Journal of Computational and Applied Mathematics 66 (1996) 27–52

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

Orthogonal matrix polynomials and applications

Ann Sinap, Walter Van Assche*

Department of Mathematics, Katholieke Universiteit Leuven, Celestijnenlaan 200B, B-3001 Heverlee, Belgium

Received 5 October 1994; revised 31 May 1995

Abstract

Orthogonal matrix polynomials, on the real line or on the unit circle, have properties which are natural generalizations of properties of scalar orthogonal polynomials, appropriately modified for the matrix calculus. We show that orthogonal matrix polynomials, both on the real line and on the unit circle, appear at various places and we describe some of them. The spectral theory of doubly infinite Jacobi matrices can be described using orthogonal 2×2 matrix polynomials on the real line. Scalar orthogonal polynomials with a Sobolev inner product containing a finite number of derivatives can be studied using matrix orthogonal polynomials on the real line. Orthogonal matrix polynomials on the unit circle are related to unitary block Hessenberg matrices and are very useful in multivariate time series analysis and multichannel signal processing. Finally we show how orthogonal matrix polynomials can be used for Gaussian quadrature of matrix-valued functions. Parallel algorithms for this purpose have been implemented (using PVM) and some examples are worked out.

Keywords: Orthogonal matrix polynomials; Block tridiagonal matrices; Block Hessenberg matrices; Multichannel time series; Quadrature

AMS classification: 42C05; 47A56; 62M10; 65D32

1. Matrix polynomials

Let $A_0, A_1, \ldots, A_n \in \mathbb{C}^{p \times p}$ be $n+1$ square matrices with complex entries, and suppose that $A_n \neq 0$. Then $P : \mathbb{C} \to \mathbb{C}^{p \times p}$ such that

$$P(z) = A_n z^n + A_{n-1} z^n + \cdots + A_1 z + A_0$$

is a matrix polynomial of degree n. If the leading coefficient is the identity, i.e., $A_n = I$, then P is a monic matrix polynomial. Obviously, when the leading coefficient A_n is nonsingular, then $A_n^{-1} P$ is a monic matrix polynomial. A matrix polynomial P is thus a polynomial in a complex variable

* Corresponding author. E-mail: walter@wis.kuleuven.ac.be. (Senior Research Associate of the Belgian National Fund for Scientific Research.)

SSDI 0377-0427(95)00193-X

with matrix coefficients and $P(z)$ is a $p \times p$ matrix for which each entry is a polynomial in z of degree at most n.

One can usually work with matrix polynomials just as with scalar polynomials. One needs to be careful though because of noncommutativity of the matrix product. In particular, in general one has $P(z)Q(z) \neq Q(z)P(z)$.

For scalar polynomials (the case $p = 1$) it is well known that a polynomial of degree n has n complex zeros (counting multiplicities). Furthermore, a polynomial is completely determined by its leading coefficient and its zeros. Something similar is true for matrix polynomials, but first one needs to give a meaning to the notion of a zero of a matrix polynomial P. A zero of a matrix polynomial (with $p > 1$) is *not* a complex number z_0 such that $P(z_0) = 0$, since this would require a simultaneous solution of p^2 polynomial equations. A complex number $z_0 \in \mathbb{C}$ is a zero of P if $P(z_0)$ is a singular matrix. This means that the zeros of a matrix polynomial P are equal to the zeros of the scalar polynomial $\det P(z)$. If the leading coefficient A_n is nonsingular, then $\det P(z)$ is a (scalar) polynomial of degree np, and thus a matrix polynomial P of degree n and $\det A_n \neq 0$ has np zeros. If z_0 is a zero of P, then $P(z_0)$ has an eigenvalue 0 and eigenvector $v_0 \in \mathbb{C}^p$. In fact, for the first-degree polynomial $P_1(z) = Iz - A$, the zeros are precisely the eigenvalues of A and $\det P_1(z)$ is the characteristic polynomial of A. For a full description of A one needs the Jordan form, which is related to the eigenvalues, their multiplicity as a zero of the characteristic polynomial, and the geometric multiplicity (dimension of the eigenspace spanned by the eigenvectors). Zeros of a matrix polynomial are thus generalizations of eigenvalues of a matrix (which occur for polynomials of degree 1), and thus it should not be surprising that for a description of a matrix polynomial P one also needs the Jordan normal form.

A (right) *Jordan chain* of length $k + 1$ for a zero z_0 of a matrix polynomial P is a sequence of p-vectors $v_0, v_1, \ldots, v_k \in \mathbb{C}^p$ with $v_0 \neq 0$ and

$$\sum_{i=0}^{j} \frac{1}{i!} P^{(i)}(z_0) v_{j-i} = 0, \quad j = 0, 1, \ldots, k.$$

The first vector v_0 is called a *root vector*. We will not use the terminology 'eigenvector' to distinguish zeros of matrix polynomials with eigenvalues of matrices. If for each $i = 1, 2, \ldots, s$ the sequence $(v_{i,0}, \ldots, v_{i,\kappa_i - 1})$ is a Jordan chain of length κ_i for the zero z_0 of the matrix polynomial P, then the set $\{(v_{i,0}, \ldots, v_{i,\kappa_i - 1})\colon i = 1, \ldots, s\}$ is a *canonical* set of Jordan chains if the root vectors $v_{1,0}, \ldots, v_{s,0}$ are linearly independent and if $\sum_{i=1}^{s} \kappa_i$ is the multiplicity of z_0 as a zero of the scalar polynomial $\det P$. Canonical sets of Jordan chains are not unique, but the number s of Jordan chains in a canonical set and the lengths $\kappa_1, \ldots, \kappa_s$ of these Jordan chains only depend on the matrix polynomial P and the zero z_0 and not on the choice of canonical set. Therefore, a matrix polynomial P can be described by its zeros and a canonical set of Jordan chains. In particular, the root vectors of a canonical set of Jordan chains are a basis for $\operatorname{Ker} P(z_0)$.

For a monic matrix polynomial P we can then define a *Jordan pair* (X, J). If $z_1, \ldots, z_k$ are the zeros of P, then for each zero z_i we choose a canonical set of Jordan chains $\{v_{j,0}^{(i)}, \ldots, v_{j,\kappa_j^{(i)}-1}^{(i)}\colon j = 1, 2, \ldots, s_i\}$ and put these vectors in a matrix

$$X_i = (v_{1,0}^{(i)} \ \cdots \ v_{1,\kappa_1^{(i)}-1}^{(i)} \ v_{2,0}^{(i)} \ \cdots \ v_{2,\kappa_2^{(i)}-1}^{(i)} \ \cdots \ v_{s_i,0}^{(i)} \ \cdots \ v_{s_i,\kappa_{s_i}^{(i)}-1}^{(i)}).$$

If z_i is a zero of multiplicity k_i, then $\sum_{j=1}^{s_i} \kappa_j^{(i)} = k_i$ and thus X_i is a matrix of size $p \times k_i$.

Furthermore, let J_i be the block diagonal matrix

$$J_i = \begin{pmatrix} J_{i,1} & \cdots & \cdots & 0 \\ \cdots & J_{i,2} & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & \cdots & J_{i,s_i} \end{pmatrix},$$

where $J_{i,j}$ is the Jordan block of size $\kappa_j^{(i)} \times \kappa_j^{(i)}$ for the zero z_i and the root vector $v_{j,0}^{(i)}$. The pair (X_i, J_i) is then the Jordan pair of P for the zero z_i. The Jordan pair for the polynomial P finally is the pair (X, J), where

$$X = (X_1 \;\; X_2 \;\; \cdots \;\; X_k), \quad J = \begin{pmatrix} J_1 & \cdots & \cdots & 0 \\ \cdots & J_2 & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & \cdots & J_k \end{pmatrix}.$$

Hence X is a $p \times np$ matrix and J is an $np \times np$ matrix.

More information on matrix polynomials can be found in [11].

2. Orthogonal matrix polynomials

We are particularly interested in orthogonal matrix polynomials and we will restrict our attention to orthogonal matrix polynomials on the real line [9] and on the unit circle [5, 8, 10, 28, 35].

2.1. *Orthogonal matrix polynomials on the real line*

First we need to introduce a matrix inner product. Let $\rho : (-\infty, \infty) \to \mathbb{R}^{p \times p}$ be a bounded symmetric matrix function such that $\rho(x) \leqslant \rho(y)$ when $x < y$. Inequality for square matrices $A \leqslant B$ means that $B - A$ is nonnegative definite. Then ρ is a matrix-valued distribution function (measure) on the real line. Such a matrix-valued measure gives rise to two different inner products: a *left* inner product and a *right* inner product. The left inner product is given by

$$\langle P, Q \rangle_{\mathrm{L}} = \int_{-\infty}^{\infty} P(x)\, \mathrm{d}\rho(x) Q(x)^{\mathrm{t}},$$

where A^{t} denotes the transpose of the matrix A. The right inner product is

$$\langle P, Q \rangle_{\mathrm{R}} = \int_{-\infty}^{\infty} P(x)^{\mathrm{t}}\, \mathrm{d}\rho(x) Q(x).$$

Note that left (right) indicates that in the inner product the left (right) member is taken as it is (without transposition). The left inner product has the following properties, which can be verified easily:

(1) $\langle P, Q\rangle_{\mathrm{L}} = \langle Q, P\rangle_{\mathrm{L}}^{\mathrm{t}}$;
(2) if $C_1, C_2 \in \mathbb{R}^{p\times p}$, then $\langle C_1P_1 + C_2P_2, Q\rangle_{\mathrm{L}} = C_1\langle P_1, Q\rangle_{\mathrm{L}} + C_2\langle P_2, Q\rangle_{\mathrm{L}}$;
(3) $\langle xP, Q\rangle_{\mathrm{L}} = \langle P, xQ\rangle_{\mathrm{L}}$;
(4) $\langle P, P\rangle_{\mathrm{L}}$ is nonnegative definite, and when $\det P \not\equiv 0$ it is positive definite;
(5) $\langle P, P\rangle_{\mathrm{L}} = 0$ if and only if $P = 0$.

Similar properties are valid for the right inner product, but property (2) becomes

(2′) if $C_1, C_2 \in \mathbb{R}^{p\times p}$, then $\langle P, Q_1C_1 + Q_2C_2\rangle_{\mathrm{R}} = \langle P, Q_1\rangle_{\mathrm{R}}C_1 + \langle P, Q_2\rangle_{\mathrm{R}}C_2$.

We can now define orthogonal matrix polynomials. Left orthonormal matrix polynomials P_n^{L} ($n = 0, 1, 2, \ldots$) are obtained by orthonormalizing $I, zI, z^2I, z^3I, \ldots$ using the left inner product. This gives

$$\int_{-\infty}^{\infty} P_n^{\mathrm{L}}(x)\,\mathrm{d}\rho(x)P_m^{\mathrm{L}}(x)^{\mathrm{t}} = \delta_{n,m}I, \quad n, m \geqslant 0.$$

Similarly right orthonormal matrix polynomials P_n^{R} ($n = 0, 1, 2, \ldots$) are obtained using the right inner product, leading to

$$\int_{-\infty}^{\infty} P_n^{\mathrm{R}}(x)^{\mathrm{t}}\,\mathrm{d}\rho(x)P_m^{\mathrm{R}}(x) = \delta_{n,m}I, \quad n, m \geqslant 0.$$

The leading coefficient for both left and right orthonormal polynomials is always nonsingular. Left orthonormal matrix polynomials are unique up to a multiplication on the left by an orthogonal matrix. Indeed, if A_n ($n \geqslant 0$) are orthogonal matrices, i.e., $A_nA_n^{\mathrm{t}} = I = A_n^{\mathrm{t}}A_n$, then

$$\langle A_nP_n^{\mathrm{L}}, A_mP_m^{\mathrm{L}}\rangle_{\mathrm{L}} = A_n\langle P_n^{\mathrm{L}}, P_m^{\mathrm{L}}\rangle_{\mathrm{L}}A_m^{\mathrm{t}} = \begin{cases} 0 & \text{if } n \neq m, \\ A_nA_n^{\mathrm{t}} = I & \text{if } n = m. \end{cases}$$

Similarly right orthonormal matrix polynomials are unique up to a multiplication on the right by an orthogonal matrix. Left and right orthogonal matrix polynomials on the real line are closely related since $P_n(x)^{\mathrm{R}} = [P_n(x)^{\mathrm{L}}]^{\mathrm{t}}$, which follows from $\langle P, Q\rangle_{\mathrm{L}} = \langle P^{\mathrm{t}}, Q^{\mathrm{t}}\rangle_{\mathrm{R}}$. Hence for orthogonal matrix polynomials on the real line it is sufficient to treat only the left orthogonal matrix polynomials.

2.2. *Orthogonal matrix polynomials on the unit circle*

We now introduce a matrix inner product on the unit circle. Let $\rho : [0, 2\pi) \to \mathbb{C}^{p\times p}$ be a bounded Hermitian matrix function such that $\rho(\theta_1) \leqslant \rho(\theta_2)$ when $\theta_1 < \theta_2$. Then ρ is a matrix-valued distribution function (measure) on $[0, 2\pi)$, which gives a matrix-valued measure on the unit circle. The left inner product now becomes

$$\langle P,\ Q\rangle_{\mathrm{L}} = \frac{1}{2\pi}\int_0^{2\pi} P(z)\,\mathrm{d}\rho(\theta)Q(z)^*, \quad z = \mathrm{e}^{\mathrm{i}\theta},$$

where A^* denotes the Hermitian conjugate of the matrix A. The right inner product is given by

$$\langle P, Q\rangle_{\mathrm{R}} = \frac{1}{2\pi}\int_0^{2\pi} P(z)^*\,\mathrm{d}\rho(\theta)Q(z), \quad z = \mathrm{e}^{\mathrm{i}\theta}.$$

Again left and right denotes on which side the matrix function is taken without Hermitian conjugation. Elementary properties of the left inner product are

(1) $\langle P, Q\rangle_{\mathrm{L}} = \langle Q, P\rangle_{\mathrm{L}}^*$;
(2) if $C_1, C_2 \in \mathbb{C}^{p\times p}$, then $\langle C_1P_1 + C_2P_2, Q\rangle_{\mathrm{L}} = C_1\langle P_1, Q\rangle_{\mathrm{L}} + C_2\langle P_2, Q\rangle_{\mathrm{L}}$;
(3) $\langle zP, Q\rangle_{\mathrm{L}} = \langle P, \bar{z}Q\rangle_{\mathrm{L}}$;
(4) $\langle P, P\rangle_{\mathrm{L}}$ is nonnegative definite, and when $\det P \not\equiv 0$ it is positive definite;
(5) $\langle P, P\rangle_{\mathrm{L}} = 0$ if and only if $P = 0$.

Similar properties are valid for the right inner product, with an obvious modification for property (2).

Left orthonormal matrix polynomials φ_n^{L} $(n = 0, 1, 2, \dots)$ are obtained by orthonormalizing the monomials using the left inner product, giving

$$\frac{1}{2\pi}\int_0^{2\pi} \varphi_n^{\mathrm{L}}(z)\,\mathrm{d}\rho(\theta)\varphi_m^{\mathrm{L}}(z)^* = \delta_{n,m}I, \quad z = \mathrm{e}^{\mathrm{i}\theta},\ n, m \geqslant 0.$$

Right orthonormal matrix polynomials φ_n^{R} $(n = 0, 1, 2, \dots)$ are obtained by using the right inner product, giving

$$\frac{1}{2\pi}\int_0^{2\pi} \varphi_n^{\mathrm{R}}(z)^*\,\mathrm{d}\rho(\theta)\varphi_m^{\mathrm{R}}(z) = \delta_{n,m}I, \quad z = \mathrm{e}^{\mathrm{i}\theta},\ n, m \geqslant 0.$$

Again the leading coefficient for both the left and the right orthogonal matrix polynomials is nonsingular. Now left orthonormal matrix polynomials are unique up to a multiplication on the left by a unitary matrix. Indeed, if A_n $(n \geqslant 0)$ are unitary matrices, i.e., $A_nA_n^* = I = A_n^*A_n$, then

$$\langle A_n\varphi_n^{\mathrm{L}}, A_m\varphi_m^{\mathrm{L}}\rangle_{\mathrm{L}} = A_n\langle \varphi_n^{\mathrm{L}}, \varphi_m^{\mathrm{L}}\rangle_{\mathrm{L}}A_m^* = \begin{cases} 0 & \text{if } n \neq m, \\ A_nA_n^* = I & \text{if } n = m. \end{cases}$$

Similarly, right orthonormal matrix polynomials are unique up to multiplication on the right by a unitary matrix.

On the unit circle there is no simple relation between left and right orthonormal polynomials (as was the case for the real line). This is due to property (3) of the left (and right) inner product. Therefore, we need to analyze both left and right orthonormal matrix polynomials when working with a measure ρ on the unit circle.

3. Recurrence relations

Scalar orthogonal polynomials on the real line and on the unit circle satisfy some simple recurrence relations (see, e.g., [4, Ch. IV] and [32, Sections 3.2 and 11.4]). Similar recurrence relations, but now with matrix coefficients, are valid for orthogonal matrix polynomials on the real line and on the unit circle.

3.1. On the real line

Theorem 3.1. *Suppose P_n^L $(n = 0, 1, 2, \ldots)$ are left orthonormal matrix polynomials with respect to the left inner product obtained by using a matrix measure ρ on the real line. Then there exist matrices $D_n \in \mathbb{R}^{p\times p}$ $(n = 1, 2, \ldots)$, with $\det D_n \neq 0$, and $E_n \in \mathbb{R}^{p\times p}$ with $E_n = E_n^t$ $(n = 0, 1, 2, \ldots)$, such that*

$$xP_n^L(x) = D_{n+1}P_{n+1}^L(x) + E_nP_n^L(x) + D_n^tP_{n-1}^L(x). \tag{3.1}$$

Proof. Expand $xP_n^L(x)$ into a Fourier series

$$xP_n^L(x) = \sum_{k=0}^{n+1} A_{k,n}P_k^L(x),$$

then the Fourier coefficients are

$$A_{k,n} = \langle xP_n^L, P_k^L\rangle_L.$$

Orthonormality then shows that $A_{k,n} = 0$ whenever $k < n - 1$, and thus only $A_{n+1,n}$, $A_{n,n}$ and $A_{n-1,n}$ remain. Furthermore properties (1) and (3) of the matrix inner product show that $A_{n+1,n} = A_{n,n+1}^t$. Hence take $A_{n+1,n} = D_{n+1}$ and $A_{n,n} = E_n$. □

A converse result also holds, namely when P_n $(n = 0, 1, 2, \ldots)$ are matrix polynomials satisfying a three-term recurrence relation as in (3.1), then these polynomials are left orthonormal matrix polynomials on the real line for a matrix measure ρ.

3.2. On the unit circle

First we introduce the *reversed polynomial*

$$\tilde{P}(z) = z^nP(1/\bar{z})^*.$$

It takes the polynomial P and reverses the order of the coefficients and also takes the Hermitian conjugates of these coefficients. In the scalar case $p = 1$ the reversed polynomial for $p_n(z)$ is usually denoted by $p_n^*(z) = z^n\bar{p}_n(1/z)$, but for matrix polynomials we use a tilde (˜), since an asterisk (*) is already used for Hermitian conjugation.

Theorem 3.2. *Suppose the left and right orthonormal matrix polynomials on the unit circle are given by*

$$\varphi_n^R(z) = K_{n,n}z^n + K_{n,n-1}z^{n-1} + \cdots + K_{n,1}z + K_{n,0},$$

$$\varphi_n^L(z) = L_{n,n}z^n + L_{n,n-1}z^{n-1} + \cdots + L_{n,1}z + L_{n,0}.$$

Then

$$\varphi_n^R(z)K_{n,n}^* = z\varphi_{n-1}^R(z)K_{n-1,n-1}^* + \tilde{\varphi}_n^L(z)(L_{n,n}^*)^{-1}K_{n,0}K_{n,n}^*, \tag{3.2}$$

$$L_{n,n}^*\varphi_n^L(z) = zL_{n-1,n-1}^*\varphi_{n-1}^L(z) + L_{n,n}^*L_{n,0}(K_{n,n}^*)^{-1}\tilde{\varphi}_n^R(z). \tag{3.3}$$

Proof. Expand

$$P_{n-1}(z) = \frac{\varphi_n^{\mathrm{R}}(z) - \tilde{\varphi}_n^{\mathrm{L}}(z)(L_{n,n}^*)^{-1}K_{n,0}}{z}$$

into a Fourier series with right orthonormal matrix polynomials

$$P_{n-1}(z) = \sum_{k=0}^{n-1} \varphi_k^{\mathrm{R}}(z) B_{k,n},$$

with Fourier coefficients

$$B_{k,n} = \langle \varphi_k^{\mathrm{R}}, P_{n-1} \rangle_{\mathrm{R}} = \langle \varphi_k^{\mathrm{R}}, \bar{z}\varphi_n^{\mathrm{R}} \rangle_{\mathrm{R}} - \langle \varphi_k^{\mathrm{R}}, \bar{z}\tilde{\varphi}_n^{\mathrm{L}}(L_{n,n}^*)^{-1}K_{n,0} \rangle_{\mathrm{R}}.$$

The first term is, using property (3) of the matrix inner product,

$$\langle \varphi_k^{\mathrm{R}}, \bar{z}\varphi_n^{\mathrm{R}} \rangle_{\mathrm{R}} = \langle z\varphi_k^{\mathrm{R}}, \varphi_n^{\mathrm{R}} \rangle_{\mathrm{R}} = 0, \quad k \leqslant n-2,$$

and $\langle \varphi_{n-1}^{\mathrm{R}}, \bar{z}\varphi_n^{\mathrm{R}} \rangle_{\mathrm{R}} = K_{n-1,n-1}^*(K_{n,n}^*)^{-1}$. For the second term we have for every $k < n$

$$\langle \varphi_k^{\mathrm{R}}, \bar{z}\tilde{\varphi}_n^{\mathrm{L}}(z) \rangle_{\mathrm{R}} = \langle \varphi_k^{\mathrm{R}}, z^{n-1}(\varphi_n^{\mathrm{L}})^* \rangle_{\mathrm{R}} = \langle z^{n-1}(\varphi_k^{\mathrm{R}})^*, \varphi_n^{\mathrm{L}} \rangle_{\mathrm{L}} = 0.$$

This proves the result. □

The coupled recurrences (3.2) and (3.3) can be simplified. Observe that $\langle \varphi_n^{\mathrm{R}}, \tilde{\varphi}_n^{\mathrm{L}} \rangle_{\mathrm{R}} = \langle \tilde{\varphi}_n^{\mathrm{R}}, \varphi_n^{\mathrm{L}} \rangle_{\mathrm{L}}$, which implies

$$K_{n,n}^{-1} L_{n,0}^* = K_{n,0}^* L_{n,n}^{-1}. \tag{3.4}$$

We can then introduce *reflection coefficients* as

$$H_n = (L_{n,n}^*)^{-1} K_{n,0} = L_{n,0}(K_{n,n}^*)^{-1}. \tag{3.5}$$

Then some matrix calculus shows that

$$(I - H_n^* H_n)^{1/2} = K_{n,n}^{-1} K_{n-1,n-1} = K_{n-1,n-1}^*(K_{n,n}^*)^{-1},$$

$$(I - H_n H_n^*)^{1/2} = (L_{n,n}^*)^{-1} L_{n-1,n-1}^* = L_{n-1,n-1} L_{n,n}^{-1},$$

and since the leading coefficients $L_{n,n}$ and $K_{n,n}$ ($n \geqslant 0$) are positive definite, it follows that $\| H_n \|_2 < 1$. The recurrence relations can now be written only in terms of the matrices H_n as

$$\varphi_n^{\mathrm{L}}(z) = (I - H_n H_n^*)^{1/2} z \varphi_{n-1}^{\mathrm{L}}(z) + H_n \tilde{\varphi}_n^{\mathrm{R}}(z),$$

$$\varphi_n^{\mathrm{R}}(z) = z\varphi_{n-1}^{\mathrm{R}}(z)(I - H_n^* H_n)^{1/2} + \tilde{\varphi}_n^{\mathrm{L}}(z) H_n.$$

Elimination of $\varphi_n^{\mathrm{R}}(z)$ from both equations then gives

$$(I - H_n H_n^*)^{1/2} \varphi_n^{\mathrm{L}}(z) = z\varphi_{n-1}^{\mathrm{L}}(z) + H_n \tilde{\varphi}_{n-1}^{\mathrm{R}}(z), \tag{3.6}$$

$$(I - H_n^* H_n)^{1/2} \tilde{\varphi}_n^{\mathrm{R}}(z) = \tilde{\varphi}_{n-1}^{\mathrm{R}}(z) + z H_n^* \varphi_{n-1}^{\mathrm{L}}(z). \tag{3.7}$$

These recurrences are matrix generalizations of the Szegő recurrence relations for scalar orthogonal polynomials on the unit circle. They show that the orthonormal matrix polynomials (both left and right) are completely determined by the reflection coefficients H_n ($n = 0, 1, 2, \dots$).

Also now a converse result holds. If polynomials φ_n^{L} and φ_n^{R} are given by coupled recurrences of the form (3.6) and (3.7), with $\|H_n\|_2 < 1$, then these polynomials are left and right orthonormal matrix polynomials on the unit circle with some matrix measure ρ. See, e.g., [5].

4. Zeros of orthogonal matrix polynomials

4.1. On the real line

From the three-term recurrence relation (3.1) we can use the recurrence coefficients and form the block tridiagonal matrix (Jacobi matrix)

$$J_n = \begin{pmatrix} E_0 & D_1 & & & & 0 \\ D_1^{\mathrm{t}} & E_1 & D_2 & & & \\ & D_2^{\mathrm{t}} & E_2 & D_3 & & \\ & & D_3^{\mathrm{t}} & E_3 & \ddots & \\ & & & \ddots & \ddots & D_{n-1} \\ 0 & & & & D_{n-1}^{\mathrm{t}} & E_{n-1} \end{pmatrix}.$$

It turns out that the zeros $x_{n,1} \leqslant x_{n,2} \leqslant \cdots \leqslant x_{n,np}$ of the left orthogonal matrix polynomial P_n^{L} are equal to the eigenvalues of the block Jacobi matrix J_n [6, 29]. This immediately has some useful consequences. Since J_n is real and symmetric, it follows that all the zeros are real. Furthermore, since the blocks are $p \times p$ matrices, the multiplicity of a zero is at most p. Finally there is an interlacing property, which follows from the inclusion principle for symmetric matrices [15, p. 189]. If we delete the last p rows and the last p columns of J_n, then we have the matrix J_{n-1} for which the eigenvalues are the zeros of the left orthogonal matrix P_{n-1}^{L}, and thus

$$x_{n,k} \leqslant x_{n-1,k} \leqslant x_{n,k+p}.$$

4.2. On the unit circle

Expand the left orthogonal matrix polynomial φ_k^{L} into a Fourier series using the left orthogonal matrix polynomials

$$z\varphi_k^{\mathrm{L}}(z) = \sum_{j=0}^{k+1} M_{k,j}\varphi_j^{\mathrm{L}}(z), \quad M_{k,j} \in \mathbb{C}^{p \times p}, \tag{4.1}$$

then $M_{k,j} = \langle z\varphi_k^{\mathrm{L}}, \varphi_j^{\mathrm{L}} \rangle_{\mathrm{L}}$. If z_0 is a zero of φ_n^{L} with root vector v_0, then

$$z_0 \varphi_k^{\mathrm{L}}(z_0) v_0 = \sum_{j=0}^{k+1} M_{k,j} \varphi_j^{\mathrm{L}}(z_0) v_0, \quad k = 0, 1, \dots, n-2,$$

$$z_0 \varphi_{n-1}^{\mathrm{L}}(z_0) v_0 = \sum_{j=0}^{n-1} M_{n-1,j} \varphi_j^{\mathrm{L}}(z_0) v_0,$$

hence z_0 is an eigenvalue of the block Hessenberg matrix

$$M_n = \begin{pmatrix} M_{0,0} & M_{0,1} & & & & 0 \\ M_{1,0} & M_{1,1} & M_{1,2} & & & \\ \vdots & \vdots & \vdots & \ddots & & \\ \vdots & \vdots & \vdots & \ddots & \ddots & \\ M_{n-2,0} & M_{n-2,1} & \cdots & \cdots & M_{n-2,n-2} & M_{n-2,n-1} \\ M_{n-1,0} & M_{n-1,1} & \cdots & \cdots & M_{n-1,n-2} & M_{n-1,n-1} \end{pmatrix}.$$

From property (3) of the left inner product and the orthogonality we have on the one hand

$$\langle z\varphi_k^{\mathrm{L}}, z\varphi_l^{\mathrm{L}} \rangle_{\mathrm{L}} = \langle \varphi_k^{\mathrm{L}}, \varphi_l^{\mathrm{L}} \rangle_{\mathrm{L}} = \delta_{k,l} I,$$

so that

$$\sum_{j=0}^{\min(k,l)+1} M_{k,j} M_{l,j}^* = \langle z\varphi_k^{\mathrm{L}}, z\varphi_l^{\mathrm{L}} \rangle_{\mathrm{L}} = \delta_{k,l} I.$$

Hence

$$M_n M_n^* = \begin{pmatrix} I_{(n-1)p} & 0 \\ 0 & I - M_{n-1,n} M_{n-1,n}^* \end{pmatrix},$$

so that M_n is almost unitary. In fact $I - M_n M_n^* = M_{n-1,n} M_{n-1,n}^*$ is nonnegative definite, which implies that there are no eigenvalues outside the closed unit disk.

The blocks $M_{n,j}$ can be expressed explicitly in terms of the reflection coefficients H_n. Indeed, from (3.2) we find

$$K_{n,n} \tilde{\varphi}_n^{\mathrm{R}}(z) = K_{n-1,n-1} \tilde{\varphi}_{n-1}^{\mathrm{R}}(z) + K_{n,n} K_{n,0}^* L_{n,n}^{-1} \varphi_n^{\mathrm{L}}(z),$$

from which one easily obtains

$$K_{n,n} \tilde{\varphi}_n^{\mathrm{R}}(z) = \sum_{j=0}^{n} K_{j,j} K_{j,0}^* L_{j,j}^{-1} \varphi_j^{\mathrm{L}}(z).$$

Inserting this into (3.3)

$$z\varphi_n^{\mathrm{L}}(z) = (L_{n,n}^*)^{-1}L_{n+1,n+1}^*\varphi_{n+1}^{\mathrm{L}}(z)$$
$$-(L_{n,n}^*)^{-1}L_{n+1,n+1}^*L_{n+1,0}(K_{n+1,n+1}^*)^{-1}K_{n+1,n+1}^{-1}\sum_{j=0}^{n+1}K_{j,j}K_{j,0}^*L_{j,j}^{-1}\varphi_j^{\mathrm{L}}(z),$$

which is the desired Fourier expansion. From this we can find the Fourier coefficients

$$M_{k,j} = -(L_{k,k}^*)^{-1}L_{k+1,k+1}^*L_{k+1,0}(K_{k+1,k+1}^*)^{-1}K_{k+1,k+1}^{-1}K_{j,j}K_{j,0}^*L_{j,j}^{-1}.$$

Now using (3.4) and (3.5)

$$M_{k,j} = -H_{k+1}\left(\prod_{i=j+1}^{k}(I - H_i^*H_i)\right)H_j^*, \quad j \leqslant k, \tag{4.2}$$

where the product runs from right ($i = j + 1$) to left ($i = k$). A formula for $M_{k,k+1}$ can be obtained in the same way, but is more easily found by comparing the leading coefficient in (4.1), giving

$$M_{k,k+1} = L_{k,k}L_{k+1,k+1}^{-1} = (I - H_{k+1}H_{k+1}^*)^{1/2}. \tag{4.3}$$

5. Some applications of orthogonal matrix polynomials on the real line

In this section we give two applications of orthogonal matrix polynomials on the real line. There are, of course, other applications such as the Lanczos method for block matrices [12]. There is also a connection with orthogonal polynomials on an algebraic harmonic curve [23] which is similar to the connection between orthogonal polynomials on a lemniscate and orthogonal matrix polynomials on the unit circle [22]. We will, however, restrict ourselves to an application in the spectral theory of doubly infinite Jacobi matrices and a connection with scalar orthogonal polynomials with respect to an inner product containing derivatives.

5.1. Spectral theory

The doubly infinite Jacobi matrix

$$J = \begin{pmatrix} \ddots & \ddots & & & & & & 0 \\ \ddots & b_{-2} & a_{-1} & & & & & \\ & a_{-1} & b_{-1} & a_0 & & & & \\ & & a_0 & b_0 & a_1 & & & \\ & & & a_1 & b_1 & a_2 & & \\ & & & & a_2 & b_2 & a_3 & \\ 0 & & & & & \ddots & \ddots & \ddots \end{pmatrix}$$

with $a_k > 0$, $b_k \in \mathbb{R}$ describes a Sturm–Liouville problem on $\mathbb{Z}$ which can be viewed as a discrete version of Sturm–Liouville differential operators on $\mathbb{R}$ [3, 25]. The semi-infinite Jacobi matrix for which $a_k = b_{k-1} = 0$ for $k \leqslant 0$ is closely related to scalar orthogonal polynomials on the real line satisfying the three-term recurrence relation

$$xp_n(x) = a_{n+1}p_{n+1}(x) + b_n p_n(x) + a_n p_{n-1}(x),$$

with boundary conditions $p_0 = 1$ and $p_{-1} = 0$. The orthogonality measure μ for these orthonormal polynomials turns out to be the spectral measure for the semi-infinite Jacobi matrix and the spectrum of this operator is the support of the measure μ. The doubly infinite Jacobi operator $J : \ell_2 \to \ell_2$ can be described by the semi-infinite Jacobi block matrix

$$\mathbb{J} = \begin{pmatrix} B_0 & A_1 & & & \\ A_1^t & B_1 & A_2 & & \\ & A_2^t & B_2 & A_3 & \\ & & A_3^t & B_3 & A_4 \\ & & & \ddots & \ddots & \ddots \end{pmatrix}$$

which contains the 2×2 matrices

$$B_0 = \begin{pmatrix} b_{-1} & a_0 \\ a_0 & b_0 \end{pmatrix}, \quad B_n = \begin{pmatrix} b_{-n-1} & 0 \\ 0 & b_n \end{pmatrix}, \quad n = 1, 2, \ldots$$

$$A_n = \begin{pmatrix} a_{-n} & 0 \\ 0 & a_n \end{pmatrix}, \quad n = 1, 2, \ldots .$$

This block Jacobi matrix $\mathbb{J}$ is closely related to matrix orthogonal polynomials on the real line. The spectral properties of this operator are now described by a matrix measure ρ, and the support of this measure, i.e., the support of the trace measure $\rho_{1,1} + \rho_{2,2}$ gives the spectrum of the operator. This was worked out in [3, 25] and applied in [33].

5.2. Sobolev-type orthogonal polynomials

Sobolev-type orthogonal polynomials p_n ($n = 0, 1, 2, \ldots$) are orthogonal polynomials with an inner product that contains a finite number of derivatives. The orthogonality is thus given by

$$\langle p_n, p_m \rangle = \int p_n(x) p_m(x) \, d\mu(x) + \sum_{i=1}^{M} \sum_{j=0}^{M_i} \lambda_{i,j} p_n^{(j)}(c_i) p_m^{(j)}(c_i) = \delta_{m,n}.$$

It can be shown that such polynomials satisfy a $(2N + 1)$-term recurrence relation

$$h(x) p_n(x) = \sum_{k=-N}^{N} c_{n,k} p_{n+k}(x),$$

where

$$h(x) = \prod_{i=1}^{M} (x - c_i)^{M_i + 1}, \quad N = M + \sum_{i=1}^{M} M_i.$$

If we expand the polynomials p_n into a series using the basis functions $\{x^k h^l(x): k = 0, 1, \dots, N-1, l \in \mathbb{N}\}$, then

$$p_{kN+l}(x) = \sum_{n=0}^{N-1} x^n R_n(p_{kN+l})(h(x)),$$

where $R_n(p_m)$ is a polynomial of degree at most $[m/n]$. Then it is shown in [7] that the matrix polynomials

$$P_n(x) = \begin{pmatrix} R_0(p_{nN})(x) & \cdots & R_{N-1}(p_{nN})(x) \\ R_0(p_{nN+1})(x) & \cdots & R_{N-1}(p_{nN-1})(x) \\ \vdots & \cdots & \vdots \\ R_0(p_{nN+N-1})(x) & \cdots & R_{N-1}(p_{nN+N-1})(x) \end{pmatrix}$$

are left orthogonal polynomials with measure $\rho = M(h^{-1}) + L\delta_0$, where

$$\mathrm{d}M(x) = \begin{pmatrix} \mathrm{d}\mu(x) & x\,\mathrm{d}\mu(x) & \cdots & x^{N-1}\,\mathrm{d}\mu(x) \\ x\,\mathrm{d}\mu(x) & x^2\,\mathrm{d}\mu(x) & \cdots & x^N\,\mathrm{d}\mu(x) \\ x^2\,\mathrm{d}\mu(x) & x^3\,\mathrm{d}\mu(x) & \cdots & x^{N+1}\,\mathrm{d}\mu(x) \\ \vdots & \vdots & \cdots & \vdots \\ x^{N-1}\,\mathrm{d}\mu(x) & x^N\,\mathrm{d}\mu(x) & \cdots & x^{2N-2}\,\mathrm{d}\mu(x) \end{pmatrix},$$

δ_0 is the matrix measure with unit mass at the origin, and

$$L = \sum_{i=1}^{M} \sum_{j=0}^{M_i} \lambda_{i,j} L(i,j),$$

where each $L(i,j)$ is a nonnegative definite rank-one matrix. This relation allows us to study the Sobolev-type orthogonal polynomials in terms of matrix orthogonal polynomials with a matrix inner product that does not contain derivatives but which consists of a rather simple measure to which a mass point at the origin is added.

6. Time series and signal processing

A multivariate time series (multichannel stationary process) is a sequence of random vectors $X_0, X_1, \dots, X_{N-1}$,

$$X_k = \begin{pmatrix} X_{k,1} \\ X_{k,2} \\ \vdots \\ X_{k,p} \end{pmatrix}.$$

We suppose $EX_k = 0$ and

$$EX_k X_{k+n}^* = (E(X_{k,i} X_{k+n,j}))_{i,j=1,\dots,p} = R_n$$

only depends on the lag n. The R_n ($n = 0,1,2,\dots$) are $p \times p$ matrices which contain information on how the random vector X_k interacts with the vectors X_{k+n} which is n units of time away. Such interaction matrices are known as the *autocovariance* matrices for the time series. One can easily show that they form a positive definite sequence of matrices, and thus they can be expressed as

$$R_n = \frac{1}{2\pi} \int_0^{2\pi} e^{in\theta} \, d\rho(\theta),$$

where ρ is some matrix measure on the unit circle (Herglotz).

6.1. Prediction

A first application of orthogonal matrix polynomials on the unit circle is when we want to predict X_N from the past $X_0, \dots, X_{N-1}$. A linear predictor for X_N is of the form

$$\hat{X}_N = \sum_{k=1}^{N} A_{N,k} X_{N-k}, \quad A_{N,k} \in \mathbb{C}^{p \times p},$$

and the linear least-squares predictor is obtained by choosing the matrix coefficients $A_{N,k}$ ($k = 1, 2, \dots, N$) in such a way that the expression

$$E(X_N - \hat{X}_N)(X_N - \hat{X}_N)^*$$

is minimized. Let $(\Omega, \mathscr{F}, P)$ be the probability space on which the random vectors X_k ($k \in \mathbb{Z}$) are living, i.e., $\mathscr{F}$ is a σ-algebra of subsets of Ω (the events) and P is a probability measure defined on $\mathscr{F}$. Then $L_2(\Omega, \mathscr{F}, P)$ is a left Hilbert module with matrix inner product

$$\langle X, Y \rangle = EXY^* = \int_\Omega XY^* \, dP(\omega).$$

On the other hand we can consider the left Hilbert module $L_2(\mathbb{T}, \rho)$ of matrix functions defined on the unit circle, with matrix inner product

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(z) \, d\rho(\theta) g(z)^*, \quad z = e^{i\theta}.$$

Both inner product spaces are closely related to each other in the sense that the mapping

$$X_k \mapsto z^k I$$

is an isometry. The minimization problem for the linear least-squares predictor in $L_2(\Omega, \mathscr{F}, P)$ thus becomes a minimization problem for polynomials in $L_2(\mathbb{T}, \rho)$, i.e., we have to minimize

$$\frac{1}{2\pi} \int_0^{2\pi} P_N(z) \, d\rho(\theta) P_N(z)^*, \quad z = e^{i\theta}$$

over all monic matrix polynomials $P_N(z)$ of degree N. This L_2-minimization problem for polynomials is known and the minimum is equal to $(L_{N,N}^* L_{N,N})^{-1}$ and is attained for the left orthogonal monic matrix polynomial $L_{N,N}^{-1}\varphi_N^{\mathrm{L}}$. Thus the isometry $X_k \mapsto z^k I$ maps the random vector $X_N - \hat{X}_N$ to the matrix polynomial $L_{N,N}^{-1}\varphi_N^{\mathrm{L}}$. Hence the prediction coefficients $A_{N,k}$ $(k = 1, 2, \ldots, N)$ are given by the coefficients of the left orthogonal matrix polynomial of degree N:

$$A_{N,k} = -L_{N,N}^{-1} L_{N,N-k}, \quad k = 1, 2, \ldots, N.$$

One way in which these prediction coefficients $A_{N,k}$ $(k = 1, 2, \ldots, N)$ and the prediction error $(L_{N,N}^* L_{N,N})^{-1}$ are computed is to use the recurrence relations (3.6)–(3.7). This gives the following algorithm:

(1) Initialization:
  (a) $L_{0,0} = R_0^{-1/2}$
  (b) $K_{0,0}^* = R_0^{-1/2}$
(2) Repeat for $m = 1, 2, \ldots, N$
  (a) $H_m = -(L_{m-1,0}R_1 + L_{m-1,1}R_2 + \cdots + L_{m-1,m-1}R_m)K_{m-1,m-1}$
  (b) $(I - H_m H_m^*)^{1/2} L_{m,m} = L_{m-1,m-1}$
  (c) $(I - H_m^* H_m)^{1/2} K_{m,0}^* = H_m^* L_{m-1,m-1}$
  (d) Repeat for $i = 1, 2, \ldots, m-1$
    (i) $(I - H_m H_m^*)^{1/2} L_{m,i} = L_{m-1,i-1} + H_m K_{m-1,m-i-1}^*$
    (ii) $(I - H_m^* H_m)^{1/2} K_{m,m-i}^* = K_{m-1,m-i-1}^* + H_m^* L_{m-1,i-1}$
  (e) $(I - H_m H_m^*)^{1/2} L_{m,0} = H_m K_{m-1,m-1}^*$
  (f) $(I - H_m^* H_m)^{1/2} K_{m,m}^* = K_{m-1,m-1}^*$.

In step (2), the operations (b)–(f) are obtained by equating the coefficients of the powers of z in (3.6) and (3.7). The formula for H_m used in (a) of step (2) follows from (4.2) with $j = 0$. For the scalar case $(p = 1)$ this algorithm was first given by Levinson in 1947. In this case one only needs to process the recurrence relation (3.6) because everything now commutes. For the multivariate case the algorithm was given by Whittle in 1965 (unaware of the identity (3.5), however) and in the same year also by Wiggens and Robinson. See [20, 21] for more details.

6.2. Frequency estimation

Another application in time series analysis consists of frequency estimation. The time series is said to be a stationary *harmonic process* if it is of the form

$$X_n = \sum_{k=1}^{m} [A_k \cos n\lambda_k + B_k \sin n\lambda_k] + Z_n,$$

where $A_k, B_k \in \mathbb{C}^p$ and Z_n is white noise. The frequencies λ_k $(k = 1, 2, \ldots, m)$ are unknown and need to be estimated from the data. This is usually done by using Fourier techniques (fast Fourier transform, periodogram). For the scalar case $(p = 1)$ it was recently suggested to use the zeros of orthogonal polynomials on the unit circle to estimate these frequencies [16–19, 26, 27]. This idea can easily be extended to matrix polynomials on the unit circle.

The underlying principle is that a purely harmonic process (without white noise) has a spectral measure ρ which is purely discrete on the unit circle, with mass points at the points $\{e^{\pm i\lambda_k}: k = 1, 2, \dots, m\}$. In this case the orthogonal polynomials can only be computed up to degree $2m$ and the polynomial of degree $2m$ has its zeros precisely at these mass points. Thus, if the white noise is small, then the orthogonal polynomials will have zeros close to these frequencies on the unit circle. Therefore, we compute zeros of the left orthogonal matrix polynomials associated with the spectral measure ρ and we will estimate the frequencies by taking the argument of those zeros which are closest to the unit circle. The orthogonal polynomials can be computed by using the Levinson–Whittle–Wiggens–Robinson algorithm given above.

We have applied this procedure to a time series containing the light variations of the white dwarf star PG 1159-035. The light intensity of this star was measured every 5 s by different telescopes. This star shows very rapid oscillations and it is of interest to locate as many frequencies as possible. We will use only a small fraction of the data, which was kindly given to us by the authors of [34], and we are grateful to them for giving us the opportunity to use their data. We have used four time periods in which measurements by four different telescopes have been done. Each of the time periods contains 1280 measurements. The first series starts at Julian Date (JD) 3 309 096.918 and ends at JD 3 315 491.918; the second series at JD 2 853 099.795 and ends at JD 2 859 494.795; the third series starts at JD 3 198 688.123 and ends at JD 3 205 083.123; and finally the last series of measurements starts at JD 3 361 849.016 and ends at JD 3 368 244.016. These four time series give us a multivariate time series with $p = 4$. In Fig. 1 we have plotted the zeros of the left orthogonal matrix polynomial of degree 50, which has $4 \times 50 = 200$ zeros. The zeros closest to the unit circle (excluding the real zero close to 1) are indicated by arrows and correspond, in decreasing order of the size of the modulus, to the frequencies $\lambda_1 = 0.062486$, $\lambda_2 = 0.056422$, and $\lambda_3 = 0.070539$. Since we have measurements every 5 s, this corresponds to an oscillation of, respectively, 1989 μHz

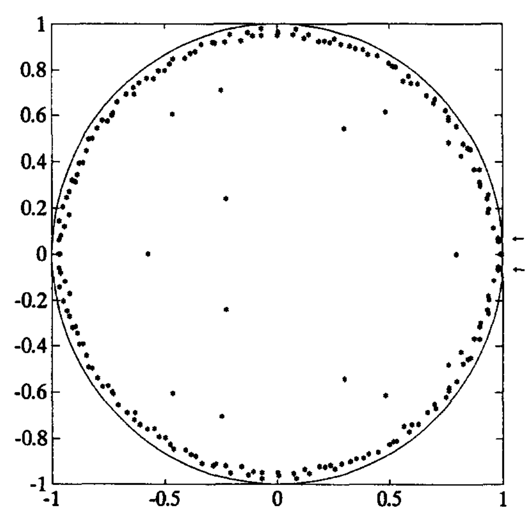

Fig. 1. Zeros of the left orthogonal matrix polynomial of degree 50 ($p = 4$). The zeros closest to the unit circle are indicated by arrows.

for the first frequency, 1796 μHz for the second frequency, and 2245 μHz for the third frequency. These frequencies are indeed close to three most dominant frequencies (1937 μHz, 1854 μHz, and 2214 μHz) with amplitudes 68.9, 61.0 and 40.1, respectively, found in [34, Table 3, p. 336]. Observe that we used only about 7 h of observations, whereas the results in [34] are based on 264 h of observations.

This method is not quite ready yet for implementation and some extra research is desired:

(1) We need criteria for the optimal degree n. This degree should not be too small, because then only small autocovariances are taken into consideration so that small frequencies may not be detected. The degree should not be too high either since then the computational work will become harder.

(2) We need criteria for closeness to the unit circle. We expect some of the zeros to move to the desired frequencies on the unit circle, but most of the other zeros will in general also have the tendency to move to the unit circle. Under rather mild conditions one can show, at least in the scalar case, that the zeros of orthogonal polynomials will asymptotically be uniformly distributed on a circle with center at the origin and radius $r \leqslant 1$ [24], and when $r = 1$, which happens for instance when the orthogonality measure has a density with a singularity on the unit circle, then most of the zeros will move to the unit circle. The zeros which tend to the frequencies, however, will usually tend faster to the circle. Of interest is thus to find a radius r_n such that the zeros of modulus greater than r_n will be selected as zeros which tend to frequencies. The existence of such a radius has been shown for the scalar case in [26, 27].

(3) In order to do statistical inference from the data, one would also like to know the statistical distribution of these estimates.

7. Quadrature

7.1. Gaussian quadrature on the real line

Let ρ be a matrix measure on the interval $[a, b]$. Then we are going to approximate the integral of matrix functions F and G by means of a sum of the form

$$\int_a^b F(x)\,\mathrm{d}\rho(x)G(x)^{\mathrm{t}} \simeq \sum_{i=1}^{k} F(x_i)\Lambda_i G(x_i)^{\mathrm{t}},$$

where $\Lambda_i \in \mathbb{R}^{p\times p}$ and x_i $(i = 1, \ldots, k)$ are points on the interval $[a, b]$. It will be convenient to choose

$$\Lambda_i = (v_{i,1} \;\; \cdots \;\; v_{i,m_i})A_i \begin{pmatrix} v_{i,1}^{\mathrm{t}} \\ \vdots \\ v_{i,m_i}^{\mathrm{t}} \end{pmatrix},$$

where $v_{i,1}, \ldots, v_{i,m_i}$ are linearly independent, nonzero vectors and $\sum_{i=1}^{k} m_i = np$.

In [31] we have shown the following theorem:

Theorem 7.1. *Let (X, J) be a Jordan pair of the orthonormal matrix polynomial $P_n^{\mathrm{L}}(x)$ on the interval $[a, b]$ with respect to the matrix measure ρ. Then we have*

$$\int_a^b F(x)\,\mathrm{d}\rho(x)G(x)^{\mathrm{t}} \simeq \sum_{i=1}^{k} F(x_i)\Lambda_i G(x_i)^{\mathrm{t}}, \tag{7.1}$$

where k is the number of different zeros x_i of the matrix polynomial P_n^{L}, m_i is the multiplicity of x_i, $v_{i,j}$ $(j = 1, 2, \ldots, m_i)$ are the vectors associated with x_i given in the matrix X_i,

$$\Lambda_i = (v_{i,1} \quad \cdots \quad v_{i,m_i})L_i^{-1}\begin{pmatrix} v_{i,1}^{\mathrm{t}} \\ \vdots \\ v_{i,m_i}^{\mathrm{t}} \end{pmatrix},$$

where

$$L_i = \begin{pmatrix} v_{i,1}^{\mathrm{t}} \\ \vdots \\ v_{i,m_i}^{\mathrm{t}} \end{pmatrix} K_{n-1}(x_i, x_i)(v_{i,1} \quad \cdots \quad v_{i,m_i}),$$

and

$$K_{n-1}(x, y) = \sum_{j=0}^{n-1} P_j^{\mathrm{L}}(y)^{\mathrm{t}} P_j^{\mathrm{L}}(x)$$

is the left reproducing kernel. This quadrature formula is exact for matrix polynomials F and G which satisfy

$$\deg F + \deg G \leqslant 2n - 1.$$

This quadrature procedure generalizes the well-known Gaussian quadrature for scalar polynomials to matrix polynomials. In order to compute the quadrature coefficients we need to know the zeros and the root vectors of the orthonormal matrix polynomial P_n^{L} on the interval $[a, b]$ with respect to the matrix measure ρ. Usually these polynomials are known by means of their recurrence coefficients. The next theorem shows how to compute the quadrature coefficients by means of the eigensystem of a symmetric block tridiagonal matrix (see [29]).

Theorem 7.2. *Let $U^{(i,j)}$ $(j = 1, 2, \ldots, m_i)$ be the eigenvectors of the matrix*

$$J_n = \begin{pmatrix} E_0 & D_1 & & & & \\ D_1^{\mathrm{t}} & E_1 & D_2 & & & \\ & D_2^{\mathrm{t}} & E_2 & D_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & D_{n-2}^{\mathrm{t}} & E_{n-2} & D_{n-1} \\ & & & & D_{n-1}^{\mathrm{t}} & E_{n-1} \end{pmatrix}$$

associated with the eigenvalue x_i, then corresponding quadrature coefficient is given by

$$\Lambda_i = (U_0^{(i,1)} \ U_0^{(i,2)} \ \cdots \ U_0^{(i,m_i)}) G_i^{-1} \begin{pmatrix} U_0^{(i,1)\mathrm{t}} \\ U_0^{(i,2)\mathrm{t}} \\ \vdots \\ U_0^{(i,m_i)\mathrm{t}} \end{pmatrix},$$

where

$$(G_i)_{s,t} = U^{(i,s)\mathrm{t}} U^{(i,t)}$$

and $U_0^{(i,j)}$ is the vector consisting of the first p components of $U^{(i,j)}$.

If $V^{(i,j)}$ $(j = 1, 2, \ldots, m_i)$ are the normalized eigenvectors, then the quadrature coefficient is given by

$$\Lambda_i = (V_0^{(i,1)} \ V_0^{(i,2)} \ \cdots \ V_0^{(i,m_i)}) \begin{pmatrix} V_0^{(i,1)\mathrm{t}} \\ V_0^{(i,2)\mathrm{t}} \\ \vdots \\ V_0^{(i,m_i)\mathrm{t}} \end{pmatrix},$$

where $V_0^{(i,j)}$ is the vector consisting of the first p components of $V^{(i,j)}$.

Since the recurrence coefficients are determined only up to a multiplication by a unitary factor, this symmetric block tridiagonal matrix can be reduced to a band matrix of order np with band width $2p + 1$. But still the computation of the eigensystem of this sparse matrix is the most time-consuming part of the algorithm. So we have to pay attention to this part when constructing a parallel algorithm.

To compute the eigensystem we have implemented the divide-and-conquer algorithm described by Arbenz in [2] on an SP1 (scalable power machine) using the message passing library PVM (parallel virtual machine). As in every divide-and-conquer algorithm the given problem is divided into smaller problems and then the solutions of the smaller problems are combined to obtain the solution of the original problem. We divide until all processors have their own small subsystem. In every node we solve the smaller problem by means of the LAPACK-routine for band-symmetric matrices. In order to compute the eigenvectors, a basis of the null space of a Weinstein matrix (usually of order $p \times p$) is determined. These vectors are orthonormalized with respect to a new inner product and then they are transformed. During the last step, when computing the eigenvectors of the original matrix, we only retain the first p components.

Example 7.3. Let $p = 2$ and let ρ be differentiable on $[-1, 1]$ with a matrix density function given by

$$W(x) = \frac{1}{\pi} \frac{1}{\sqrt{1 - x^2}} \begin{pmatrix} 1 & x \\ x & 1 \end{pmatrix}.$$

Table 1
Relative errors for Example 7.3

n	Serial	Nproc = 2	Nproc = 4	Nproc = 8
50	$0.3970 \cdot 10^{-13}$	$0.3159 \cdot 10^{-13}$	$0.5051 \cdot 10^{-13}$	$0.3578 \cdot 10^{-05}$
100	$0.1174 \cdot 10^{-13}$	$0.1212 \cdot 10^{-12}$	$0.2359 \cdot 10^{-13}$	$0.8023 \cdot 10^{-13}$
150	$0.8065 \cdot 10^{-13}$	$0.1247 \cdot 10^{-13}$	$0.8169 \cdot 10^{-13}$	$0.1130 \cdot 10^{-12}$
200	$0.4687 \cdot 10^{-13}$	$0.5924 \cdot 10^{-14}$	$0.5280 \cdot 10^{-13}$	$0.4685 \cdot 10^{-07}$
250	$0.6007 \cdot 10^{-13}$	$0.3024 \cdot 10^{-13}$	$0.1089 \cdot 10^{-12}$	$0.1571 \cdot 10^{-08}$
300	$0.8211 \cdot 10^{-14}$	$0.5967 \cdot 10^{-12}$	$0.9385 \cdot 10^{-13}$	$0.3918 \cdot 10^{-13}$
350	$0.1070 \cdot 10^{-13}$	$0.1295 \cdot 10^{-12}$	$0.3935 \cdot 10^{-12}$	$0.3048 \cdot 10^{-12}$
400	$0.6849 \cdot 10^{-13}$	$0.1416 \cdot 10^{-12}$	$0.1440 \cdot 10^{-12}$	$0.2775 \cdot 10^{-09}$
450	$0.9229 \cdot 10^{-13}$	$0.5767 \cdot 10^{-12}$	$0.4589 \cdot 10^{-12}$	$0.7490 \cdot 10^{-09}$
500	$0.8117 \cdot 10^{-13}$	$0.7398 \cdot 10^{-12}$	$0.4195 \cdot 10^{-12}$	$0.5287 \cdot 10^{-01}$

Let F be the matrix polynomial of degree 30

$$F(x) = \sum_{k=0}^{30} \begin{pmatrix} 4-5k & 5-5k \\ 7-5k & 8-5k \end{pmatrix} x^k,$$

and G the matrix polynomial of degree 20

$$G(x) = \sum_{k=0}^{20} \begin{pmatrix} 2 & -5+2k \\ 5+2k & 2k \end{pmatrix} x^k.$$

To get an idea of the errors, we have computed the integral

$$\int_a^b F(x)W(x)G(x)^{\mathrm{t}}\,\mathrm{d}x$$

by means of Mathematica with a precision of 30 digits and compared these results with the results of the implementation of the Gaussian quadrature rules. The relative errors are given in Table 1.

When using 8 processors, the large errors are due to errors in the computation of determinants to compute the eigenvalues of a matrix out of the solution of two smaller problems (see [29]). It would be better to compute these determinants in higher precision.

The execution times (in seconds) are given in Table 2. The speedup is given between square brackets.

It is true that there is more communication when the number of processors increases, but the reason for the small speedup on 8, and even on 4 processors, is the decreasing number of Mflop/s as the number of processors increases. For the serial algorithm this number is about 64, while for the parallel algorithm on resp. 2, 4, and 8 processors this number becomes resp. 31, 9, 5. The reason for this decrease is the fact that we have a lot of tests in the step where we compute the eigenvalues out of the eigensystem of two smaller problems. These tests are very time-consuming: a comparison of two floating-point numbers is about 7 times slower than a multiply–add instruction.

Table 2
Execution times (in seconds) for Example 7.3. Speedup is given between square brackets

n	Serial	Nproc = 2	Nproc = 4	Nproc = 8
50	0.317	0.492 [0.64]	0.714 [0.44]	0.880 [0.36]
100	2.133	1.564 [1.36]	1.635 [1.30]	2.523 [0.85]
150	6.571	3.060 [2.15]	3.120 [2.11]	3.914 [1.68]
200	14.819	5.730 [2.59]	5.481 [2.70]	5.525 [2.68]
250	28.318	9.502 [2.98]	8.469 [3.34]	8.934 [3.17]
300	47.544	14.334 [3.32]	13.832 [3.44]	14.653 [3.24]
350	73.985	20.520 [3.61]	20.312 [3.64]	21.756 [3.40]
400	108.793	28.279 [3.85]	30.580 [3.55]	30.828 [3.53]
450	153.062	38.062 [3.98]	39.752 [3.85]	40.045 [3.82]
500	208.011	50.078 [4.15]	52.079 [3.99]	52.629 [3.95]

Example 7.4. Let $p = 5$ and let the matrix density function on the interval $[-1, 1]$ be given by

$$W(x) = \begin{pmatrix} \frac{1}{2} & \frac{1}{8}x & 0 & 0 & 0 \\ \frac{1}{8}x & \frac{1}{2} & \frac{1}{4}x & 0 & 0 \\ 0 & \frac{1}{4}x & \frac{1}{2} & \frac{1}{4}\sqrt{2}x & 0 \\ 0 & 0 & \frac{1}{4}\sqrt{2}x & \frac{1}{2} & \frac{1}{20}\sqrt{10}x \\ 0 & 0 & 0 & \frac{1}{20}\sqrt{10}x & \frac{1}{2} \end{pmatrix}.$$

Let F be the matrix polynomial of degree 30

$$F(x) = \sum_{k=0}^{30} \begin{pmatrix} 5-5k & 7-5k & 9-5k & 11-5k & 13-5k \\ 8-5k & 10-5k & 12-5k & 14-5k & 16-5k \\ 11-5k & 13-5k & 15-5k & 17-5k & 19-5k \\ 14-5k & 16-5k & 18-5k & 20-5k & 22-5k \\ 17-5k & 19-5k & 21-5k & 23-5k & 25-5k \end{pmatrix} x^k,$$

and G the matrix polynomial of degree 20

$$G(x) = \sum_{k=0}^{20} \begin{pmatrix} 2k & -5+2k & -10+2k & -15+2k & -20+2k \\ 5+2k & 2k & -5+2k & -10+2k & -15+2k \\ 10+2k & 5+2k & 2k & -5+2k & -10+2k \\ 15+2k & 10+2k & 5+2k & 2k & -5+2k \\ 20+2k & 15+2k & 10+2k & 5+2k & 2k \end{pmatrix} x^k.$$

The relative errors are given in Table 3.

Table 3
Relative errors for Example 7.4

n	Serial	Nproc = 2	Nproc = 4	Nproc = 8
50	$0.1370 \cdot 10^{-13}$	$0.4352 \cdot 10^{-14}$	$0.1192 \cdot 10^{-13}$	$0.1469 \cdot 10^{-13}$
100	$0.2443 \cdot 10^{-13}$	$0.2621 \cdot 10^{-12}$	$0.1284 \cdot 10^{-12}$	$0.9523 \cdot 10^{-14}$
150	$0.9016 \cdot 10^{-14}$	$0.2184 \cdot 10^{-13}$	$0.7654 \cdot 10^{-13}$	$0.2292 \cdot 10^{-01}$
200	$0.5892 \cdot 10^{-13}$	$0.1949 \cdot 10^{-12}$	$0.1069 \cdot 10^{-12}$	$0.4538 \cdot 10^{-03}$
250	$0.1177 \cdot 10^{-12}$	$0.3222 \cdot 10^{-12}$	$0.1399 \cdot 10^{-12}$	$0.1767 \cdot 10^{-02}$
300	$0.1197 \cdot 10^{-12}$	$0.3764 \cdot 10^{-13}$	$0.5447 \cdot 10^{-11}$	$0.8572 \cdot 10^{-01}$

Table 4
Execution times (in seconds) for Example 7.4. Speedup is given between square brackets

n	Serial	Nproc = 2	Nproc = 4	Nproc = 8
50	4.393	4.747 [0.79]	5.605 [0.78]	6.988 [0.63]
100	31.297	21.045 [1.49]	22.244 [1.41]	22.865 [1.37]
150	104.634	51.868 [2.02]	56.467 [1.85]	56.371 [1.86]
200	236.321	100.217 [2.36]	109.911 [2.15]	110.634 [2.14]
250	468.341	169.953 [2.76]	185.709 [2.52]	183.944 [2.55]
300	781.134	260.569 [3.00]	281.707 [2.77]	285.843 [2.73]

The execution times (in seconds) and the speedup (between square brackets) are given in Table 4. Similar conclusions as in Example 7.3 can be made for this example.

7.2. Gaussian quadrature on the unit circle

Let ρ be a matrix measure defined on the unit circle. Then we are going to approximate the integral of matrix functions F and G by means of a sum of the form

$$\frac{1}{2\pi}\int_0^{2\pi} F(z)\,\mathrm{d}\rho(\theta)G(z)^* \simeq \sum_{i=1}^{k} F(z_i)\Lambda_i G(z_i)^*,$$

where $\Lambda_i \in \mathbb{C}^{p\times p}$, and z^i $(i = 1,\ldots,k)$ are points on the unit circle. Again it will be convenient to choose

$$\Lambda_i = (v_{i,1} \quad \cdots \quad v_{i,m_i})A_i \begin{pmatrix} v_{i,1}^* \\ \vdots \\ v_{i,m_i}^* \end{pmatrix},$$

where the nonzero vectors $v_{i,1},\ldots,v_{i,m_i}$ are linearly independent, and $\sum_{i=1}^{k} m_i = np$. We would like to have a formula which is exact for as many Laurent polynomials as possible. We denote by $\Lambda_{-m,n}$ the set of Laurent polynomials of the form $\sum_{k=-m}^{n} a_k z^k$.

We restrict attention to the left inner product $\langle\cdot,\cdot\rangle_{\mathrm{L}}$, but everthing can be repeated for the right inner product $\langle\cdot,\cdot\rangle_{\mathrm{R}}$. Para-orthogonal matrix polynomials

$$B_n(z, W_n) = \phi_n^{\mathrm{L}}(z) + W_n\tilde{\phi}_n^{\mathrm{R}}(z),$$

where W_n is a unitary matrix, are going to play an important role in the quadrature theory. The zeros of the polynomials $B_n(z, W_n)$ are the eigenvalues of a unitary block lower Hessenberg matrix, their multiplicity is less than or equal to p and the length of the corresponding Jordan chains is equal to 1 (see [30]).

Theorem 7.5. *Let (X, J) be a Jordan pair of the para-orthogonal matrix polynomial $B_n(z, W_n) = \phi_n^{\mathrm{L}}(z) + W_n\tilde{\phi}_n^{\mathrm{R}}(z)$, where $\{\phi_n^{\mathrm{L}}(z)\}$ are the left orthonormal matrix polynomials and $\{\phi_n^{\mathrm{R}}(z)\}$ are the right orthonormal matrix polynomials, and let W_n be a unitary matrix. Then we have*

$$\langle F, G\rangle_{\mathrm{L}} = \frac{1}{2\pi}\int_0^{2\pi} F(z)\,\mathrm{d}\rho(\theta)G(z)^* \simeq \sum_{i=1}^{k} F(z_i)\Lambda_i G(z_i)^*, \quad z = \mathrm{e}^{\mathrm{i}\theta},$$

where k is the number of different zeros z_i of $B_n(z, W_n)$, m_i is the multiplicity of z_i, $v_{i,j}$ $(j = 1, 2, \ldots, m_i)$ are the vectors associated with z_i given in the matrix X_i and

$$\Lambda_i = (v_{i,1} \quad \cdots \quad v_{i,m_i})K_i^{-1}\begin{pmatrix} v_{i,1}^* \\ \vdots \\ v_{i,m_i}^* \end{pmatrix},$$

where

$$K_i = \begin{pmatrix} v_{i,1}^* \\ \vdots \\ v_{i,m_i}^* \end{pmatrix} S_{n-1}^{\mathrm{L}}(z_i, z_i)(v_{i,1} \quad \cdots \quad v_{i,m_i})$$

and

$$S_{n-1}^{\mathrm{L}}(z_1, z_2) = \sum_{j=0}^{n-1} \phi_j^{\mathrm{L}}(z_2)^*\phi_j^{\mathrm{L}}(z_1).$$

This quadrature formula is exact for Laurent matrix polynomials

$$F \in \Lambda_{-s,t} \quad \text{and} \quad G \in \Lambda_{-(n-1-t),(n-1-s)},$$

where $0 \leqslant s, t \leqslant n-1$.

To compute the quadrature coefficients, we need to know the zeros and root vectors of the para-orthogonal polynomials. The next theorem shows how to compute the quadrature coefficients by means of the eigensystem of a unitary block lower Hessenberg matrix.

Theorem 7.6. *Let $U^{(i,j)}$ $(j = 1,2,\dots,m_i)$ be the eigenvectors of the matrix*

$$M_n = \begin{pmatrix} M_{0,0} & M_{0,1} & & & & \\ M_{1,0} & M_{1,1} & M_{1,2} & & & \\ \vdots & \vdots & \ddots & \ddots & & \\ \vdots & \vdots & \vdots & \ddots & \ddots & \\ M_{n-2,0} & M_{n-2,1} & \cdots & \cdots & M_{n-2,n-2} & M_{n-2,n-1} \\ M'_{n-1,0} & M'_{n-1,1} & \cdots & \cdots & M'_{n-1,n-2} & M'_{n-1,n-1} \end{pmatrix}$$

associated with the eigenvalue z_i. Then the associated quadrature coefficient is given by

$$\Lambda_i = (U_0^{(i,1)}\ U_0^{(i,2)}\ \cdots\ U_0^{(i,m_i)})G_i^{-1}\begin{pmatrix} U_0^{(i,1)*} \\ U_0^{(i,2)*} \\ \vdots \\ U_0^{(i,m_i)*} \end{pmatrix},$$

where

$$(G_i)_{s,t} = U^{(i,s)*}U^{(i,t)}$$

and $U_0^{(i,j)}$ is the vector consisting of the first p components of $U^{(i,j)}$.

If $V^{(i,j)}$ are the normalized eigenvectors, then the quadrature coefficient is given by

$$\Lambda_i = (V_0^{(i,1)}\ V_0^{(i,2)}\ \cdots\ V_0^{(i,m_i)})\begin{pmatrix} V_0^{(i,1)*} \\ V_0^{(i,2)*} \\ \vdots \\ V_0^{(i,m_i)*} \end{pmatrix},$$

where $V_0^{(i,j)}$ is the vector consisting of the first p components of $V^{(i,j)}$.

The blocks of the matrix M_n are nearly all the same as those given by (4.2), only those of the last row are different:

$$M'_{n-1,j} = -((I - H_nH_n^*)^{-1/2}H_n + (I - H_nH_n^*)^{1/2}(W_n^* + H_n^*)^{-1}) \\ \times (I - H_n^*H_n)^{1/2}\cdots(I - H_{j+1}^*H_{j+1})^{1/2}H_j^*, \quad 0 \leqslant j \leqslant n-1.$$

This modification of the last row makes the block lower Hessenberg matrix unitary, and thus its eigenvalues are on the unit circle.

In order to compute an approximation of a given matrix integral, we have generalized (see [30]) the divide-and-conquer method described in [1, 13, 14]. Also in this complex case we solve the smaller problems by means of a routine from the LAPACK-library and then we combine the solutions of the smaller problems to that of a bigger problem. Again we only compute the first p components of the orthonormalized eigenvectors.

Table 5
Relative errors for Example 7.7

n	Serial	Nproc = 2	Nproc = 4
50	$0.18837 \cdot 10^{-14}$	$0.34917 \cdot 10^{-14}$	$0.26711 \cdot 10^{-11}$
100	$0.14999 \cdot 10^{-14}$	$0.60959 \cdot 10^{-14}$	$0.32993 \cdot 10^{-12}$
150	$0.12240 \cdot 10^{-14}$	$0.45497 \cdot 10^{-14}$	$0.11105 \cdot 10^{-02}$
200	$0.56191 \cdot 10^{-14}$	$0.48696 \cdot 10^{-14}$	$0.53183 \cdot 10^{-11}$
250	$0.32259 \cdot 10^{-14}$	$0.21086 \cdot 10^{-14}$	$0.61844 \cdot 10^{-09}$
300	$0.41379 \cdot 10^{-14}$	$0.13186 \cdot 10^{-13}$	$0.39806 \cdot 10^{-08}$
350	$0.49843 \cdot 10^{-14}$	$0.50799 \cdot 10^{-12}$	$0.19537 \cdot 10^{-05}$
400	$0.42870 \cdot 10^{-14}$	$0.18732 \cdot 10^{-12}$	$0.12672 \cdot 10^{-07}$

Table 6
Execution times (in seconds) for Example 7.7. Speedup is given between square brackets

n	Serial	Nproc = 2	Nproc = 4
50	3.250	0.901 [3.61]	1.167 [2.78]
100	19.651	4.220 [4.66]	4.444 [4.42]
150	61.649	11.944 [5.16]	9.326 [6.61]
200	139.604	25.073 [5.57]	16.867 [8.28]
250	302.348	47.806 [6.32]	28.705 [10.53]
300	451.373	79.764 [5.66]	43.958 [10.27]
350	703.105	118.779 [5.92]	62.243 [11.30]
400	1158.375	173.916 [6.66]	84.741 [13.67]

Example 7.7. Let $p = 2$, $W_n = I$ and

$$\rho(\theta) = \begin{pmatrix} 1 & \cos\theta \\ \cos\theta & 1 \end{pmatrix}.$$

In order to get an idea of the errors, we take $F = G = I$ and this led to the relative errors given in Table 5.

The execution times (in seconds) and the speedup (between square brackets) are given in Table 6.

Due to the great similarity between the implementation of the Gaussian quadrature rule on the real line and on the unit circle, the same remarks can be made in this case. But since we are dealing with complex operations, the multiply–add instructions require more time and this leads to a better performance on 4 processors compared with the performance on 2 processors.

References

[1] G.S. Ammar, W.B. Gragg and L. Reichel, An implementation of a divide and conquer algorithm for the unitary eigenproblem, *ACM Trans. Math. Software* **18** (3) (1992) 292–307.

[2] P. Arbenz, Divide and conquer algorithms for the bandsymmetric eigenvalue problem, *Parallel Comput.* **18** (1992) 1105–1128.

[3] Yu.M. Berezanskiĭ, *Expansions in Eigenfunctions of Self-adjoint Operators*, Transl. Math. Monographs, Vol. 17 (AMS, Providence, RI, 1968).

[4] T.S. Chihara, *An Introduction to Orthogonal Polynomials* (Gordon and Breach, New York, 1978).

[5] Ph. Delsarte, Y. Genin and Y. Kamp, Orthogonal polynomial matrices on the unit circle, *IEEE Trans. Circuits Systems* **25** (3) (1978) 149–160.

[6] A.J. Durán and P. López-Rodriguez, Orthogonal matrix polynomials: zeros and Blumenthal's theorem, manuscript.

[7] A.J. Durán and W. Van Assche, Orthogonal matrix polynomials and higher order recurrences, *Linear Algebra Appl.* **219** (1995) 261–280.

[8] J.S. Geronimo, Matrix orthogonal polynomials on the unit circle, *J. Math. Phys.* **22** (7) (1981) 1359–1365.

[9] J.S. Geronimo, Scattering theory and matrix orthogonal polynomials on the real line, *Circuits Systems Signal Process.* **1** (1982) 471–495.

[10] I. Gohberg, Ed., *Orthogonal Matrix-Valued Polynomials and Applications*, Operator Theory: Advances and Applications, Vol. **34** (Birkhäuser, Basel, 1988).

[11] I. Gohberg, P. Lancaster and L. Rodman, *Matrix Polynomials* (Academic Press, New York, 1982).

[12] G. Golub and R. Underwood, The block Lanczos method for computing eigenvalues, in: J.R. Rice, Ed., *Mathematical Software III* (Academic Press, New York, 1977) 364–377.

[13] W.B. Gragg and L. Reichel, A divide and conquer algorithm for the unitairy eigenproblem, in: M.T. Heath, Ed., *Hypercube Multiprocessors* (SIAM, Philadelphia, PA, 1987) 639–647.

[14] W.B. Gragg and L. Reichel, A divide and conquer algorithm for unitary and orthogonal eigenproblem, *Numer. Math.* **57** (1990) 695–718.

[15] R.A. Horn and C.R. Johnson, *Matrix Analysis* (Cambridge Univ. Press, Cambridge, 1985, reprinted 1993).

[16] W.B. Jones and O. Njåstad, Applications of Szegő polynomials to digital signal processing, *Rocky Mountain J. Math.* **21** (1991) 387–436.

[17] W.B. Jones, O. Njåstad and E.B. Saff, Szegő polynomials associated with Wiener–Levinson filters, *J. Comput. Appl. Math.* **32** (1990) 387–406.

[18] W.B. Jones and E.B. Saff, Szegő polynomials and frequency analysis, in: G. Anastassiou, Ed., *Approximation Theory* (Marcel Dekker, New York, 1992) 341–352.

[19] W.B. Jones, W. Thron, O. Njåstad and H. Waadeland, Szegő polynomials applied to frequency analysis, *J. Comput. Appl. Math.* **46** (1993) 217–228.

[20] T. Kailath, *Modern Signal Processing* (Springer, Berlin, 1985).

[21] T. Kailath, A. Vieira and M. Morf, Inverses of Toeplitz operators, innovations, and orthogonal polynomials, *SIAM Rev.* **20** (1978) 106–119.

[22] F. Marcellán Español and I. Rodríguez González, A class of matrix orthogonal polynomials on the unit circle, *Linear Algebra Appl.* **121** (1989) 233–241.

[23] F. Marcellán and G. Sansigre, On a class of matrix orthogonal polynomials on the real line, *Linear Algebra Appl.* **181** (1993) 97–109.

[24] H.N. Mhaskar and E.B. Saff, On the distribution of zeros of polynomials orthogonal on the unit circle, *J. Approx. Theory* **63** (1990) 30–38.

[25] E.M. Nikishin, Discrete Sturm–Liouville operators and some problems of function theory, *Trudy Sem. Petrovsk.* **10** (1984) 3–77; *J. Soviet Math.* **35** (1986) 2679–2744.

[26] K. Pan, Asymptotics for Szegő polynomials associated with Wiener–Levinson filters, *J. Comput. Appl. Math.* **46** (1993) 387–394.

[27] K. Pan and E.B. Saff, Asymptotics for zeros of Szegő polynomials associated with trigonometric polynomial signals, *J. Approx. Theory* **71** (1992) 239–251.

[28] L. Rodman, Orthogonal matrix polynomials, in: P. Nevai, Ed., *Orthogonal Polynomials: Theory and Practice*, NATO ASI Series C, Vol. 295 (Kluwer, Dordrecht, 1990) 345–362.

[29] A. Sinap, Gaussian quadrature for matrix valued functions on the real line, *J. Comput. Appl. Math.* **65** (1995) 369–385.

[30] A. Sinap, Gaussian quadrature for matrix valued functions on the unit circle, *Electron. Trans. Numer. Anal.* **3** (1995) 95–115.

[31] A. Sinap and W. Van Assche, Polynomial interpolation and Gaussian quadrature for matrix-valued functions, *Linear Algebra Appl.* **207** (1994) 71–115.

[32] G. Szegő, *Orthogonal Polynomials*, AMS Colloq. Publ., Vol. **23** (AMS, Providence, RI, 4th ed., 1975).

[33] W. Van Assche, Weak convergence of orthogonal polynomials, *Indag. Math.* (*N.S.*) **6** (1) (1995). 7–23.

[34] D.E. Winget et al., Asteroseismology of the DOV star PG 1159-035 with the whole earth telescope, *Astrophys. J.* **378** (1991) 326–346.

[35] D.C. Youla and N. Kazanjian, Bauer-type factorization of positive matrices and the theory of matrix polynomials orthogonal on the unit circle, *IEEE Trans. Circuits Systems* **25** (1978) 57–69.