ARCHIVUM MATHEMATICUM (BRNO)
Tomus 36 (2000), 499–505, CDDE 2000 issue

STURM-LIOUVILLE DIFFERENCE EQUATIONS AND BANDED MATRICES

Werner Kratz

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

Abstract. In this paper we consider *discrete* Sturm-Liouville eigenvalue problems of the form

$$L(y)_k := \sum_{\mu=0}^{n} (-\Delta)^\mu \{ r_\mu(k) \Delta^\mu y_{k+1-\mu} \} = \lambda \rho(k) y_{k+1}$$

$$\textit{for } \ 0 \le k \le N-n \ \textit{ with } \ y_{1-n} = \cdots = y_0 = y_{N+2-n} = \cdots = y_{N+1} = 0,$$

where N and n are integers with $1 \le n \le N$ and with the assumptions that $r_n(k) \neq 0$, $\rho(k) > 0$ for all k. These problems correspond to eigenvalue problems for symmetric, banded matrices $\mathcal{A} \in \mathbb{R}^{(N+1-n)\times(N+1-n)}$ with band-width $2n+1$. We present the following results: - a formula for the chracteristic polynomial of $\mathcal{A}$, which yields a *recursion* for its calculation - an *oscillation theorem*, which generalizes Sturm's well-known theorem on Sturmian chains, and - an inversion formula, which shows that *every* symmetric, banded matrix corresponds uniquely to a Sturm-Liouville eigenvalue problem of the above form.

AMS Subject Classification. 39A10, 39A12, 65F15, 15A18

Keywords. Sturm-Liouville equations, banded matrices, eigenvalue problems; Sturmian chains.

1. Introduction

We consider *discrete* Sturm-Liouville eigenvalue problems (with eigenvalue parameter λ) of the form

$$L(y)_k := \sum_{\mu=0}^{n} (-\Delta)^{\mu} \{r_{\mu}(k) \Delta^{\mu} y_{k+1-\mu}\} = \lambda \rho(k) y_{k+1} \tag{1}$$

for $0 \le k \le N-n$, where $\Delta y_k = y_{k+1} - y_k$, and with the boundary conditions

$$y_{1-n} = \cdots = y_0 = y_{N+2-n} = \cdots = y_{N+1} = 0, \tag{2}$$

where N and n are fixed integers with $1 \le n \le N$ and where we always assume that

$$r_n(k) \neq 0 \quad \textit{for all } k. \tag{3}$$

These problems correspond to eigenvalue problems for symmetric, banded matrices $\mathcal{A}$ of size $(N+1-n) \times (N+1-n)$ with band-width $2n+1$. In particular, $\mathcal{A}$ is *tridiagonal* in the case $n = 1$.

In this paper we essentially formulate and discuss our results while detailed proofs will be given in a forthcoming paper. The following theorems will be presented:

- a formula for the characteristic polynomial of $\mathcal{A}$ (Theorem 1). This result yields also a *recursion* for its calculation. In the case $n = 1$ we obtain the well-known algorithm, which is commonly used in numerical analysis to handle eigenvalue problems for tridiagonal matrices (cf. [[4], pp. 305; [8], pp. 134; [9], pp. 299]).
- an *oscillation theorem* (Theorem 2). This result generalizes Sturm's well-known theorem on Sturmian chains (cf. e.g. [[4], Theorem 8.5-1 or [8], Sätze 4.8 and 4.9]).
- an *inversion formula* (Theorem 3). This identity can be used to calculate the matrix $\mathcal{A}$ when the discrete Sturm-Liouville operator from equation (1) is given and vice versa. Hence, *every* symmetric, banded matrix with bandwidth $2n+1$ corresponds uniquely to such a Sturm-Liouville operator.

Our method and most of our results have continuous counterparts along the lines of the book [6] (cf. also [7]).

2. Discrete Sturm-Liouville equations and associated Hamiltonian systems

In this section we give the connection between discrete Sturm-Liouville equations and Hamiltonian difference systems (cf. [[1], Proposition 5]), and we introduce the important notions of *conjoined bases* and *focal points* of it (cf. [[1], Definitions 1 and 3]). Moreover, the *multiplicity* of focal points is defined according to [3]. It will turn out that these multiplicities always equal one for Hamiltonian systems, which we treat here, i.e. which originate from Sturm-Liouville equations.

Lemma 1. *A vector $y = (y_k)_{1-n}^{N+1} \in \mathbb{R}^{N+1-n}$ solves the Sturm-Liouville difference equation* (1) *for $0 \le k \le N-n$ if and only if (x,u) solves the Hamiltonian difference system*

$$\Delta x_k = A x_{k+1} + B_k u_k\,,\ \Delta u_k = (C_k - \lambda \tilde{C}_k) x_{k+1} - A^T u_k \tag{4}$$

for $0 \le k \le N$, where we use the following **notation:**
A, B_k, C_k, $\tilde{C}_k$ are $n \times n$-matrices defined by

$$A = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 \end{pmatrix}\,,\ B_k = \frac{1}{r_n(k)} B \quad \textit{with} \ \ B = \mathit{diag}\,(0, \dots, 0, 1)\,,$$

$$C_k = \mathit{diag}\,(r_0(k), \dots, r_{n-1}(k))\,,\ \tilde{C}_k = \rho(k)\tilde{C} \quad \textit{with} \ \ \tilde{C} = \mathit{diag}\,(1, 0, \dots, 0)\,,$$

for $0 \le k \le N$, and $x_k = (x_k^{(\nu)})_{\nu=0}^{n-1}$, $u_k = (u_k^{(\nu)})_{\nu=0}^{n-1} \in \mathbb{R}^n$ are defined by

$$x_k^{(\nu)} = \Delta^\nu y_{k-\nu},\ u_k^{(\nu)} = \sum_{\mu=\nu+1}^{n} (-\Delta)^{\mu-\nu-1}\{r_\mu(k)\Delta^\mu y_{k+1-\mu}\}$$

for $0 \le \nu \le n-1\,, 0 \le k \le N+1$ with suitably chosen $y_{N+2}, \dots, y_{N+n+1}$ (which are used for $u_{N+2-n}, \dots, u_{N+1}$).

Definition 1. Assume that (3) holds.

(i) A pair $(X, U) = (X_k, U_k)_{k=0}^{N+1}$ is called a conjoined basis of (4), if the real $n \times n$-matrices X_k, U_k solve (4) for $0 \le k \le N$, and if

$$X_0^T U_0 = U_0^T X_0 \ \text{ and } \ \operatorname{rank}(X_0^T, U_0^T) = n \quad \text{holds.}$$

(ii) Suppose that (X, U) is a conjoined basis of (4) and let $0 \le k \le N$. We say that X has no focal point in the interval $(k, k+1]$ if

$$\operatorname{Ker} X_{k+1} \subset \operatorname{Ker} X_k \ \text{ and } \ D_k := X_k X_{k+1}^\dagger \tilde{A} B_k \ge 0 \quad \text{holds,}$$

where $\tilde{A} := (I - A)^{-1}$. Moreover, if $\operatorname{Ker} X_{k+1} \subset \operatorname{Ker} X_k$ and $D_k \not\ge 0$, then $\operatorname{ind} D_k$ is called the multiplicity of the focal point of X in the interval $(k, k+1)$.

Remark 1.

(i) For a matrix M we denote by $\operatorname{Ker} M$ the kernel of M, $\operatorname{ind} M$ denotes the index of M, i.e., the number of negative eigenvalues of M, provided M is symmetric (and real), and $M^\dagger$ denotes the Moore-Penrose inverse of M. Moreover, $M \ge 0$ means that M is symmetric (and real) and non-negative definite. Observe that D_k is symmetric, if $\operatorname{Ker} X_{k+1} \subset \operatorname{Ker} X_k$ (cf. [[1], Proposition 1]).

(ii) For our Sturm-Liouville difference equations the multiplicity of focal points, which we defined only in case $\operatorname{Ker} X_{k+1} \subset \operatorname{Ker} X_k$, always equals 1, because $\operatorname{rank} D_k \le \operatorname{rank} B = 1$.

3. Associated quadratic functionals and banded matrices

For $y = (y_k)_{k=1}^{N+1-n} \in \mathbb{R}^{N+1-n}$ we define a quadratic functional $\mathcal{F}$, which corresponds to the Sturm-Liouville operator $L(y)$ from equation (1), by

$$\mathcal{F}(y) := \sum_{k=0}^{N} \sum_{\mu=0}^{n} r_\mu(k)(\Delta^\mu y_{k+1-\mu})^2 \,,$$

where we assume (2), i.e., $y_{1-n} = \cdots = y_0 = y_{N+2-n} = \cdots = y_{N+1} = 0$.

Lemma 2. *The following formulas hold.*

(i) $\mathcal{F}(y) = y^T \mathcal{A} y$, *where* $\mathcal{A} \in \mathbb{R}^{(N+1-n)\times(N+1-n)}$ *is a symmetric, banded matrix with band-width* $2n+1$, *which is defined by*

$$a_{k+1,k+1+t} = (-1)^t \sum_{\mu=t}^{n} \sum_{\nu=t}^{\mu} \binom{\mu}{\nu} \binom{\mu}{\nu - t} r_\mu(k+\nu)$$

for $0 \le t \le n$ *and* $0 \le k \le N-n-t$.

(ii) $(\mathcal{A}y)_{k+1} = L(y)_k$ *for* $0 \le k \le N-n$ *with* $L(y)_k$ *given by* (1).

Observe that $\mathcal{A}$ is a tridiagonal $N \times N$-matrix in the case $n = 1$. In the sequel we use the **notation**:

$\mathcal{A}_{N+1} = \mathcal{A} \in \mathbb{R}^{(N+1-n)\times(N+1-n)}$ is the symmetric, banded matrix as defined in Lemma 2, and $\mathcal{A}_k \in \mathbb{R}^{(k-n)\times(k-n)}$ is defined correspondingly for $n+1 \le k \le N+1$. Moreover, let $\mathcal{A}(\lambda) := \mathcal{A} - \lambda \mathcal{D}$ with
$\mathcal{D} := \operatorname{diag}(\rho(0), \ldots, \rho(N-1))$, and as before, $\mathcal{A}_k(\lambda)$ is defined accordingly.

The following statement follows directly from Lemma 2.

Corollary 1. *The discrete Sturm-Liouville eigenvalue problem* (1) *and* (2) *from Section 1 is equivalent with the algebraic eigenvalue problem (matrix pencil)*

$$\mathcal{A}y = \lambda \mathcal{D} y \quad \text{or} \quad \mathcal{A}(\lambda) y = 0.$$

4. Results

We assume throughout that (X, U) is the so-called *principal solution* of (4), i.e., $X = X_k(\lambda)$, $U_k = U_k(\lambda)$ satisfy (4) with

$$X_0 \equiv 0,\ U_0 \equiv I\,. \tag{5}$$

Moreover, as in the previous sections, $y = (y_k)_{k=1}^{N+1-n} \in \mathbb{R}^{N+1-n}$ satisfies (2), i.e., $y_{1-n} = \cdots = y_0 = y_{N+2-n} = \cdots = y_{N+1} = 0$, and

$$\mathcal{F}(y) = \sum_{k=0}^{N} \sum_{\mu=0}^{n} r_\mu(k)(\Delta^\mu y_{k+1-\mu})^2 \,,\ \ D_k = X_k X_{k+1}^\dagger \tilde{A} B_k \,(= D_k(\lambda)).$$

First, we cite some auxiliary results mainly from [1].

4.1. Auxiliary results

Lemma 3. *The following assertions hold, provided* (3) *and* (5) *are fulfilled.*

(i) $X_0, \dots, X_n$ *are independent of* λ.
(ii) $\det X_k = 0$, $D_k = 0$, $Ker\, X_{k+1} \subset Ker\, X_k$ *for* $k = 0, \dots, n-1$.
(iii) $\det X_n = \{r_n(0) \cdots r_n(n-1)\}^{-1} \neq 0$.
(iv) $\det X_k(\lambda) \neq 0$ *for* $n \leq k \leq N+1$, *if* λ *is sufficiently small, provided* $\rho(k) > 0$ *for* $0 \leq k \leq N-n$.
(v) $D_k(\lambda) = \dfrac{1}{r_n(k)} \dfrac{\det X_k(\lambda)}{\det X_{k+1}(\lambda)} B$, *provided* $\det X_{k+1}(\lambda) \neq 0$, *for* $n \leq k \leq N$.

Proof. The assertions (i) and (iii) are derived in a forthcoming paper. The assertion (ii) is contained in [[1], Proposition 6], and (iv) follows from [[1], Satz 9], because

$$\mathcal{F}(y) - \lambda \sum_{k=0}^{N-n} \rho(k) y_{k+1}^2 > 0 \quad for \quad \lambda \leq \lambda_0,$$

if $y \neq 0$ and $\rho(k) > 0$ for $0 \leq k \leq N-n$. Finally, the assertion (v) is shown in [[2], Lemma 4.1].

Observe that $X_k(\lambda)$, $U_k(\lambda)$ are matrix-polynomials in λ, so that $D_k(\lambda)$ is a rational function of λ as follows from Lemma 3 (v). Hence, if $\rho(k) > 0$ for all k, then $\det X_k(\lambda) \neq 0$ for $n \leq k \leq N+1$ and all $\lambda \in \mathbb{R} \setminus \mathcal{N}$ with a *finite* set $\mathcal{N}$. The next result follows from [[1], Proposition 1] and Lemma 3.

Lemma 4. *(Picone's identity) Suppose* (2), (3), *and* (5), *and assume that* $\det X_k(\lambda) \neq 0$ *for* $n \leq k \leq N+1$. *Then*

$$\mathcal{F}(y) - \lambda \sum_{k=0}^{N-n} \rho(k) y_{k+1}^2 = \sum_{k=n}^{N} z_k^T D_k z_k,$$

where $z_k = u_k - U_k(\lambda) X_k^{-1}(\lambda) x_k$ *with* x_k, u_k *as in Lemma 1.*

The next statement with the notation of Section 3 follows immediately from Lemma 3 and Lemma 4.

Corollary 2. *Under the assumptions of Lemma 4*

$$y^T(\mathcal{A}_{N+1} - \lambda \mathcal{D}) y = \sum_{k=n}^{N} r_n(k) \frac{\det X_{k+1}(\lambda)}{\det X_k(\lambda)} w_{k+1-n}^2,$$

where $w_\nu = y_\nu + \sum_{\mu=\nu+1}^{\nu+n} \alpha_\mu y_\mu$ *with suitable coefficients* $\alpha_\mu = \alpha_\mu(\nu, \lambda)$. *Hence,* $w = Ty$ *with* $T = \begin{pmatrix} 1 & \star & \star \\ \vdots & \ddots & \star \\ 0 & \cdots & 1 \end{pmatrix}$, *so that* $\det T = 1$.

4.2. Main results

First, the Lemmas 3 and 4 with Crollary 2 yield our first result, which states a formula for the characteristic polynomial of $\mathcal{A}$ and its recursive calculation.

Theorem 1. *(Recursion) Assume* (3), (5), *and suppose that*

$$\rho(k) > 0 \quad for \;\; 0 \le k \le N - n \tag{6}$$

holds. Then, with the notation of Section 3,

$$det(\mathcal{A} - \lambda\mathcal{D}) = r_n(0) \cdots r_n(N) det X_{N+1}(\lambda) \tag{7}$$

for all $\lambda \in \mathbb{R}$. *Moreover, by* (4) *and* (5), $X_{N+1}(\lambda)$ *is given by the recursion*

$$X_{k+1} = \tilde{A}(X_k + B_k U_k)\,,\; U_{k+1} = (C_k - \lambda\tilde{C}_k)X_{k+1} + (I - A^T)U_k$$

for all $0 \le k \le N$ *with* $X_0 = 0$, $U_0 = I$.

Proof. By Lemma 3 and Lemma 4 we have that

$$\begin{aligned}\det \mathcal{A}(\lambda) =& r_n(n)\frac{\det X_{n+1}(\lambda)}{\det X_n(\lambda)} \cdots r_n(N)\frac{\det X_{N+1}(\lambda)}{\det X_N(\lambda)} \\ =& r_n(0)\cdots r_n(N)\det X_{N+1}(\lambda).\end{aligned}$$

Next, the general oscillation theorem for Hamiltonian systems from reference [3] implies a corresponding result here.

Theorem 2. *(Oscillation) Under the assumptions of Theorem 1 let* $\lambda \in \mathbb{R}$ *with* $det X_k(\lambda) \neq 0$ *for* $n \le k \le N+1$. *Then, the number of eigenvalues (including multiplicities) of the eigenvalue problem* (1), (2) *from Section 1, which are less than* λ, *equals the number of focal points of* $X(\lambda)$ *in the interval* $(0, N+1]$.

Remark 2. Observe first, that the multiplicity of an eigenvalue λ is given by the rank of the kernel of $X_{N+1}(\lambda)$. Hence, it is an integer in $\{1, \dots, n\}$. Moreover, by Remark 1, the focal points of $X(\lambda)$ are all simple, i.e., of multiplicity one, and their number in $(0, N+1]$ equals the number of the elements of the set

$$\{k : n \le k \le N \;\; with \;\; r_n(k)\frac{\det X_{k+1}(\lambda)}{\det X_k(\lambda)} < 0\}.$$

The next corollary is just another formulation of Theorem 2. It generalizes the well-known theorem of Sturm on "Sturmian chains"(cf. [[4], Theorem 8.5-1 and [8], Sätze 4.8 and 4.9). Moreover, it yields the Poincaré separation theorem for banded matrices (cf. [[5], 4.3.16 Corollary]).

Corollary 3. *Under the assumptions of Theorem 2 and the previous notation define polynomials* $f_k(t)$ *by*

$$f_k(t) := det \mathcal{A}_k(t) \quad for \;\; n+1 \le k \le N+1 \quad and \quad f_n(t) \equiv 1. \tag{8}$$

Then the number of zeros of $f_{N+1}(t)$ *(including multiplicities), which are less than* λ, *equals the number of sign changes of* $\{f_k(\lambda)\}$ *for* $n \le k \le N+1$, *i.e.,* $\{f_k(\lambda)\}$ *is a "Sturmian chain".*

Proof. The assertion follows from Theorem 1 and Theorem 2, because

$$f_k(\lambda) = r_n(0) \cdots r_n(k-1) \det X_k(\lambda)$$

for $n \leq k \leq N+1$, so that

$$\frac{f_{k+1}(\lambda)}{f_k(\lambda)} = r_n(k) \frac{\det X_{k+1}(\lambda)}{\det X_k(\lambda)} .$$

Finally, we have the following inversion formula, where the "easy" part is the assertion (i) of Lemma 2, while the main formula will be proved in detail via generating functions in a forthcoming paper as already mentioned in the introduction.

Theorem 3. *(Inversion) The following inversion formulas hold:*

$$r_\mu(k+\mu) = \tag{9}$$

$$(-1)^\mu \sum_{s=\mu}^{n} \left\{ \binom{s}{\mu} a_{k+1,k+1+s} + \sum_{l=1}^{s-\mu} \frac{s}{l} \binom{\mu+l-1}{l-1} \binom{s-l-1}{s-\mu-l} a_{k+1-l,k+1-l+s} \right\},$$

for $0 \leq \mu \leq n$ and all k, if and only if the $a_{\mu\nu}$ are given by

$$a_{k+1,k+1+t} = (-1)^t \sum_{\mu=t}^{n} \sum_{\nu=t}^{\mu} \binom{\mu}{\nu} \binom{\mu}{\nu-t} r_\mu(k+\nu) \tag{10}$$

for $0 \leq t \leq n$ and all k.

References

1. M. Bohner, *Zur Positivität diskreter quadratischer Funktionale*, Dissertation, Ulm, (1995).
2. M. Bohner and O. Došlý, *Positivity of block tridiagonal matrices*, SIAM J. Matrix Anal. Appl. **20** (1999), 182–195.
3. M. Bohner, O. Došlý and W. Kratz, *An oscillation theorem for discrete eigenvalue problems*, to appear.
4. G.H. Golub and C.F. Van Loan, *Matrix computations*, John Hopkins University Press, Baltimore, (1983).
5. R.A.Horn and C.A. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge, (1991).
6. W. Kratz, *Quadratic functionals in variational analysis and control theory*, Akademie Verlag, Berlin, (1995).
7. W.T. Reid, *Sturmian theory for ordinary differential equations*, Springer Verlag, New York, (1980).
8. H.R. Schwarz, H. Rutishauser and E. Stiefel, *Numerik symmetrischer Matrizen*, Teubner Verlag, Stuttgart, (1972).
9. J.H. Wilkinson, The algebraic eigenvalue problem, Clarendon Press, Oxford, (1965).