2(\lambda) = \tilde{n}_2(\lambda) \text{ and } n_1(\lambda) + n_3(\lambda) = \tilde{n}_1(\lambda) \quad \text{for} \quad \lambda \in \mathbf{R} \setminus \mathcal{N},$$

which completes the proof. $\square$

Of course, Remark 3 (i) and (ii) apply here accordingly. We summarize the conclusions. The assumption $\lim_{\lambda \to -\infty} n_2(\lambda) = 0$ means that $\det \Lambda(\lambda) \not\equiv 0$, and $\lim_{\lambda \to -\infty}(n_1(\lambda) + n_3(\lambda)) = 0$ means that (X, U) has no focal points in the interval $(0, N+2]$, where $\Lambda(\lambda)$ and $X_{N+2}(\lambda)$ are defined by (14) and (11). Using the above construction, the last assertion is equivalent with the positivity of a corresponding quadratic form via the Reid roundabout theorem [**6**, Theorem 1].

4.2 The general case. For $N \in \mathbf{N}$ we consider the following discrete eigenvalue problem

(E_g)

$$\begin{cases} x_{k+1} = \mathcal{A}_k x_k + \mathcal{B}_k u_k, u_{k+1} = \mathcal{C}_k x_k - \lambda \mathcal{W}_k x_{k+1} + \mathcal{D}_k u_k \text{ for } 0 \le k \le N \\ \text{with the boundary conditions } R_1 \begin{pmatrix} -x_0 \\ x_{N+1} \end{pmatrix} + R_2 \begin{pmatrix} u_0 \\ u_{N+1} \end{pmatrix} = 0, \end{cases}$$

where R_1 and R_2 are real $2n \times 2n$ matrices such that

$$\text{(A6)} \qquad \operatorname{rank}(R_1, R_2) = 2n \quad \text{and} \quad R_1 R_2^T = R_2 R_1^T$$

holds. As in Section 4.1, a matrix $S_1 \in \mathbf{R}^{2n \times 2n}$ exists such that

$$\text{(16)} \qquad R_1 R_2^T = R_2 S_1 R_2^T \quad \text{and} \quad S_1 \text{ is symmetric.}$$

Theorem 4. *Assume* (A1), (A6), *and let* $Z = (X, U)$, $\tilde{Z} = (\tilde{X}, \tilde{U})$ *be the principal solution and the associated solution at* 0 *of* (1) *according to Definition* 1 (ii). *Moreover, suppose that* (A5) *of Theorem* 3 *holds. Then the assertion* (12) *of Theorem* 3 *holds, where* $\mathcal{N}$, $n_1(\lambda)$, $n_2(\lambda)$, $n_3(\lambda)$, *via the principal solution, are given as in Theorem* 3, *but we consider the above eigenvalue problem* (E_g) *instead of* (E_s), *and we define* $\Lambda(\lambda)$ *and* $D_{N+1}(\lambda)$ *instead of* (14) *and* (11) *by*

$$(17)\quad \begin{cases} \Lambda(\lambda) := R_1 \bar{X}_{N+1}(\lambda) + R_2 \bar{U}_{N+1}(\lambda) \in \mathbf{R}^{2n\times 2n} \quad \textit{with} \\ \bar{X}_k(\lambda) := \begin{pmatrix} 0 & I \\ X_k(\lambda) & \tilde{X}_k(\lambda) \end{pmatrix}, \bar{U}_k(\lambda) := \begin{pmatrix} I & 0 \\ U_k(\lambda) & \tilde{U}_k(\lambda) \end{pmatrix} \\ \qquad\qquad \textit{for} \quad 0 \le k \le N+1, \\ D_{N+1}(\lambda) := \bar{X}_{N+1}(\lambda) \bar{X}^{\dagger}_{N+2}(\lambda) \mathcal{B} \quad \textit{and} \\ \bar{X}_{N+2}(\lambda) := (\mathcal{B} S_1 + I - \mathcal{B}\mathcal{B}^{\dagger}) \bar{X}_{N+1}(\lambda) + \mathcal{B} U_{N+1}(\lambda), \\ \mathcal{B} := R_2^T R_2 \quad \textit{and} \quad S_1 \quad \textit{as} \quad \textit{in} \quad (16). \end{cases}$$

Proof. We introduce a "big" eigenvalue problem $(\bar{E}_s)$ of size $4n \times 4n$ with separated boundary conditions, which is equivalent with (E_g)

$$(\bar{E}_s)\quad \begin{cases} x_{k+1} = \bar{\mathcal{A}}_k x_k + \bar{\mathcal{B}}_k u_k, u_{k+1} = \bar{\mathcal{C}}_k x_k - \lambda \bar{\mathcal{W}}_k x_{k+1} + \bar{\mathcal{D}}_k u_k \\ \qquad\qquad \text{for } 0 \le k \le N, \\ \text{with } R_0^* x_0 + R_0 u_0 = 0,\ R_{N+1}^* x_{N+1} + R_{N+1} u_{N+1} = 0, \end{cases}$$

where the $2n \times 2n$ matrices occurring are defined as follows:

$$\bar{\mathcal{A}}_k = \begin{pmatrix} I & 0 \\ 0 & \mathcal{A}_k \end{pmatrix}, \quad \bar{\mathcal{B}}_k = \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{B}_k \end{pmatrix}, \quad \bar{\mathcal{C}}_k = \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{C}_k \end{pmatrix},$$
$$\bar{\mathcal{D}}_k = \begin{pmatrix} I & 0 \\ 0 & \mathcal{D}_k \end{pmatrix}, \quad \bar{\mathcal{W}}_k = \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{W}_k \end{pmatrix},$$
$$R_0^* = \begin{pmatrix} I & I \\ 0 & 0 \end{pmatrix}, \quad R_0 = \begin{pmatrix} 0 & 0 \\ -I & I \end{pmatrix}, \quad R_{N+1}^* = R_1, \quad R_{N+1} = R_2.$$

A simple calculation shows that (A1) holds for the big difference system correspondingly. The definitions of $R_0^*, R_0, R_{N+1}^*, R_{N+1}$ and (A6) imply that (A4) holds. Moreover, by (17), $\bar{X}_0(\lambda) \equiv -R_0^T$, $\bar{U}_0(\lambda) \equiv R_0^{*T}$ and $(\bar{X}_k(\lambda), \bar{U}_k(\lambda))_{k=0}^{N+1}$ satisfies the big difference system of our eigenvalue problem $(\bar{E}_s)$. Hence the assumptions of Theorem 3 are satisfied for $(\bar{E}_s)$. Since $\operatorname{rank} X_k(\lambda) = \operatorname{rank} \bar{X}_k(\lambda)$ and

$\operatorname{Ker} X_{k+1}(\lambda) \subset \operatorname{Ker} X_k(\lambda)$ if and only if $\operatorname{Ker} \bar{X}_{k+1}(\lambda) \subset \operatorname{Ker} \bar{X}_k(\lambda)$, we obtain Theorem 4 directly from Theorem 3, provided we prove that

$$\text{(18)} \qquad \operatorname{ind} D_k(\lambda) = \operatorname{ind} \bar{D}_k(\lambda) \quad \text{for all} \quad 0 \le k \le N, \quad \lambda \in \mathbf{R} \setminus \mathcal{N},$$

where $D_k(\lambda) = X_k(\lambda) X_{k+1}^{\dagger}(\lambda) \mathcal{B}_k$, $\bar{D}_k(\lambda) = \bar{X}_k(\lambda) \bar{X}_{k+1}^{\dagger}(\lambda) \bar{\mathcal{B}}_k$ are defined as usual. We show that

$$\text{(19)} \qquad \bar{D}_k(\lambda) = \begin{pmatrix} 0 & 0 \\ 0 & D_k(\lambda) \end{pmatrix} \quad \text{for all} \quad 0 \le k \le N, \quad \lambda \in \mathbf{R} \setminus \mathcal{N}$$

holds, which implies (18).

For the proof of (19), let $0 \le k \le N$, $\lambda \in \mathbf{R} \setminus \mathcal{N}$, and put $\bar{X}_k^{\dagger}(\lambda) = \begin{pmatrix} * & P \\ * & Q \end{pmatrix}$. Then, by our notation,

$$\begin{aligned} \bar{D}_k(\lambda) &= \begin{pmatrix} 0 & I \\ X_k(\lambda) & \tilde{X}_k(\lambda) \end{pmatrix} \begin{pmatrix} 0 & P\mathcal{B}_k \\ 0 & Q\mathcal{B}_k \end{pmatrix} \\ &= \begin{pmatrix} 0 & Q\mathcal{B}_k \\ 0 & X_k P \mathcal{B}_k + \tilde{X}_k Q \mathcal{B}_k \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & X_k P \mathcal{B}_k \end{pmatrix}, \end{aligned}$$

since $Q\mathcal{B}_k = 0$ by the symmetry of $\bar{D}_k(\lambda)$ (cf. Remark 1). Since $\operatorname{Ker} X_{k+1}(\lambda) \subset \operatorname{Ker} X_k(\lambda)$ (see (C2)) we have the following formulas (where we omit the argument λ): $\mathcal{B}_k = X_{k+1} X_{k+1}^{\dagger} \mathcal{B}_k$ (see Remark 1 (i)) $X_k = X_k X_{k+1}^{\dagger} X_{k+1}$ (see [**6**, Remark 1 (v)] or [**4**, Lemma A5]) and $P = (S^{-1/2} X_{k+1})^{\dagger} S^{-1/2}$ with $S = I + \tilde{X} \tilde{X}^T$ (see [**4**, Lemma A6] or [**5**, Remark 8]). Using these identities and the basic property of Moore-Penrose inverses $X = X X^{\dagger} X$, we obtain that

$$\begin{aligned} X_k P \mathcal{B}_k &= X_k X_{k+1}^{\dagger} X_{k+1} ((S^{-1/2} X_{k+1})^{\dagger} S^{-1/2}) X_{k+1} X_{k+1}^{\dagger} \mathcal{B}_k \\ &= X_k X_{k+1}^{\dagger} S^{1/2} (S^{-1/2} X_{k+1}) (S^{-1/2} X_{k+1})^{\dagger} (S^{-1/2} X_{k+1}) X_{k+1}^{\dagger} \mathcal{B}_k \\ &= X_k X_{k+1}^{\dagger} S^{1/2} S^{-1/2} X_{k+1} X_{k+1}^{\dagger} \mathcal{B}_k \\ &= X_k X_{k+1}^{\dagger} \mathcal{B}_k \\ &= X_k X_{k+1}^{\dagger} \mathcal{B}_k \\ &= D_k(\lambda), \end{aligned}$$

which completes the proof. □

REFERENCES

1. R.P. Agarwal, *Difference equations and inequalities*, Dekker, New York, 1992.

2. C.D. Ahlbrandt and A.C. Peterson, *Discrete Hamiltonian systems*: *Difference equations, continued fractions, and Riccati equations*, Kluwer, Boston, 1996.

3. A. Ben-Israel and T.N.E. Greville, *Generalized inverses*: *Theory and applications*, John Wiley & Sons, Inc., New York, 1974.

4. M. Bohner, *Zur Positivität diskreter quadratischer Funktionale*, Ph.D. Thesis, Universität Ulm, 1995; Engl. Edition: *On positivity of discrete quadratic functionals*.

5. ———, *Linear Hamiltonian difference systems*: *Disconjugacy and Jacobi-type conditions*, J. Math. Anal. Appl. **199** (1996), 804–826.

6. M. Bohner and O. Došlý, *Disconjugacy and transformations for symplectic systems*, Rocky Mountain J. Math. **27** (3) (1997), 707–743.

7. M. Bohner, O. Došlý and W. Kratz, *Discrete Reid roundabout theorems*, Dynam. Systems Appl., Special Issue on "Discrete and Continuous Hamiltonian systems" (R.P. Agarwal and M. Bohner, eds.) **8** (1999), 345–352.

8. M. Bohner and A. Peterson, *Dynamic equations on time scales*: *An introduction with applications*, Birkhäuser, Boston, 2001.

9. S. Elaydi, *An introduction to difference equations*, Undergraduate Texts in Math., Springer-Verlag, New York, 1996.

10. F.R. Gantmacher, *The theory of matrices*, Vol. 1, Chelsea Publ. Co., New York, 1959. (Originally published in Moscow, 1954).

11. S. Hilger, *Analysis on measure chains–a unified approach to continuous and discrete calculus*, Results Math. **18** (1990), 18–56.

12. R.A. Horn and C.R. Johnson, *Matrix analysis*, Cambridge Univ. Press, Cambridge, 1991.

13. W. Kratz, *An index theorem for monotone matrix-valued functions*, SIAM J. Matrix Anal. Appl. **16**(1), (1995).

14. ———, *Quadratic functionals in variational analysis and control theory*, Akademie Verlag, Berlin, 1995.

15. ———, *Sturm-Liouville difference equations and banded matrices*, Arch. Math. (Brno) **36** (2000), 499–505.

16. ———, *Banded matrices and difference equations*, Linear Algebra Appl. **337**, (2001), 1–20.

University of Missouri-Rolla, Department of Mathematics and Statistics, 115 Rolla Building, Rolla, Missouri 65409-0020
E-mail address: bohner@umr.edu

Masaryk University Brno, Department of Mathematics, Faculty of Science, Janáčkovo Nám. 2a, CZ-66295 Brno, Czech Republic
E-mail address: dosly@math.muni.cz

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