Formal orthogonal polynomials and rational approximation for arbitrary bilinear forms

Adhemar Bultheel and Marc Van Barel

Report TW 163, December 1991

Department of Computer Science, K.U. Leuven

Abstract

Classically, formal orthogonal polynomials are studied with respect to a linear functional, which gives rise to a moment matrix with a Hankel structure. Moreover, in most situations, the moment matrix is supposed to be strongly regular.

This implies a number of algebraic properties which are well known, like for example the existence of a three-term recurrence relation (characterised by a tridiagonal Jacobi matrix), Padé approximation properties etc.

In this note we shall investigate how these formal algebraic properties generalize for moment matrices with no special structure. Subsequently, we shall look especially at the case of a moment matrix with an indefinite Hankel structure and with a nonsymmetric indefinite Toeplitz structure.

Keywords : formal orthogonal polynomials, bilinear forms, rational approximation, Padé approximation.
AMS(MOS) Classification : Primary : 15A63 Secondary : 26C10, 41A21

Formal orthogonal polynomials and rational approximation for arbitrary bilinear forms

Adhemar Bultheel and Marc Van Barel
Department of Computer Science
K.U. Leuven

December 4, 1991

Abstract

Classically, formal orthogonal polynomials are studied with respect to a linear functional, which gives rise to a moment matrix with a Hankel structure. Moreover, in most situations, the moment matrix is supposed to be strongly regular.

This implies a number of algebraic properties which are well known, like for example the existence of a three-term recurrence relation (characterised by a tridiagonal Jacobi matrix), Padé approximation properties etc.

In this note we shall investigate how these formal algebraic properties generalize for moment matrices with no special structure. Subsequently, we shall look especially at the case of a moment matrix with an indefinite Hankel structure and with a non-symmetric indefinite Toeplitz structure.

1 Introduction

In classical situations, formal orthogonal polynomials are studied with respect to some linear functional. The "inner product" (p, q) of two polynomials is defined as

$$(p, q) = \mu(pq)$$

where μ is a linear functional defined on the set of polynomials through the moments

$$\mu(z^k) = \mu_k, \quad k = 0, 1, 2, \ldots$$

This implies that $(z^i, z^j) = (z^{i+j}, 1) = (1, z^{i+j}) = \mu_{i+j}$ and therefore, we get a moment matrix for the basis $\{z^k : k = 0, 1, \ldots\}$ which has a Hankel structure:

$$M = [\mu_{i,j}] = [(z^i, z^j)] = [\mu_{i+j}], \quad (i, j = 0, 1, \ldots).$$

In the so called normal situation, all leading $(n + 1) \times (n + 1)$ submatrices M_n are regular. (The matrix M is then called strongly regular.) This situation is well understood. A classical treatment is [1]. We summarize some of the classical properties.

If the moment matrix is strongly regular, then we can use the Gram-Schmidt procedure to orthogonalize the system of polynomials $1, z, z^2, \ldots$ to generate the orthogonal system

$a_0, a_1, a_2, \ldots$, which may be chosen to be monic polynomials. They satisfy $(a_i, a_j) = \delta_{ij} D_{ii}$, $D_{ii} \neq 0$.

These (monic) orthogonal polynomials satisfy a three-term recurrence relation

$$a_{n+1}(z) = (z + \gamma_{n+1}) a_n(z) + b_{n+1} a_{n-1}(z), \quad a_{-1} = 0, a_0 = 1,$$

where the coefficients are given by

$$\gamma_{n+1} = -\frac{(z a_n, a_n)}{(a_n, a_n)} \quad \text{and} \quad b_{n+1} = -\frac{(z a_n, a_{n-1})}{(a_{n-1}, a_{n-1})}.$$

This three-term recurrence relation is reflected in the Jacobi matrix being a tridiagonal matrix

$$J = \text{tridiag} \begin{pmatrix} -b_2 & -b_3 & -b_4 & \cdots \\ -\gamma_1 & -\gamma_2 & -\gamma_3 & \cdots \\ 1 & 1 & 1 & \cdots \end{pmatrix}$$

because the recurence relation corresponds to writing

$$z[a_0(z), a_1(z), a_2(z), \ldots] = [a_0(z), a_1(z), a_2(z), \ldots] J. \tag{1.1}$$

Many of the properties can be formulated in matrix form (see e.g., [4, 5]). Therefore we introduce the following notations. If p is a polynomial $p = \sum_0^n p_i z^i$, then p will denote its stacking vector, i.e., the column matrix of its coefficients $\mathsf{p} = [p_0, p_1, \ldots]^T$ so that $p(z) = [1, z, z^2, \ldots]\mathsf{p}$. It then also follows that for arbitrary polynomials p and q

$$(p, q) = \mathsf{p}^T M \mathsf{q}.$$

Cutting the stacking vectors to the appropriate length, we may consider the unit upper triangular matrices of sizes $(n+1) \times (n+1)$

$$A_n = [\mathsf{a}_0 \mathsf{a}_1 \cdots \mathsf{a}_n]$$

by which the orthogonality of the first $n+1$ polynomials can be expressed as

$$A_n^T M_n A_n = D_n, \quad D_n = \text{diag}(D_{00}, \ldots, D_{nn}).$$

The finite dimensional version of the orthogonality relation (1.1) reads

$$F_n A_n = A_n T_n$$

with T_n the leading $(n+1) \times (n+1)$ dimensional part of the Jacobi matrix and F_n the Frobenius or companion matrix for the polynomial $a_{n+1}(z)$, i.e.,

$$F_n = F(a_{n+1}) = \begin{bmatrix} 0 & 0 & \cdots & 0 & \\ 1 & & & & \\ & 1 & & & -\underline{\mathsf{a}}_{n+1} \\ & & \ddots & & \\ & & & 1 & \end{bmatrix}, \quad \mathsf{a}_{n+1}^T = [\underline{\mathsf{a}}_{n+1}^T \; 1].$$

Another way of expressing the orthogonality of $a_{n+1}(z)$ to all powers z^k for $k = 0, 1, \ldots, n$ is as follows:

$$M_n \underline{\mathsf{a}}_{n+1} = -[\mu_{0,n+1} \cdots \mu_{n,n+1}]^T \tag{1.2}$$

which can be used to get

$$M_n F_n = M_n^<$$

where as before F_n is the companion matrix for a_{n+1} and $M_n^<$ is the $(n+1)\times(n+1)$ leading part of the left shifted (= up shifted) moment matrix $M^< = [\mu_{i+j+1}]$. We then easily find that

$$A_n^T M_n^< A_n = A_n^T M_n F_n A_n = A_n^T M_n A_n T_n = D_n T_n.$$

The approximating properties are related to the formal series

$$g(z) = \sum_{k=0}^{\infty} \mu_k z^{-(k+1)}$$

(the symbol of the Hankel matrix M). Using the formal expansion $(z-\zeta)^{-1} = \sum_{k=0}^{\infty} \zeta^k z^{-(k+1)}$, we can formally set $g(z) = (1,(z-\zeta)^{-1}) = \mu(1/(z-\zeta))$. We consider rational forms

$$g_n(z) = \frac{c_n(z)}{a_n(z)} = \sum_{k=0}^{\infty} \mu_k^n z^{-(k+1)}.$$

The denominator is the nth orthogonal polynomial and the numerator is the nth polynomial of the second kind, defined by

$$c_n(z) = \mu\left(\frac{a_n(z) - a_n(\zeta)}{z-\zeta}\right).$$

It is then well known that $g_n(z)$ is a Padé approximant for $g(z)$ because

$$\mu_k = \mu_k^n \text{ for } k = 0, 1, \ldots, 2n-1.$$

In this paper we shall now generalize these results for a general moment matrix, with no special restriction whatsoever. The strongly regular case for a general hermitian positive definite moment matrix was discussed in [6]. The degenerate case of a Hankel matrix where the condition of strong regularity has been dropped is studied extensively in [3]. This is a special case of the general situation. Another special case that is interesting to look at is when the moment matrix has a Toeplitz structure. For the strongly regular case, such a study has been done in [2]. Also this will be generalized to a Toeplitz matrix which is not strongly regular.

2 Formal orthogonality: general moment matrix

Consider a complex bilinear form $(\cdot,\cdot)$ defined on the set of complex polynomials $\mathrm{C}[z]$, which is linear in the second factor and conjugate linear in the first one:

$$\begin{aligned}(a_1 p_1 + a_2 p_2, q) &= \overline{a}_1(p_1, q) + \overline{a}_2(p_2, q)\\ (p, a_1 q_1 + a_2 q_2) &= a_1(p, q_1) + a_2(p, q_2).\end{aligned}$$

It is defined by its moments

$$\mu_{ij} = (z^i, z^j), \quad i,j = 0,1,2,\ldots \tag{2.1}$$

If it is hermitian, then $\mu_{ji} = \overline{\mu}_{ij}$ but we shall not preassume this symmetry in general. The moments are collected in a moment matrix $M = [\mu_{ij}]$. The submatrix obtained by selecting rows $0, 1, \ldots, k-1$ and columns $0, 1, \ldots, l-1$ from M will be denoted as M_{kl}. In general we shall not suppose that M is positive definite nor that it is strongly regular. The latter means that the leading submatrices $M_{\nu\nu}$ are not necessarily regular for all ν.

If p is a complex polynomial $p = \sum_0^n p_i z^i \in \mathbf{C}[z]$, then $\mathbf{p}$ will denote its stacking vector, i.e., the column matrix of its coefficients $\mathbf{p} = [p_0, p_1, \ldots]^T$ so that $p(z) = [1, z, z^2, \ldots]\mathbf{p}$. It then also follows that

$$(p, q) = \mathbf{p}^H M \mathbf{q}, \quad p, q \in \mathbf{C}[z]$$

where ^{H} means complex conjugate transpose.

If the moment matrix were strongly regular, then we could use the two-sided Gram-Schmidt procedure to orthogonalize the system of polynomials $1, z, z^2, \ldots$ to generate the biorthogonal system $\{b_i, a_i\}_{i \geq 0}$, which may be chosen to be monic polynomials. Biorthogonality means that the polynomials satisfy $(b_i, a_j) = \delta_{ij} D_{ii}$, $D_{ii} \neq 0$. The set $\{b_i\}$ is the set of left orthogonal polynomials and $\{a_i\}$ is the set of right orthogonal polynomials.

In general however, we can only guarantee the existence of a block biorthogonal system. Let us define this more precisely. Suppose the submatrices $M_{\nu\nu}$ of the moment matrix M for $\nu = \nu_1, \nu_2, \ldots, \nu_N$ are the successive regular ones. If N is finite, we define $\nu_{N+1} = \infty$ and in any case set $\nu_0 = 0$. Then $\alpha_{n+1} = \nu_{n+1} - \nu_n$ for $n = 0, 1, \ldots$ will be the number of polynomials that will be grouped in block n.

In what follows, we shall concentrate on the right orthogonal polynomials (ROPs). Of course, similar results can be obtained for the left orthogonal polynomials (LOPs).

So let block n consist of the α_{n+1} ROPs

$$\mathbf{a}_n = [a_{\nu_n}, a_{\nu_n+1}, \ldots, a_{\nu_n+\alpha_{n+1}-1}].$$

We shall also use the notation a_n^k to denote the kth polynomial of block n, thus $a_n^k = a_{\nu_n+k-1}$, $k = 1, \ldots, \alpha_{n+1}$. Using a similar notation for the LOPs, one has

$$(\mathbf{b}_i, \mathbf{a}_j) = [(b_i^k, a_j^l)] = \delta_{ij} D_{jj}$$

where D_{jj} is some regular $\alpha_{j+1} \times \alpha_{j+1}$ matrix and δ_{ij} represents the Kronecker delta. The first ROP a_n^1 of block n has degree ν_n and it is right orthogonal to all polynomials of lower degree:

$$(z^i, a_n^1(z)) = 0 \text{ for } i = 0, 1, \ldots, \nu_n - 1.$$

Note however, that (z^{ν_n}, a_n^1) can be zero. Following Draux [3], we shall call it a regular ROP. It is unique if monically normalized since its stacking vector has to satisfy a linear system

$$M_{n-1} \underline{\mathbf{a}}_n^1 = -[\mu_{0,\nu_n}, \ldots, \mu_{\nu_n-1,\nu_n}]^T \tag{2.2}$$

where $\mathbf{a}_n^1 = [\underline{\mathbf{a}}_n^{1\,T}\ 1]^T$ and $M_{n-1} = M_{\nu_n,\nu_n}$ is a regular submatrix of M. The remaining polynomials in block n however are not uniquely defined. We shall require that they have precise degree i.e., a_n^k has degree $\nu_n + k - 1$ (we could e.g., make them monic) but they are only required to be orthogonal to all polynomials in previous blocks but not to previous polynomials in the same block. The latter would be impossible for all polynomials in a block because of the singularity of the leading submatrices when one is inside a block.

There is a way to make them uniquely defined though, when apart from being monic, we require a_n^k to be not only orthogonal to all previous blocks, but also to all polynomials in the

same block, with the exception of one. Thus a_n^k is orthogonal to all a_m^i for $m = 0, 1, \ldots, n-1$, $i = 1, \ldots, \alpha_{m+1}$ and also to a_n^l, $l = 1, \ldots, \alpha_{n+1}$, $l \neq l_{k,n}$ with $\{l_{k,n} : k = 1, \ldots, \alpha_{n+1}\} = \{1, \ldots, \alpha_{n+1}\}$. That is, we require $D_{nn} = (\mathbf{a}_n, \mathbf{a}_n)$ to be a permuted diagonal matrix. This is always possible since one can show that for every regular matrix D, there exists a unit upper triangular matrix R and a unit lower triangular matrix L such that LDR is a permuted diagonal.

We shall however not need the latter more strict definition of block biorthogonality. When we say block right orthogonal we mean that the polynomials in a block are of strict degree (monic) and right orthogonal to the polynomials in the previous left blocks where the start position and the size of each block is dictated by the rank structure of M through the indices ν_n (the block indices) and α_{n+1} (the block sizes).

The formal Fourier expansions with respect to a biorthogonal polynomial system with left/right blocks $\{\mathbf{b}_k, \mathbf{a}_k\}$, $k = 0, 1, \ldots$, can be given as

$$a(z) = \sum_{k=0}^{\infty} \mathbf{a}_k(z) c_k, \text{ with } c_k = D_{k,k}^{-1}(\mathbf{b}_k, a)$$

Let M_n be a regular submatrix of M of size ν_{n+1} and suppose $M_n^{-1} = [m_{ij}]$. Then define the *kernel function* $k_n(x, y)$ to be

$$k_n(x, y) = \sum_{i,j=0}^{\nu} x^j m_{ij} \overline{y}^i, \quad \nu = \nu_{n+1} - 1.$$

For this kernel, the following properties hold.

Theorem 2.1 *Let M be a moment matrix with some block index ν_n and M_n the corresponding regular submatrix. Let $k_n(x, y)$ be the kernel as defined above. Then with $\nu = \nu_{n+1} - 1$*

$$k_n(x, y) = -\frac{1}{\det M_n} \det \begin{bmatrix} & & 1 \\ & M_n & \overline{y} \\ & & \vdots \\ & & \overline{y}^{\nu} \\ 1 & x \cdots & x^{\nu} \quad 0 \end{bmatrix} = \sum_{i=0}^{n} \mathbf{a}_i(x) D_{ii}^{-1} \mathbf{b}_i(y)^H \tag{2.3}$$

where $\{\mathbf{b}_k, \mathbf{a}_k\}$ denotes the kth blocks of left/right orthogonal polynomials and as before $(\mathbf{b}_i, \mathbf{a}_j) = \delta_{ij} D_{jj}$.

Moreover, for any polynomials of degree at most $\nu = \nu_{n+1} - 1$, the kernel $k_n(x, y)$ is reproducing:

$$(a(x), k_n(x, y)) = \overline{a(y)}.$$

Proof. For (2.3), use the definition

$$k_n(x, y) = [1, x, x^2, \ldots, x^{\nu}] M_n^{-1} [1, y, \ldots, y^{\nu}]^H \tag{2.4}$$

and a property of determinants to find that the determinant of the bordered matrix in (2.3) equals

$$\det M_n \det(0 - [1 \ldots x^{\nu}] M_n^{-1} [1 \ldots y^{\nu}]^H)$$

which proves the first result.

If the stacking vectors for the ROPs in the blocks $\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_n$ are collected in the unit upper triangular matrix A_n, (and similarly for the LOPs) then we know that $B_n^H M_n A_n = D_n = \mathrm{diag}(D_{00}, \ldots, D_{nn})$. Filling this into (2.4) gives the second expression.

For the reproducing property note that

$$(a(x), k_n(x,y)) = \sum_{i=0}^{n} (a(x), \mathbf{a}_i(x)) D_{ii}^{-1} \mathbf{b}_i(y)^H$$

which is the (formal) Fourier expansion of $\overline{a(y)}$ because if we set $a(y) = \sum d_i \mathbf{b}_i(y)^T$ with $d_i \in \mathbf{C}^{1 \times \alpha_{i+1}}$, then we find from the orthogonality that

$$\overline{d}_i = (a(x), \mathbf{a}_i(x)) D_{ii}^{-1},$$

which proves the theorem. □

The generalization of the Jacobi matrix for classical orthogonal polynomials becomes an upper Hessenberg matrix when the moment matrix has no special structure. The block two-sided Gram-Schmidt procedure will generate a possible set of monic block left/right orthogonal polynomials by the following procedure. Start with $a_0 = 1$ and then generate

$$a_{\nu+1}(z) = z a_\nu(z) - \sum_{i=1}^{k-1} \mathbf{a}_i(z) (\mathbf{b}_i(\zeta), \mathbf{a}_i(\zeta))^{-1} (\mathbf{b}_i(\zeta), \zeta a_\nu(\zeta))$$

for the polynomials in block k, i.e. for $\nu_k \leq \nu + 1 < \nu_{k+1}$. (A similar relation holds for the LOPs). If you place all the generated polynomials of block k in the vector $\mathbf{a}_k = [a_{\nu_k}, \ldots, a_{\nu_{k+1}-1}]$ then you know that the block is complete as soon as the matrix $(\mathbf{b}_k, \mathbf{a}_k)$ is regular. This defines the block size α_{k+1} and together with the previous ones also the block index $\nu_{k+1} = \sum_{i=1}^{k+1} \alpha_i$. For $n > k$ this can be summarized in the relation

$$z\mathbf{a}_k(z) = [\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_n][T_{0,k}^T \ldots T_{k,k}^T, T_{k+1,k}^T, 0 \ldots 0]^T. \tag{2.5}$$

The block $T_{i,k}$ is a matrix of size $\alpha_{i+1} \times \alpha_{k+1}$, the diagonal block is square and has the form of a companion matrix, i.e., ones on the subdiagonal and its last column is equal to $(\mathbf{b}_k, \mathbf{a}_k)^{-1} (\mathbf{b}_k, z a_{\nu_{k+1}-1})$ and the subdiagonal block $T_{k+1,k}$ is zero except for its right top element which is 1. This last column is special because it relates the first polynomial a_{k+1}^1 of the next block with the previous blocks $\mathbf{a}_0, \ldots, \mathbf{a}_k$. Writing this out for $k = 0, \ldots, n$ and translating this to the matrices of stacking vectors we get

$$F_n A_n = A_n T_n \tag{2.6}$$

where as before A_n is the unit upper triangular matrix containing the stacking vectors for the first $n+1$ blocks, $F_n = F(a_{n+1}^1)$ is the companion matrix for the first polynomial a_{n+1}^1 of block $n+1$, and T_n is a unit upper Hessenberg matrix

$$T_n = \begin{bmatrix} T_{0,0} & T_{0,1} & T_{0,2} & \cdots & T_{0,n} \\ T_{1,0} & T_{1,1} & T_{1,2} & & T_{1,n} \\ & T_{2,1} & T_{2,2} & \ddots & T_{2,n} \\ & & \ddots & \ddots & \vdots \\ & & & & T_{n-1,n} \\ & & & T_{n,n-1} & T_{n,n} \end{bmatrix}.$$

The companion matrix F_n appears because we gave a finite dimensional version. When $n \to \infty$, we could at least formally, write the infinite relation

$$z[\mathbf{a}_0(z), \mathbf{a}_1(z), \ldots] = [\mathbf{a}_0(z), \mathbf{a}_1(z), \ldots]T_\infty.$$

The operator T_∞ is the generalization of the Jacobi matrix (here written for monic polynomials). In the classical case it is tridiagonal, and as we shall see in the next section, it is block tridiagonal for a general Hankel moment matrix, reducing to a scalar tridiagonal when it is strongly regular.

Recall that $M = [\mu_{ij}]$ and $M_n = M_{\nu_{n+1},\nu_{n+1}}$ its (regular) leading $\nu_{n+1} \times \nu_{n+1}$ submatrix. Now define the left shifted moment matrix $M_n^< = [\mu_{i,j+1}]$ and $M_n^< = M^<_{\nu_{n+1},\nu_{n+1}}$ its leading $\nu_{n+1} \times \nu_{n+1}$ submatrix. Then the orthogonality of a_{n+1}^1 can be expressed as

$$M_n F_n = M_n^< \tag{2.7}$$

where as before F_n represents the companion matrix $F_n = F(a_{n+1}^1)$. We also know that, because of the block biorthogonality

$$B_n^H M_n A_n = D_n \tag{2.8}$$

where A_n (B_n) is the unit upper triangular matrix of stacking vectors for the ROPs (LOPs). Hence, by combining (2.6–2.8) we find

$$\begin{aligned} B_n^H M_n^< A_n &= B_n^H M_n F_n A_n & (2.9)\\ &= B_n^H M_n A_n T_n \\ &= D_n T_n. & (2.10) \end{aligned}$$

As a direct consequence, we find the following determinant expressions for a_{n+1}^1.

Theorem 2.2 *The regular ROP a_{n+1}^1 of block $n+1$ is given by the following determinant expressions*

$$\begin{aligned} a_{n+1}^1(z) &= \det(zI - F_n) \\ &= \det M_n^{-1} \det(zM_n - M_n^<) \\ &= \det(zI_n - T_n) \\ &= \det M_n^{-1} \det \begin{bmatrix} & & \mu_{0,\nu+1} \\ & M_n & \vdots \\ & & \mu_{\nu,\nu+1} \\ 1 & z \ \cdots & z^{\nu+1} \end{bmatrix}, \quad \nu + 1 = \nu_{n+1} \end{aligned}$$

where we have used the notation that was introduced above.

Proof. The first relation holds by definition of companion matrix. The second one follows by multiplying from the left with M_n and using $M_n^< = M_n F_n$. This result, multiplied from the left with B_n^H and from the right with A_n gives the next line because $\det B_n = \det A_n = 1$ and $zB_n^H M_n A_n - B_n^H M_n^< A_n = zD_n - D_n T_n$ while $\det M_n = \det D_n$. For the last line, note that the determinant on the right equals

$$\det M_n(z^{\nu+1} - [1 \ z \ \cdots \ z^\nu] M_n^{-1} [\mu_{0,\nu+1} \cdots \mu_{\nu,\nu+1}]^T).$$

The rest follows from (2.2). □

We now give some general rational approximations that can be obtained from the elements we have collected so far.

We associate with row i of the moment matrix M the following formal series

$$g^{[i]}(z) = \sum_{k=0}^{\infty} \frac{\mu_{i,k}}{z^{k+1}}, \quad i = 0, 1, 2, \ldots \tag{2.11}$$

We can write this formally as

$$g^{[i]}(z) = (\zeta^i, (z - \zeta)^{-1}). \tag{2.12}$$

Furthermore, using the notation introduced above, define

$$\begin{aligned} g_n^{[i]}(z) &= e_i^T M_n (z M_n - M_n^<)^{-1} M_n e_0 \\ &= \sum_{k=0}^{\infty} \frac{\mu_{i,k}^n}{z^{k+1}}, \quad i = 0, 1, 2, \ldots, \nu_{n+1} - 1 \end{aligned} \tag{2.13}$$

where e_i is the ith unit vector. Then $\mu_{i,k} = \mu_{i,k}^n$ for $k = 0, 1, \ldots, \nu_{n+1}$. We formulate this as a theorem.

Theorem 2.3 *Let $g^{[i]}$ be defined by (2.11) and the rational function $g_n^{[i]}$ by (2.13). With the notation introduced before, the latter can also be expressed as*

$$g_n^{[i]}(z) = e_i^T B_n^{-H} D_n (zI - T_n)^{-1} e_0 \tag{2.14}$$

or in the rational form

$$g_n^{[i]}(z) = c_{n+1}^{[i]}(z) / a_{n+1}^1(z) \tag{2.15}$$

with the numerator a polynomial of the second kind given by

$$c_{n+1}^{[i]}(z) = (\zeta^i, \frac{a_{n+1}^1(z) - a_{n+1}^1(\zeta)}{z - \zeta}). \tag{2.16}$$

It has the determinant expression

$$c_{n+1}^{[i]}(z) = \det M_n^{-1} \det \begin{bmatrix} M_n e_0 & M_n^< \\ 0 & u_n \end{bmatrix} \tag{2.17}$$

with u_n the polynomial part of

$$[z \; z^2 \; \ldots \; z^{\nu_{n+1}}] g^{[i]}(z).$$

Moreover, $g^{[i]}$ and $g_n^{[i]}$ correspond in their initial terms:

$$\mu_{i,k} = \mu_{i,k}^n \text{ for } k = 0, 1, \ldots, \nu_{n+1}.$$

Proof. The alternative expression (2.14) for $g_n^{[i]}$ can be easily found from (2.13) when using (2.8) and (2.10).

To prove the correspondence of the expansions of $g^{[i]}$ and $g_n^{[i]}$, we can start from (2.14) and use a formal expansion of $(zI - T_n)^{-1}$.

$$g_n^{[i]}(z) = e_i^T B_n^{-H} D_n (\sum_{k=0}^{\infty} \frac{T_n^k}{z^{k+1}}) e_0.$$

Because $A_n^{-1}F_nA_n = T_n$, as we can see from (2.6), we also have

$$T_n^k e_0 = A_n^{-1}F_n^k A_n e_0 = A_n^{-1}F_n^k e_0$$

because A_n is unit upper triangular. Thus

$$\begin{aligned} g_n^{[i]}(z) &= e_i^T B_n^{-H} D_n A_n^{-1} \sum_{k=0}^{\infty} z^{-(k+1)} F_n^k e_0 \\ &= e_i^T M_n \sum_{k=0}^{\infty} z^{-(k+1)} F_n^k e_0 \\ &= [\mu_{i,0}, \ldots, \mu_{i,\nu}] \sum_{k=0}^{\infty} z^{-(k+1)} F_n^k e_0 \end{aligned}$$

where $\nu = \nu_{n+1} - 1$. It is easily verified that

$$F_n^k e_0 = \begin{cases} e_k & \text{for } k = 0, \ldots, \nu = \nu_{n+1} - 1 \\ -\underline{\mathbf{a}}_{n+1}^1 & \text{for } k = \nu_{n+1} \end{cases}$$

This proves that

$$g_n^{[i]}(z) = \sum_{k=0}^{\nu} z^{-(k+1)} \mu_{i,k} - e_i^T M_n \underline{\mathbf{a}}_{n+1}^1 z^{-(\nu+2)} + \cdots$$

Since also

$$e_i^T M_n \underline{\mathbf{a}}_{n+1}^1 = -\mu_{i,\nu+1},$$

we have proved the correspondence.

The expression (2.16) follows from the formal equalities

$$(\zeta^i, \frac{a_{n+1}^1(z) - a_{n+1}^1(\zeta)}{z - \zeta}) = g^{[i]}(z) a_{n+1}^1(z) - (\zeta^i, \frac{a_{n+1}^1(\zeta)}{z - \zeta}) \tag{2.18}$$

where we have used the expression (2.12) for $g^{[i]}$. Expanding $(z - \zeta)^{-1}$ as $\sum \zeta^k z^{-(k+1)}$, one sees that the second term is of the form $\sum_{k=0}^{\infty} \eta_k z^{-(k+1)}$ with η_k the coefficient of $z^{-(k+1)}$ in the expansion of $g^{[i]}a_{n+1}^1$. Therefore the left hand side of (2.18) is polynomial. Since a_{n+1}^1 has degree ν_{n+1} and $\mu_{ij} = \mu_{ij}^n$ for $j = 0, \ldots, \nu_{n+1}$, it follows that the polynomial part of $g_n^{[i]}a_{n+1}^1$ (which is $c_{n+1}^{[i]}$ by definition) and of $g^{[i]}a_{n+1}^1$ is the same. This proves (2.16).

The determinant expression (2.17) for the numerator can be found as follows. First note that

$$\det \begin{bmatrix} zM_n - M_n^{<} & M_n e_0 \\ e_i^T M_n & 0 \end{bmatrix} = \det(zM_n - M_n^{<})(-e_i^T M_n (zM_n - M_n^{<})^{-1} M_n e_0)$$

The first factor equals $\det M_n a_{n+1}^1$ and the second one equals $-g_n^{[i]}$. Hence the numerator $c_{n+1}^{[i]}$ is given by

$$c_{n+1}^{[i]}(z) = -\det M_n^{-1} \begin{bmatrix} zM_n - M_n^{<} & M_n e_0 \\ e_i^T M_n & 0 \end{bmatrix}.$$

This gives the desired result after some elementary matrix operations that do not change the determinant. □

The previous theorem shows a kind of vector Padé approximation. Indeed, when we consider the vector with components $g^{[i]}(z)$, $i = 0, \ldots, \nu - 1$ where $\nu = \nu_{n+1}$ and the vector of rational approximants $g_n^{[i]}(z)$, then, since $\deg c_{n+1}^{[i]} < \deg a_{n+1}^1$, the number of parameters in the latter is ν^2 (for the numerator coefficients) + ν (for the monic denominator), giving a total of $(\nu+1)\nu$ parameters, which is precisely the number of coefficients that is fitted in the ν series. The block two-sided Gram-Schmidt procedure computes the denominators of the approximants for increasing n. This generates approximants of increasing degree ν_n, but also the number of series and the number of approximants is increasing! However, it is possible to keep this number constant when we choose a special moment matrix. We shall not go into the details here. An extreme example is the case where the moment matrix is Hankel, in which case only one series is fitted, the others being shifted versions. The general idea is that we can think of the moment matrix consisting of horizontal stripes (possibly obtained after row permutation) where we assume each stripe to have a Hankel structure. Then the number of different stripes corresponds to the number of series that are fitted. The number of rows in each stripe is related to the number of extra coefficients that are fitted (on top of the number that is always fitted in the general case) for the series that is associated with that stripe.

3 Formal orthogonality: the Hankel case

Now we shall suppose that the moment matrix has a Hankel structure, i.e., $\mu_{ij} = \mu_{i+j+1}$. Classically, such a situation appears with real moments and real orthogonal polynomials. If we allow the moments and the polynomials to be complex, the symmetry of the moment matrix implies that if $a_k(z) = [1\ z\ z^2 \cdots]\mathbf{a}_k$ is the kth ROP, then we can choose $b_k(z) = \overline{a}_k(z) = [1\ z\ z^2 \cdots]\overline{\mathbf{a}}_k$ to be the kth LOP. The real situation has been studied extensively by Draux in [3]. We shall give a very short survey here.

It now holds that $(zf(z), g(z)) = (f(z), zg(z))$. This translates in the matrix relation

$$Z^T M = M^< = MZ$$

where Z is the downshift operator and $M^<$ is the left (= up) shifted moment matrix obtained by deleting the first column (or the first row) in M.

The determination of the block sizes can now be obtained as follows.

Theorem 3.1 *Suppose ν_k is some block index for the moment matrix M and a_k^1, the first polynomial of block k which is by definition orthogonal to z^i for $i = 0, \ldots, \nu$ where $\nu+1 = \nu_k$. Now suppose it is over-orthogonal in the sense that also $(z^i, a_k^1) = 0$ for $i = \nu+1, \nu+2, \ldots, \nu+\alpha-1$ but nonzero for $i = \nu+\alpha$, then the size of block k will be α.*

Proof. We may indeed choose $a_{\nu+i}(z) = z^{i-1}a_{\nu+1}(z)$ for $i = 1, \ldots, \alpha$ since it follows from the above mentioned relation that $a_{\nu+i}$ is orthogonal to all z^j for $j = 0, 1, \ldots, \nu+\alpha-i$. Hence to all polynomials in the previous blocks as long as $1 \leq i \leq \alpha$. Collecting these polynomials in a block

$$\mathbf{a}_k = [a_{\nu+1} \ldots, a_{\nu+\alpha}]$$

we find $(\overline{\mathbf{a}}_k, \mathbf{a}_k) = D_{kk}$ which has a lower triangular Hankel structure with nonzero diagonal element. Thus we can not take less than α elements in a block because then D_{kk} would not be regular. In other words the block size of block k is $\alpha_{k+1} = \alpha$. □

As a consequence of the above construction we can see that the orthogonality of $z^{\alpha_{k+1}} a_k^1$ to z^i is automatic for $i = 0, \ldots, \nu_k - 2$. This implies that the general relation (2.5) reduces to

$$z\mathbf{a}_k(z) = [\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_n][0 \ldots 0, T_{k-1,k}^T, T_{k,k}^T, T_{k+1,k}^T, 0 \ldots 0]^T \tag{3.1}$$

where $T_{k,k}$ and $T_{k+1,k}$ have the same structure as in the general case, while $T_{k-1,k}$ is all zero with the exception of the right top element. The relation (2.6) thus becomes

$$F_n A_n = A_n T_n$$

where as before F_n is the companion matrix of a_{n+1}^1 and T_n has now a block triangular form

$$T_n = \begin{bmatrix} T_{0,0} & T_{0,1} & & & \\ T_{1,0} & T_{1,1} & T_{1,2} & & \\ & T_{2,1} & T_{2,2} & \ddots & \\ & & \ddots & \ddots & T_{n-1,n} \\ & & & T_{n,n-1} & T_{n,n} \end{bmatrix}.$$

If M is strongly regular, all its blocks have size 1×1, and then this is an ordinary tridiagonal matrix which is known as the (nonsymmetric) Jacobi matrix which reflects the three term recurrence relation. In this regular case each block consists of one polynomial $\mathbf{a}_k = a_k = a_k^1$ and

$$a_{n+1}^1(z) = (z + \gamma_{n+1}) a_n^1(z) + b_{n+1} a_{n-1}^1(z)$$

corresponds to $T_{n-1,n} = -b_{n+1}$, $T_{n,n} = -\gamma_{n+1}$. For the block case, this generalizes to

$$a_{n+1}^1(z) = a_{n+1}(z) a_n^1(z) + b_{n+1} a_{n-1}^1(z)$$

where $a_{n+1}(z)$ is a monic polynomial of degree α_{n+1} for which $T_{n,n}$ is the companion matrix, while $-b_{n+1}$ is the right top element of $T_{n-1,n}$. For the initial conditions one may choose $a_{-1}^1 = 0$ and $a_0^1 = 1$.

With this generalized 'three term recurrence relation' we may associate the continued fraction

$$\cfrac{b_1}{a_1(z)} + \cfrac{b_2}{a_2(z)} + \cfrac{b_3}{a_3(z)} + \cdots$$

whose convergents are c_k^1 / a_k^1, $k = 1, 2, \ldots$ where the initial conditions are

$$\begin{bmatrix} c_{0,-1} & c_{0,0} \\ a_{0,-1} & a_{0,0} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

We shall see that these convergents are the approximants that were discussed in the previous section in the more general situation. Because the $g^{[i]}$ are shifted versions of each other when M is a Hankel matrix, we now only consider one formal series

$$g(z) = \sum_{k=1}^{\infty} \mu_k z^{-k}. \tag{3.2}$$

The approximants (take $i = 0$ in the general situation) are now

$$\begin{aligned} g_n(z) &= e_0^T M_n (z M_n - M_n^<)^{-1} M_n e_0 \\ &= \sum_{k=1}^{\infty} \mu_k^n z^{-k} \end{aligned} \tag{3.3}$$

We know from the general theory of the previous section that g_n fits $\nu+2=\nu_{n+1}+1$ of the coefficients of g, i.e., the coefficients of z^k for $k=0,\ldots,\nu_{n+1}$. We can verify however that the correspondence of the coefficients goes further now. Like in the general theorem 2.3, we find

$$g_n = e_0^T M_n \sum_{k=0}^{\infty} z^{-(k+1)} F_n^k e_0.$$

Use

$$M_n F_n = M_n^{<} \text{ and } e_j^T M_n^{<} = e_{j+1}^T M_n, \quad 0 \le j < \nu$$

ν times which will result in

$$e_0^T M_n F_n^{\nu} = e_{\nu}^T M_n.$$

Hence

$$g_n(z) = \sum_{k=1}^{\nu+1} \mu_k z^{-k} + \sum_{k=1}^{\nu+1} e_{\nu}^T M_n F_n^k e_0 z^{-(\nu+1+k)} + \cdots$$

But as we know from the previous section it holds for $1 \le k \le \nu+1$ that

$$\begin{aligned} e_{\nu}^T M_n F_n^k e_0 &= [\mu_{\nu,0} \cdots \mu_{\nu,\nu}] F_n^k e_0 \\ &= \mu_{\nu,k} = \mu_{\nu+k+1}. \end{aligned}$$

Hence

$$g_n(z) = \sum_{k=1}^{2(\nu+1)} \mu_k z^{-k} + \cdots$$

This proves that μ_k and μ_k^n correspond for at least $k=1,\ldots,2(\nu+1)=2\nu_{n+1}$. Because a_{n+1}^1 is over-orthogonal by α_{n+2} as we have seen, the fit is even better:

$$\mu_k = \mu_k^n, \text{ for } k = 1,\ldots,2\nu_{n+1}+\alpha_{n+2} = 2\nu_{n+2}-\alpha_{n+2}.$$

4 Formal orthogonality: the Toeplitz case

Now we shall suppose that the moment matrix M is Toeplitz, i.e., $\mu_{i,j}=\mu_{j-i}$, but in general it need not be hermitian or symmetric. Thus we should distinguish between left and right orthogonality. Again, for simplicity we shall only consider ROPs. As in the general case, if the matrix M is not strongly regular, we have to compute the ROPs in blocks. However, the blocks that will appear here in a natural way are *not* the same as the blocks defined in the general case. In the Toeplitz case, we shall group the ROPs in blocks that correspond naturally to the way they are recursively computed. The blocks for the LOPs and the blocks for the ROPs need not be the same. The block structure corresponding to the dimensions of the regular leading submatrices of M is a refining of the block structure of both the ROPs and the LOPs. By this we mean that the block indices of the blocks of ROPs (LOPs) are a subset of the block indices of M. Indeed, it follows from the definition of the Iohvidov indices below that the submatrices $M_n = M_{\nu,\nu}$ ($\nu=\nu_{n+1}$) are regular. However, it may happen that intermediate submatrices $M_{\nu,\nu}$ with $\nu_n < \nu < \nu_{n+1}$ are regular too. This situation is different from the general case of section 2 wich shows that the blocks need not be the same.

Because in this section we shall always refer to blocks of ROPs, we shall use the same terminology (block index, block size, ...) and the same notation ($\nu_k, \alpha_k, \ldots$) since confusion is very unlikely.

Definition 4.1 (blocks, Iohvidov indices) *The block indices are defined as* $\nu_n = \sum_{k=1}^{n} \alpha_k$, $\nu_0 = 0$. *The block sizes* α_k *are obtained as* $\alpha_{n+1} = \alpha_{n+1}^{+} + \alpha_{n+1}^{-} + 1$. *The Iohvidov indices* $\alpha_{n+1}^{\pm}$ *of block* n *are defined by*

$$\begin{aligned} \alpha_{n+1}^{-} &= \min\{\alpha \in \mathbf{N} : (1, z^{\alpha+1} a_{\nu+1}) \neq 0\} \\ \alpha_{n+1}^{+} &= \min\{\alpha \in \mathbf{N} : (z^{\nu+\alpha+1}, a_{\nu+1}) \neq 0\} \end{aligned}$$

where $\nu + 1 = \nu_n$.

To describe the blocks of ROPs we shall exclude some trivial situations first. First exclude the case where $\alpha_1^{+} = \infty$ since this means that M is strictly upper triangular and then there do not exist (monic) orthogonal polynomials unless $\alpha_1^{-} = \infty$. The case $\alpha_1^{-} = \infty$ is also trivial, since then there is an infinite block of orthogonal polynomials which are just the powers of z.

Hence we suppose $\alpha_1^{+} + \alpha_1^{-} < \infty$. Then block number 0 contains $\nu + 1 = \nu_1 = \alpha_1^{+} + \alpha_1^{-} + 1$ polynomials $a_k = z^k$, $k = 0, 1, \ldots, \nu$. The first monic polynomial of block number 1, i.e., $a_{\nu+1} = a_1^1$ has a stacking vector $\mathbf{a}_1^1 = [\underline{\mathbf{a}}_1^{1T} \;\; 1]^T$ that can be found by solving the system

$$M_0 \underline{\mathbf{a}}_1^1 = -[\mu_{\nu+1} \cdots \mu_1]^T, \quad M_0 = M_{\nu_1, \nu_1}.$$

Note that the matrix of this system is regular and has a band of zeros of width $\alpha_1^{+} + \alpha_1^{-}$ (which is 0 in the regular case). Hence this system can be solved very efficiently.

Now for the general situation of block n. Since $M_{n-1} = M_{\nu,\nu}$ ($\nu = \nu_n$) is regular, it follows that the first monic ROP of block n, i.e., a_n^1, will be uniquely defined. For the construction of the remaining ones, we introduce the following notation:

$$\tilde{a}_n(z) = z^{\alpha_{n+1}^{-}+1} a_n^1(z).$$

By definition of α_{n+1}^{-}, we know that the polynomials $a_{\nu_n+k}(z) = z^k a_n^1(z)$ are orthogonal to all the previous ones for $k = 0, 1, \ldots, \alpha_{n+1}^{-}$. This is a consequence of the Toeplitz structure of M which implies $(zf(z), zg(z)) = (f(z), g(z))$. The polynomial $\tilde{a}_n$ is the first one in block n (except for a_n^1 of course) which is obtained by shifting the previous one and is not orthogonal to all polynomials of lower degree.

Before treating the generic case where both Iohvidov indices are finite, we shall deal with the cases of infinite α^{+} and/or α^{-} first.

Theorem 4.1 *Let* M *be a Toeplitz matrix and suppose* $\nu + 1 = \nu_n$ *is a block index. Hence* $M_{\nu,\nu}$ *is regular and the (monic) regular orthogonal polynomial* $a_{\nu+1}$ *is uniquely defined. Let the Iohvidov index for that block* α_{n+1}^{-} *be infinite. Then* $a_{\nu+k+1}(z) = z^k a_{\nu+1}(z)$, $k = 0, 1, \ldots$ *is a set of ROPs.*

If moreover, the other Iohvidov index $\alpha_{n+1}^{+} > 0$, *then there is no leading submatrix* $M_{k,k}$ *that is regular for some* $k > \nu$, *i.e., the block size* $\alpha_{n+1} = \infty$.

Proof. The right orthogonality follows immediately from the definition of α_{n+1}^{-}.

When we express the fact that $[a_0, a_1, \ldots, a_\nu | a_{\nu+1}, z a_{\nu+1}, \ldots]$ is right orthogonal to $[1, z, z^2, \ldots]$, in a matrix form (A is unit upper triangular and contains the stacking vectors) we see that $MA = R$ with R a matrix of the form

$$R = \begin{bmatrix} R_{00} & 0 \\ R_{10} & R_{11} \end{bmatrix}$$

with R_{00} regular of size $\nu+1$ and R_{11} infinite lower triangular Toeplitz. The diagonal element of R_{11} is zero if $\alpha^+_{n+1} > 0$. Thus it follows from $MA = R$ with A unit upper triangular that all leading submatrices M_{kk} are singular for $k > \nu$. □

The previous theorem treated the case where $\alpha^-_{n+1} = \infty$. The next one discusses $\alpha^+_{n+1} = \infty$.

Theorem 4.2 *Let M be Toeplitz, with block index $\nu + 1 = \nu_n$ and corresponding Iohvidov indices $\alpha^+_{n+1} = \infty$ and $\alpha^-_{n+1} < \infty$, then all $M_{k,k}$ are singular for $k > \nu$.*

An infinite block of ROPs can be computed as follows. Set $a_{\nu+k+1} = z^k a_{\nu+1}$ for $k = 0, 1, \ldots, \alpha^-_{n+1}$. Shifting the last of these once more will give $\tilde{a}_n$. Setting $\mu = \nu + \alpha^-_{n+1} + 2$, we can now find a nonzero constant c_n and monic polynomials d_j such that

$$a_{\mu+j}(z) = z^j \tilde{a}_{n-1}(z) c_n + \tilde{a}_n(z) d_j(z), \quad j = 0, 1, \ldots$$

Proof. With $\mathsf{a}_{\nu+1}$ the stacking vector of $a_{\nu+1}$, it follows from the definition of α^+_{n+1}, which is infinity, that $M\mathsf{a}_{\nu+1} = 0$. Hence, using again the unit upper triangular matrix A whose first $\nu+1$ columns are the stacking vectors of $a_0, \ldots, a_{\nu+1}$ and whose remaining part corresponds to columns of the unit matrix, then the singularity of $M_{k,k}$ for $k > \nu$ follows immediately from $MA = R$, since R contains a zero column ($M\mathsf{a}_{\nu+1}$).

The first α^-_{n+1} ROPs are allright by definition of α^-_{n+1}.

The remaining polynomials can be obtained as indicated. For $j = 0$, we know that $(z^k, \tilde{a}_n) = 0$ for $k = 1, 2, \ldots$, but it is nonzero for $k = 0$. A similar observation holds for $\tilde{a}_{n-1}$ namely $(z^k, \tilde{a}_{n-1}) = 0$ for $k = 1, \ldots, \nu_n - 1$ and nonzero for $k = 0$. Thus the constant $c_n \neq 0$ exists so that $a_\mu = \tilde{a}_{n-1} c_n + \tilde{a}_n$ is right orthogonal to z^k for $k = 0, 1, \ldots, \nu_n - 1$

For $j > 0$, suppose that $a_{\mu+i}$ for $i = 0, \ldots, j-1$ have been computed as ROPs, orthogonal to z^k, $k = 0, 1, \ldots, \nu_n - 1$. Hence $z^j a_\mu$ is right orthogonal to z^k, $k = j, \ldots, \nu_n + j - 1$. By adding multiples of $z^i \tilde{a}_n$, $i = 0, \ldots, j-1$, we can satisfy the orthogonality relation.

The polynomials $a_{\mu+j}$ thus obtained are monic and of the correct degree. This follows from $\deg \tilde{a}_{n-1} < \deg \tilde{a}_n$. The latter inequality follows from the definitions of ν_n and α_n and the Iohvidov indices being nonnegative. □

Now we are ready to treat the generic situation in the following theorem.

Theorem 4.3 *Let the moment matrix M be Toeplitz and let $\nu = \nu_n$ be a block index. Let the corresponding Iohvidov indices α^+_{n+1} and α^-_{n+1} be finite. Then the $\alpha_{n+1} = \alpha^+_{n+1} + \alpha^-_{n+1} + 1$ orthogonal polynomials for block n can be obtained as follows*

$$a_{\nu+i}(z) = z^i a_\nu(z), \text{ for } i = 0, \ldots, \alpha^-_{n+1} \tag{4.1}$$

and for $\mu = \nu + \alpha^-_{n+1} + 1 = \deg \tilde{a}_n$ (with $\tilde{a}_n$ as defined in the previous theorem)

$$a_{\mu+j}(z) = z^j \tilde{a}_{n-1}(z) c_n + \tilde{a}_n(z) d_j(z), \quad j = 0, \ldots, \alpha^+_{n+1} - 1 \tag{4.2}$$

with $c_n \neq 0$, a constant, and d_j a monic polynomial of degree j.

The first polynomial of the next block satisfies

$$a^1_{n+1}(z) = z^{\alpha^+_{n+1}} \tilde{a}_{n-1}(z) c_n + a^1_n(z) q_n(z) \tag{4.3}$$

where again $c_n \neq 0$ is a constant and q_n is a monic polynomial of degree α_{n+1}.

Proof. The choice of the first $\alpha_{n+1}^- + 1$ orthogonal polynomials is again justified by the definition of the Iohvidov index α_{n+1}^-.

For the remaining α_{n+1}^+ polynomials, take $j = 0$ first. We know that $(z^k, \tilde{a}_n) = 0$ for $k = 1, \ldots, \nu_{n+1} - 1$ but it is nonzero for $k = 0$ and similarly, $(z^k, \tilde{a}_{n-1}) \neq 0$ for $k = 0$ and zero for $k = 1, \ldots, \nu_n - 1$. Hence there exists a constant c_n such that

$$a_\mu = \tilde{a}_{n-1} c_n + \tilde{a}_n$$

is orthogonal to all z^k for $k = 0, 1, \ldots, \nu_n - 1$. If more generally, $j > 0$, then suppose $z^j a_\mu$ is orthogonal to all z^k for $k = j, \ldots, \nu_n + j - 1$. By adding multiples of $z^i \tilde{a}_n(z)$, $i = 0, 1, \ldots, j-1$, we can satisfy the remaining orthogonality conditions. This proves (4.2).

For the computation of a_{n+1}^1, the first polynomial of block $n+1$, we observe that $z^{\alpha_{n+1}} a_{n-1}^1(z) = z^{\alpha_{n+1}^+} \tilde{a}_{n-1}(z)$ is orthogonal to z^k for $\alpha_{n+1}^+ < k \leq \nu_n - 1 + \alpha_{n+1}^+$. By adding multiples of $z^i \tilde{a}_n$, for $i = 0, 1, \ldots, \alpha_{n+1}^+$, we can construct a polynomial

$$\begin{aligned} \hat{a}(z) &= z^{\alpha_{n+1}^+} \tilde{a}_{n-1}(z) c_n + \tilde{a}_n(z) r(z) \\ &= z^{\alpha_{n+1}^+} \tilde{a}_{n-1}(z) c_n + z^{\alpha_{n+1}^- + 1} a_n^1(z) r(z) \end{aligned} \tag{4.4}$$

which is orthogonal to z^k for $0 \leq k \leq \nu_n + \alpha_{n+1}^+ - 1$ and where c_n is a nonzero constant and $r(z)$ a monic polynomial of degree α_{n+1}^+. The remaining α_{n+1}^- orthogonality conditions can be satisfied by adding multiples of $z^l a_n^1(z)$ for $l = 0, 1, \ldots, \alpha_{n+1}^-$ which are indeed orthogonal to z^k for $0 \leq k \leq \nu_n + \alpha_{n+1}^+ + l - 1$. This then leads to the form proposed in the theorem: $a_{n+1}^1(z) = \hat{a}(z) + s(z) a_n^1(z)$ with $s(z)$ of degree α_{n+1}^- and where $q_n(z) = s(z) + z^{\alpha_{n+1}^- + 1} r(z)$. □

Note that the polynomials d_j can be simply obtained by the recursion

$$d_j(z) = z d_{j-1}(z) + d_j(0).$$

The recurrence relation that we have just derived for the block orthogonal polynomials gives rise just like in the general case to a relation of the form (2.6), i.e.,

$$F_n A_n = A_n T_n$$

but unlike in the Hankel case, the upper Hessenberg matrix T_n does not simplify considerably. Although it has a special structure, it is complicated to describe. There is however another formulation, which does give simple matrices. We should then replace the previous relation by

$$F_n A_n U_n = A_n S_n$$

where we have replaced T_n by $S_n U_n^{-1}$. The matrix U_n is a simple upper block bidiagonal matrix and S_n is a unit upper Hessenberg matrix which is block bidiagonal too with a very simple structure. Why is it that we have this extra complication with the expression $F_n A_n = A_n T_n$? In the left-hand side, F_n acts like a shift operator, i.e., a_k in the kth column of A_n is replaced by $z a_k$. The right-hand side expresses these polynomials in terms of the a_j. However, by our recurrence relation, we have expressed this in terms of $\tilde{a}_k$-polynomials too. The latter do not correspond to columns of the A_n matrix. Thus we have to express them explicitly in terms of a_k's and this makes T_n so complicated. But, the $\tilde{a}_n$ polynomial is a shifted version of one of the a_k and it is therefore natural to bring that part of the recurrence

which refers to one of these $\tilde{a}_k$ to the left-hand side, where these shifted polynomials are indeed available. This is the philosophy of the next manipulations of the formulas. Basically, this decomposition corresponds to bringing the part of the recursion that corresponds to the d_j part to the left hand side.

As before, we use the notations $\nu + 1 = \nu_n$ and we shall give an index to μ: $\mu_n = \nu_n + \alpha^-_{n+1} + 1$. We now rewrite the different updating schemes.

The expression (4.1) can be reformulated as

$$z a_{\nu_n+i}(z) = a_{\nu_n+i+1} \text{ for } i = 0, \ldots, \alpha^-_{n+1} - 1. \tag{4.5}$$

The update (4.2) for $j = 0$ gives

$$z(a_{\mu_n-1} + a_{\mu_{n-1}-1} c_n) = a_{\mu_n} \tag{4.6}$$

(recall $d_0 = 1$ and $\tilde{a}_n = z^{\alpha^-_{n+1}+1} a^1_n$).
For $j > 0$ in (4.2), use $d_j(z) = z d_{j-1}(z) + d_j(0)$ to get

$$a_{\mu_n+j}(z) = z\left(z^{j-1}\tilde{a}_{n-1}(z) + \tilde{a}_n(z) d_{j-1}(z)\right) + \tilde{a}_n(z) d_j(0).$$

The expression between big brackets is a_{μ_n+j-1}. With the definition of $\tilde{a}_n$, we can bring this into the form

$$z\left(a_{\mu_n+j-1}(z) + a_{\mu_n-1} d_j(0)\right) = a_{\mu_n+j}(z) \tag{4.7}$$

Finally, the update (4.3) will be rearranged. First note that the polynomial r in (4.4) is in fact the polynomial $d_{\alpha^+_{n+1}}$ that one would get by continuing (4.2) one step further. So, for simplicity of notation, set $d_{\alpha^+_{n+1}-1}(z) = t(z)$, so that $r(z) = z t(z) + r(0)$. We then get ($q_n(z) = s(z) + z^{\alpha^-_{n+a}+1} r(z)$)

$$\begin{aligned} a^1_{n+1}(z) &= z^{\alpha^+_{n+1}} \tilde{a}_{n-1} c_n + a^1_n z^{\alpha^-_{n+1}+1} r + a^1_n s \\ &= z\left(z^{\alpha^+_{n+1}-1}\tilde{a}_{n-1}(z) c_n + \tilde{a}_n t(z)\right) + \tilde{a}_n(z) r(0) + a^1_n(z) s(z) \\ &= z\left(a_{\mu_n+\alpha^+_{n+1}-1}(z) + a_{\mu_n-1}(z) r(0)\right) + a^1_n(z) s(z) \end{aligned}$$

($\tilde{a}_n = z a_{\mu_n-1}$). This can be rearranged into the form

$$z\left(a_{\nu_{n+1}-1}(z) + a_{\mu_n-1}(z) r(0)\right) = a^1_{n+1}(z) - a^1_n(z) s(z). \tag{4.8}$$

Collecting now the relations (4.5—4.8) into one matrix relation, we get

$$F_n A_n U_n = A_n S_n.$$

$F_n = F(a_{0,n+1})$ is still the companion matrix for the (monic) polynomial a^1_{n+1} of degree ν_{n+1} and A_n is the unit upper triangular matrix whose columns are the stacking vectors of the polynomials a_k for $k = 0, \ldots, \nu_{n+1} - 1$. The unit upper Hessenberg matrix T_n is decomposed as $T_n = S_n U_n^{-1}$ where now S_n is a simplified unit upper Hessenberg and U_n is unit upper triangular. Puzzling the pieces together will reveal that U_n is block bi-diagonal

$$U_n = \begin{bmatrix} U^{00}_n & U^{01}_n & & & \\ & U^{11}_n & U^{12}_n & & \\ & & \ddots & \ddots & \\ & & & \ddots & U^{n-1,n}_n \\ & & & & U^{nn}_n \end{bmatrix}$$

with $U_n^{n-1,n}$ a typical off diagonal block of size $\alpha_n \times \alpha_{n+1}$ having the form

$$U_n^{n-1,n} = \begin{array}{c} \\ \alpha_n^- \\ 1 \\ \alpha_n^+ \end{array} \begin{array}{c} \begin{array}{ccc} \alpha_{n+1}^- & 1 & \alpha_{n+1}^+ \end{array} \\ \begin{bmatrix} 0 & 0 & 0 \\ 0 & c_n & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

and U_n^{nn}, a typical diagonal block of size $\alpha_{n+1} \times \alpha_{n+1}$ of the form

$$U_n^{nn} = \begin{array}{c} \\ \alpha_{n+1}^- \\ 1 \\ \alpha_{n+1}^+ \end{array} \begin{array}{c} \begin{array}{ccc} \alpha_{n+1}^- & 1 & \alpha_{n+1}^+ \end{array} \\ \begin{bmatrix} I & 0 & 0 \\ 0 & 1 & \mathbf{r}^T \\ 0 & 0 & I \end{bmatrix} \end{array}$$

where $\mathbf{r}$ is the stacking vector of the polynomial $r(z)$. As for the matrix S_n, this is also block bi-diagonal

$$S_n = \begin{bmatrix} S_n^{00} & & & \\ S_n^{10} & S_n^{11} & & \\ & \ddots & \ddots & \\ & & S_n^{n,n-1} & S_n^{nn} \end{bmatrix}$$

with a typical block S_n^{nn} of size $\alpha_{n+1} \times \alpha_{n+1}$ of the form

$$S_n^{nn} = \begin{array}{c} \\ 1 \\ \alpha_{n+1}^- \\ \alpha_{n+1}^+ \end{array} \begin{array}{c} \begin{array}{ccc} \alpha_{n+1}^- & \alpha_{n+1}^+ & 1 \end{array} \\ \begin{bmatrix} 0 & 0 & -s_0 \\ & & -s_1 \\ I & 0 & \vdots \\ & & -s_{\alpha_{n+1}^-} \\ 0 & I & 0 \end{bmatrix} \end{array}$$

which is in fact the companion matrix for the polynomial $z^{\alpha_{n+1}^-} + s(z)$.

The subdiagonal blocks are zero everywhere except for the right top element which is 1.

We can now derive the following determinant formula for a_{n+1}^1:

$$\begin{aligned} a_{n+1}^1(z) &= \det(zI - F_n) \\ &= \det(zI - A_n S_n U_n^{-1} A_n^{-1}) \\ &= \det(zI - S_n U_n^{-1}) \\ &= \det(zU_n - S_n) \end{aligned}$$

because U_n and A_n are unit upper triangular. The latter relation gives a determinant expression in terms of the simple matrices U_n and S_n.

The rational approximants that can be obtained for the Toeplitz case are described as follows. Define

$$\begin{aligned} g^+(z) &= \mu_0 + \mu_{-1} z + \mu_{-2} z^2 + \cdots \\ -g^-(z) &= \mu_1 z^{-1} + \mu_2 z^{-2} + \cdots \end{aligned}$$

Let $\nu+1=\nu_n$, then from $M_{\nu_n,\nu_n+1}\mathbf{a}_n^1=[0,\ldots,0]^T$, we know that

$$\mathbf{c}_n=\begin{bmatrix}\mu_0 & 0 & & 0\\ \mu_{-1} & \mu_0 & & \\ & & \ddots & \vdots\\ \mu_{-\nu} & & \cdots\ \mu_0 & 0\end{bmatrix}\mathbf{a}_n^1=-\begin{bmatrix}0 & \mu_1 & \mu_2 & \cdots & \mu_{\nu+1}\\ & 0 & \mu_1 & & \\ & & \ddots & & \\ & & & 0 & \mu_1\end{bmatrix}\mathbf{a}_n^1$$

is the stacking vector of a polynomial of degree at most ν which is obtained as the polynomial part of $g^-(z)a_{n+1}^1(z)$. Hence

$$\begin{aligned}g^-(z)a_n^1(z)-c_n(z) &= *z^{-1-\alpha_{n+1}^-}+\text{lower order terms}\\ g^+(z)a_n^1(z)-c_n(z) &= *z^{\nu+1+\alpha_{n+1}^+}+\text{higher order terms.}\end{aligned}$$

We can easily check by induction that $a_n^1(0)\neq 0$, so that

$$\begin{aligned}g^-(z)-c_n(z)/a_n^1(z) &= *z^{-\nu-2-\alpha_{n+1}^-}+\text{lower order terms}\\ g^+(z)-c_n(z)/a_n^1(z) &= *z^{\nu+1+\alpha_{n+1}^+}+\text{higher order terms.}\end{aligned}$$

The rational approximant $g_n=c_n/a_n^1$ has a total of $2(\nu+1)=2\nu_n$ parameters and fits in the series $g^\pm$ together a total of $2\nu_n+\alpha_{n+1}-1$ coefficients. It can therefore be considered as a two point Padé approximant for the two series $g^\pm$.

References

[1] C. Brezinski. *Padé-type approximants and general orthogonal polynomials*, volume 50 of *ISNM*. Birkhäuser Verlag, Basel, 1980.

[2] A. Bultheel. *Laurent series and their Padé approximations*, volume OT 27 of *Operator Theory: Advances and applications*. Birkhäuser Verlag, Basel, Boston, 1987.

[3] A. Draux. *Polynômes orthogonaux formels – applications*, volume 974 of *Lecture Notes Mathematics*. Springer-Verlag, Berlin, 1983.

[4] W.B. Gragg. The Padé table and its relation to certain algorithms of numerical analysis. *SIAM Rev.*, 14:1–62, 1972.

[5] W.B. Gragg. Matrix interpretations and applications of the continued fraction algorithm. *Rocky Mountain J. Math.*, 4(2):213–225, 1974.

[6] W.B. Gragg. Positive definite Toeplitz matrices, the Arnoldi process for isometric operators, and Gaussian quadrature on the unit circle. In E.S. Nikolaev, editor, *Numerical methods in linear algebra*, pages 16–32, Moscow, 1982. Moscow University Press. (in Russian).