A Geronimus transformation for matrix orthogonal polynomials

Juan Carlos García-Ardila,* Luis E. Garza,† and Francisco Marcellán ‡

Abstract

We consider matrix polynomials orthogonal with respect to a bilinear form $[\cdot,\cdot]_W$ such that

$$[P(t)W(t), Q(t)W(t)]_W = \int_{\Im} P(t) d\mu Q(t)^T, \quad P, Q \in \mathbb{P}^{p\times p}(\mathbb{R}),$$

where μ is a symmetric, positive definite matrix of measures supported in some infinite subset $\Im$ of the real line, and $W(t)$ is a matrix polynomial of degree 1. We obtain a connection formula between the sequences of matrix polynomials orthogonal with respect to $[\cdot,\cdot]_W$ and μ, as well as a relation between the corresponding block Jacobi matrices. A non-symmetric bilinear form is also considered.

2010 Mathematics Subject Classification: Primary 42C05; Secondary 15A23.

Keywords and Phrases: Matrix orthogonal polynomials, Block Jacobi matrices, Matrix Geronimus transformation, Block Cholesky decomposition, Block LU decomposition.

1 Introduction

Let us denote by $\mathbb{C}$ and $\mathbb{R}$ the sets of complex and real numbers, respectively. Let p and q be positive integers. $\mathbb{C}^{p\times q}$ will denote the set of $p \times q$ matrices with complex entries. If $A \in \mathbb{C}^{p\times q}$, then we denote by A^T and A^* the transpose and conjugate transpose of A, respectively.

Definition 1. *Let $A_0, A_1 \cdots, A_N \in \mathbb{R}^{p\times p}$ and let us assume that A_N is nonsingular. Then*

$$P_N(t) = A_N t^N + A_{N-1} t^{N-1} + \cdots + A_1 t + A_0$$

is said to be a matrix polynomial of degree N. The matrix polynomial is said to be monic if $A_N = I_p$, the $p \times p$ identity matrix. The set of matrix polynomials with coefficients in $\mathbb{R}^{p\times p}$ is denoted by $\mathbb{P}^{p\times p}(\mathbb{R})$.

*Departamento de Matemáticas, Universidad Carlos III de Madrid, Avenida Universidad 30, Leganés, Spain jugarcia@math.uc3m.es

†Facultad de Ciencias, Universidad de Colima, Bernal Díaz del Castillo 340, Colima garzaleg@gmail.com

‡Instituto de Ciencias Matemáticas (ICMAT) and Departamento de Matemáticas, Universidad Carlos III de Madrid, Avenida Universidad 30, Leganés, Spain pacomarc@ing.uc3m.es

Definition 2. *We say that* $t_0 \in \mathbb{C}$ *is a zero of a matrix polynomial* $P_N(t)$ *if* $\det[P_N(t_0)] = 0$. *The set of the zeros of* $P_N(t)$ *will be denoted* $\mathcal{Z}(P_N)$.

Clearly, from the above definition, $P_N(t)$ has at most Np zeros. Furthermore, if there exist matrices Z_i, $i = 1 \ldots N$, such that $P_N(t)$ can be written as (this is not always possible)

$$P_N(t) = (tI_p - Z_1)(tI_p - Z_2) \cdots (tI_p - Z_N),$$

then the spectrum of any Z_i, $i = 1 \ldots N$, belongs to the set of zeros of $P_N(t)$, and we have

$$\cup_{i=1}^{N} \sigma(Z_i) = \mathcal{Z}(P_N).$$

Definition 3. *Given the monic matrix polynomial* $P_N(t) = t^N + A_{N-1}t^{N-1} + \cdots + A_1 t + A_0$ *we define the operator* $P_N : \mathbb{C}^{p\times p} \longrightarrow \mathbb{C}^{p\times p}$ *by*

$$P_N(Z) = Z^N + A_{N-1}Z^{N-1} + \cdots + A_1 Z + A_0.$$

We say $P_N(Z)$ *is the evaluation of the polynomial* P_N *at the matrix* Z.

Remark 1. *Notice that the order of the factors is very important because of the non-commutativity for the product of matrices. An important fact is the following (see [18]).* Z_1 *satisfies* $P_N(Z_1) = 0_{p\times p}$ *if and only if there exists a monic matrix polynomial* $q(t)$ *of degree* $N-1$ *such that* $P_N(t) = q(t)(tI_p - Z_1)$. *Nevertheless, if* $P_N(t) = (tI_p - Z_2)q(t)$, *then in general it is not true that* $P_N(Z_2) = 0_{p\times p}$, *as is shown in the following example. Take*

$$P_2(t) = (tI_2 - Z_2)(tI_2 - Z_1) = I_2 t^2 - (Z_2 + Z_1)t + Z_2 Z_1$$

with

$$Z_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \qquad Z_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Then,

$$P_2(Z) = Z^2 - (Z_2 + Z_1)Z + Z_2 Z_1,$$

and by direct computation we get

$$P_2(Z_1) = 0_{p\times p} \quad \textit{but} \quad P_2(Z_2) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The above example shows that in general is not true that if $P_N(t)$ *can be written as* $P_N(t) = (tI_p - Z_1) \cdots (tI_p - Z_N)$ *then* $P_N(Z) = (Z - Z_1) \cdots (Z - Z_N)$.

Let $\mu = (u_{i,j})_{i,j=0}^{p-1}$ be a positive definite Hermitian matrix of measures supported in $\Im \subset \mathbb{R}$ and let us introduce the inner product

$$\langle P(t), Q(t) \rangle_L = \int_{\Im} P(t) d\mu Q^T(t), \quad P(t), Q(t) \in \mathbb{P}^{p\times p}(\mathbb{R}). \tag{1}$$

This is known in the literature (see [22, 23]) as left inner product. Alternatively, we can define the right inner product

$$\langle P(t), Q(t)\rangle_R = \int_{\mathfrak{I}} P^T(t)d\mu Q(t),$$

but since $\langle R, Q\rangle_L = \left\langle R^T, Q^T\right\rangle_R$ for all $R, Q \in \mathbb{P}^{p\times p}(\mathbb{R})$, it suffices to study the left inner product. The following proposition follows easily from the definition.

Proposition 2. *Let $\langle\cdot,\cdot\rangle_L$ be defined as in* (1)*. Then,*

i) $\langle P(t), Q(t)\rangle_L^T = \langle Q(t), P(t)\rangle_L$, *for all* $P(t), Q(t) \in \mathbb{P}^{p\times p}(\mathbb{R})$.

ii) $\langle C_1P_1(t) + C_2P_2(t), Q(t)\rangle_L = C_1\langle P_1(t), Q(t)\rangle_L + C_2\langle P_2(t), Q(t)\rangle_L$, *for* $C_1, C_2 \in \mathbb{R}^{p\times p}$ *and* $Q(t), P_1(t), P_2(t) \in \mathbb{P}^{p\times p}(\mathbb{R})$.

iii) *For $P(t) \in \mathbb{P}^{p\times p}(\mathbb{R})$, the matrix $\langle P(t), P(t)\rangle_L$ is always positive semi-definite and it is positive definite if $det[P(t)] \neq 0$.*

iv) *$\langle P(t), P(t)\rangle_L = 0_p$ if only if $P(t) \equiv 0_p$, where 0_p is a $p \times p$ matrix with all entries equal to zero.*

By using a generalization of the Gram-Schmidt orthogonalization process for the basis $(t^nI_p)_{n\in\mathbb{N}}$ of $\mathbb{P}^{p\times p}(\mathbb{R})$, we can obtain a sequence of matrix polynomials $(P_n(t))_{n\in\mathbb{N}}$ such that the degree of $P_n(t)$ is n, its leading coefficient is a nonsingular matrix,

$$\langle P_n(t), P_m(t)\rangle_L = \delta_{n,m}S_n, \quad n, m \geqslant 0,$$

where $\delta_{n,m}$ is the Kronecker delta, and S_n is a positive definite $p \times p$ matrix for every $n \geqslant 0$. $(P_n(t))_{n\in\mathbb{N}}$ is said to be a sequence of matrix orthogonal polynomials associated with $\langle\cdot,\cdot\rangle_L$. Without loss of generality, we will assume that $(P_n(t))_{n\in\mathbb{N}}$ is a sequence of monic matrix orthogonal polynomials. The above implies immediately that such a sequence satisfies the three term recurrence relation [8, 23]

$$tP_n(t) = P_{n+1}(t) + B_nP_n(t) + A_nP_{n-1}(t), \quad n \geq 0,$$

with initial conditions $P_{-1}(t) = 0, P_0(t) = I_p$, and where A_n, $n \geq 1$, B_n, $n \geq 0$ are $p \times p$ matrices with A_n a nonsingular matrix and $B_n = B_n^T$. In matrix form, we have $tP = JP$, where $P = [P_0^T(t), P_1^T(t)\cdots]^T$ and

$$J = \begin{pmatrix} B_0 & I_p & & & \\ A_1 & B_1 & I_p & & \\ & A_2 & B_2 & I_p & \\ & & \ddots & \ddots & \ddots \end{pmatrix}.$$

J is called the (monic) block Jacobi matrix associated with the sequence $(P_n(t))_{n\in\mathbb{N}}$.

On the other hand, since the matrices S_n are positive definite, for each $n \in \mathbb{N}$ there exists a unique positive definite matrix K_n such that $S_n = K_n^2$, i.e. K_n is the square root of $\langle P_n(t), P_n(t)\rangle_L$ (see Theorem 7.2.6 in [15]). The sequence of matrix polynomials $(Q_n(t))_{n\in\mathbb{N}}$ defined by $Q_n(t) = K_n^{-1}P_n(t)$ is called a sequence of orthonormal matrix polynomials with respect to $\langle\cdot,\cdot\rangle_L$, since we have

$$\langle Q_n(t), Q_m(t)\rangle_L = \int_{\mathfrak{I}} Q_n d\mu Q_m^T = \delta_{n,m} I_p, \quad n, m \geqslant 0.$$

$(Q_n(t))_{n\in\mathbb{N}}$ satisfies the symmetric three term recurrence relation

$$tQ_n(t) = C_{n+1}Q_{n+1}(t) + E_nQ_n(t) + C_n^T Q_{n-1}(t). \quad n \geq 0, \quad Q_{-1}(t) = 0, \quad Q_0(t) = I_p, \tag{2}$$

where the matrices C_n, $n \geq 1$, are non singular and $E_n = E_n^T$, $n \geq 0$. Notice that the sequence of orthonormal matrix polynomials is not unique, since given a sequence of unitary matrices $(U_n)_{n\in\mathbb{N}}$, the sequence $(U_nQ_n(t))_{n\in\mathbb{N}}$ is also a sequence of orthonormal polynomials. As a consequence, the matrices C_n can be chosen as lower triangular (see [12]). Using spectral theory, A. I. Aptekarev and E. M. Nikishin showed in [1] that if a sequence of matrix polynomials $(Q_n(t))_{n\in\mathbb{N}}$ satisfies (2) then there exists a symmetric matrix of measures μ such that the sequence $(P_n(t))_{n\in\mathbb{N}}$ is orthonormal with respect to μ, i.e. an analogue for matrix polynomials of the classical Favard's theorem in the scalar case.

On the other hand, let $H = (m_{i,j})_{i,j=0}^{\infty}$, with

$$m_{i,j} = \left\langle t^i I_p, t^j I_p\right\rangle_L = \int_{\mathfrak{I}} t^{i+j} d\mu, \quad i, j \geqslant 0,$$

be the matrix of moments associated with μ, with respect to the canonical basis $(t^n I_p)_{n\in\mathbb{N}}$. Notice that H is a positive definite block Hankel matrix

$$H = \begin{pmatrix} m_{0,0} & m_{0,1} & \dots & m_{0,n} & \dots \\ m_{1,0} & m_{1,1} & \dots & m_{1,n} & \dots \\ \vdots & \vdots & \vdots & \vdots & \dots \\ m_{n,0} & m_{n,1} & \dots & m_{n,n} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

which can be written also as

$$H = \int_{\mathfrak{I}} \chi(t) d\mu \chi(t)^T,$$

where $\chi(t) = (I_p, tI_p, \dots)^T$. Since μ is positive definite, there exist a block lower triangular matrix T with blocks I_p in its diagonal and a block diagonal matrix $D = diag\{S_0, S_1, \dots\}$ such that $H = T^{-1}DT^{-T}$ (block Cholesky factorization) [2]. As a consequence, $P = T\chi(t)$ and the orthogonality of the sequence $(P_n(t))_{n\in\mathbb{N}}$ can also be expressed as

$$\int_{\mathfrak{I}} P d\mu P^T = T\left(\int_{\mathfrak{I}} \chi(t) d\mu \chi^T(t)\right) T^T = D.$$

The sequence $(P_n(t))_{n\in\mathbb{N}}$ can be computed in terms of the entries of the matrix H (see [11], [16]) by

$$\begin{array}{rcl} P_0(t) & = & I_p, \\ P_n(t) & = & (-Y_n^T H_{n-1}^{-1}, I_p)R_n(t)v_n, \quad n \geqslant 1, \end{array}$$

where $Y_n = (m_{0,n}, m_{1,n}, \ldots, m_{n-1,n})^T$, H_k is the $(k+1)p \times (k+1)p$ principal leading sub-matrix of H,

$$v_0 \quad := \quad I_p, \quad v_n := \begin{pmatrix} I_p \\ 0_{np\times p} \end{pmatrix} = \begin{pmatrix} v_{n-1} \\ 0_p \end{pmatrix}, \quad n \geq 1,$$

and $R_n : \mathbb{R} \to \mathbb{R}^{(n+1)p\times(n+1)p}$ is defined by

$$R_n(t) := (I_{(n+1)p} - tT_n)^{-1}, \qquad n \geq 0,$$

with

$$T_0 := 0_p, \quad T_n := \begin{pmatrix} 0_{p\times np} & 0_p \\ I_{np} & 0_{np\times p} \end{pmatrix}, \qquad n \geq 0.$$

Notice that, since $R_n(t)v_n = [I_p, tI_p, \ldots, t^n I_p]^T$, then (3) can be written as

$$P_n(t) = t^n I_p - Y_n^T H_{n-1}^{-1}[I_p, tI_p, \ldots, t^{n-1}I_p]^T.$$

Furthermore, if $(r_n(t))_{n\in\mathbb{N}}$ is another ordered basis of monic polynomials for $\mathbb{P}^{p\times p}(\mathbb{R})$ with associated moments $\tilde{\mu}_{i,j} = \left\langle r_i(t), r_j(t)\right\rangle_L$ and $\tilde{Y}_n^T$ and $\tilde{H}_{j-1}$ are defined as above, then it is easy to show that

$$P_n(t) = r_n(t) - \tilde{Y}_n^T \tilde{H}_{n-1}^{-1}[r_0(t), r_1(t), \ldots, r_{n-1}(t)]^T. \tag{3}$$

For convenience, in the sequel we will use the basis $(r_n(t))_{n\in\mathbb{N}}$ with $r_n(t) = (tI_p - A)^n$, i.e. we will define the new moments $(\mu_{i,j})_{i,j\in\mathbb{N}}$ associated with $d\mu$ by $\mu_{i,j} = \left\langle (tI_p - A)^i, (tI_p - A)^j\right\rangle_L$, for every $i, j \in \mathbb{N}$ and we will denote by H_A the corresponding block Hankel matrix.
More generally, a sesquilinear form $\langle \cdot, \cdot \rangle$ on the space $\mathbb{P}^{p\times p}(\mathbb{R})$ is a function

$$\langle \cdot, \cdot \rangle : \mathbb{P}^{p\times p}(\mathbb{R}) \times \mathbb{P}^{p\times p}(\mathbb{R}) \longrightarrow \mathbb{R}^{p\times p}$$

such that

i) $\langle AP(t), Q(t)\rangle = A\langle P(t), Q(t)\rangle$ for all $A \in \mathbb{R}^{p\times p}$.

ii) $\langle P(t), Q(t)\rangle = \langle Q(t), R(t)\rangle^T$.

iii) $\langle R(t) + P(t), Q(t)\rangle = \langle R(t), Q(t)\rangle + \langle P(t), Q(t)\rangle$.

$\langle\cdot,\cdot\rangle$ is said to be non degenerate if the leading principal sub-matrices of the corresponding Hankel matrix of moments with respect to any basis are nonsingular, and nontrivial if $\langle R(t), R(t)\rangle$ is a positive definite matrix for all $R(t) \in \mathbb{P}^{p\times p}(\mathbb{R})$ with nonsingular leading coefficient. Indeed, any non degenerate sesquilinear form will have a sequence of matrix monic orthogonal polynomials [2].

Definition 4. *Given a matrix bilinear form $[\cdot,\cdot]$, The sequences of polynomials $(P_n^{[1]}(t))_{n\in\mathbb{N}}$ and $(P_n^{[2]}(t))_{n\in\mathbb{N}}$ are said to be bi-orthogonal with respect to $[\cdot,\cdot]$ if*

$$[P_n^{[1]}(t), P_k^{[2]}(t)] = \delta_{n,k} K_n,$$

where $K_n \neq 0_{p\times p}$. Here, it is important to notice the order of the polynomials in the bilinear form i.e., if $n \neq k$ then $[P_n^{[2]}(t), P_k^{[1]}(t)]$ could be different from $0_{p\times p}$. In particular, if $[\cdot,\cdot]$ is a sesquilinear form, then $P_n^{[2]}(t) = B_n P_n^{[1]}(t)$, where B_n is a nonsingular matrix of size $p \times p$.

In the sequel, we will use quasi-determinants to obtain connection formulas between some families of orthogonal (resp. bi-orthogomal) polynomials. They constitute a generalization of the determinants when the entries of the matrix belong to a non-commutative field, and share several properties with them. In the simplest case of a 2×2 block matrix there are four quasi-determinants

$$\begin{vmatrix} \boxed{a_{1,1}} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1} - a_{1,2}a_{2,2}^{-1}a_{2,1}, \qquad \begin{vmatrix} a_{1,1} & \boxed{a_{1,2}} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,2} - a_{1,1}a_{2,1}^{-1}a_{2,2},$$
$$\begin{vmatrix} a_{1,1} & \boxed{a_{1,2}} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{2,1} - a_{2,2}a_{1,2}^{-1}a_{1,1}, \qquad \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & \boxed{a_{2,2}} \end{vmatrix} = a_{2,2} - a_{2,1}a_{1,1}^{-1}a_{1,2}.$$

Notice that on each case the quasi-determinant related to the boxed block is just the Schur complement of the opposite block. We will also use quasi-determinants for 3×3 block matrices. In this case, the Sylvester's theorem for quasi-determinants gives

$$\begin{vmatrix} \boxed{a_{1,1}} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} = \begin{vmatrix} \boxed{a_{1,1}} & a_{1,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} - \begin{vmatrix} \boxed{a_{1,2}} & a_{1,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} \begin{vmatrix} \boxed{a_{2,2}} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix}^{-1} \begin{vmatrix} \boxed{a_{2,1}} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{vmatrix}, \tag{4}$$

when the right side expression makes sense. For more information on quasi-determinants, we refer the reader to [13].

In the scalar case ($p = 1$), linear spectral transformations of measures supported on the real line have been studied in the literature in connection with integrable systems ([24], [27]). The most studied cases are the so-called canonical transformations: the Christoffel transformation consists in the multiplication of the measure by a polynomial of degree 1 (see [3, 6, 17, 21, 24, 27]), the Uvarov transformation consists in the addition of a Dirac delta measure (see [3, 17, 19, 27]) and Geronimus transformation consists in the division by a polynomial of degree 1 and the addition of a Dirac delta measure at the zero of the polynomial (see [9, 10, 14, 17, 20]). In [9] is proved that all multiple Geronimus transformations yield a simple Geronimus transformation for matrices of measures. The Uvarov transformation for the matrix case has been analyzed in [25]. More recently, the analysis of perturbations of a finite number of entries in the matrix of moments (which can be seen as a particular case of an Uvarov perturbation) was considered in [7].

In this contribution, we will consider a Geronimus transformation for matrix orthogonal polynomials. The structure of the manuscript is as follows. In Section 2, we define a symmetric bilinear form that represents a Geronimus transformation of a matrix of measures. We obtain a connection formula. A relation between the corresponding block Jacobi matrices is also deduced. In Section 3, we generalize the above results to the case of a non-symmetric bilinear form and obtain connection formulas for some bi-orthogonal families associated with such a transformation. Some final remarks and open problems are stated in Section 4.

2 A Geronimus transformation for symmetric matrix sesquilinear forms

Let $W(t) = tI_p - A$ be a matrix monic polynomial with zeros outside $\Im$ the support of μ. We define a sesquilinear form $[\cdot,\cdot]_W$ on the linear space $\mathbb{P}^{p\times p}(\mathbb{R})$ as follows

$$[P(t)(tI_p - A), Q(t)(tI_p - A)]_W = \int_{\Im} P(t)d\mu Q(t)^T, \quad P, Q \in \mathbb{P}^{p\times p}(\mathbb{R}). \tag{5}$$

Clearly, this definition does not determine uniquely the bilinear form $[\cdot,\cdot]_W$. Indeed, its corresponding moments $(\hat{\mu}_{i,j})_{i,j\in\mathbb{N}}$, given in terms of the basis $((tI_p - A)^n)_{n\in\mathbb{N}}$, are defined by $\hat{\mu}_{i,j} = [(tI_p - A)^i, (tI_p - A)^j]_W = \mu_{i-1,j-1}$ for $i, j \geqslant 1$, but the moments

$$[I_p, (tI_p - A)^k]_W = \hat{\mu}_{0,k} = \hat{\mu}_{k,0}^T = [(tI_p - A)^k, I_p]_W^T, \quad k = 0, 1 \ldots, \tag{6}$$

are arbitrary.

Furthermore, if $[\cdot,\cdot]_W$ is at least non degenerate, then there exists a sequence of monic matrix orthogonal polynomials $(\hat{P}_n)_{n\in\mathbb{N}}$ if and only if its corresponding block matrix of moments $\hat{H}_A = (\hat{\mu}_{i,j})_{i,j\in\mathbb{N}}$ has nonsingular leading principal submatrices. Moreover, we have

$$\hat{P}_n(t) = (tI_p - A)^n - \hat{Y}_n^T(\hat{H}_A)_{n-1}^{-1}[I_p, tI_p - A, \ldots, (tI_p - A)^{n-1}]^T, \quad n \geqslant 1,$$

where $\hat{Y}_n^T = [\hat{\mu}_{n,0}, \hat{\mu}_{n,1}, \ldots, \hat{\mu}_{n,n'1}] = [\hat{\mu}_{n,0}, \mu_{n-1,0}, \ldots, \mu_{n-1,n-2}]$. In such a case, $\hat{H}_A$ has the block Cholesky factorization $\hat{H}_A = \hat{T}_A^{-1}\hat{D}_A\hat{T}_A^{-T}$ (see [2]), so its associated sequence of monic orthogonal polynomials can also be obtained by $\hat{P} = \hat{T}_A\chi_A(t)$, where $\hat{P} = [\hat{P}_0, \hat{P}_1, \ldots]^T$ and $\chi_A(t) = [I_p, tI_p - A^T, (tI_p - A^T)^2, \ldots]^T$. In the next proposition, we provide an equivalent condition for the existence of $(\hat{P}_n)_{n\in\mathbb{N}}$.

Remark 3. *Let $\hat{H}$ be the moment matrix associated with $\langle\cdot,\cdot\rangle_W$, with respect to the canonical basis i.e $\hat{H} = \langle\chi(t), \chi(t)\rangle_W$. If $\hat{H}$ has a block Cholesky matrix factorization $\hat{H} = \hat{T}^{-1}\hat{D}\hat{T}^{-T}$, then* *the sequence of matrix monic orthogonal polynomials $(\hat{P}_n)_{n\in\mathbb{N}}$ can be written as $\hat{P} = \hat{T}\chi(t)$. Since $((tI_p - A)^n)_{n\in\mathbb{N}}$ is also a basis of $P^{p\times p}(\mathbb{R})$, then there exists a block lower triangular matrix L_A such that $\chi(t) = L_A\chi_A(t)$, and since the sequence of monic orthogonal polynomials is unique, then $\hat{T}_A L_A = \hat{T}$.*

Proposition 4. *Let $[\cdot,\cdot]_W$ be as defined in* (5)*. A necessary and sufficient condition for the existence of $(\hat{P}_n)_{n\in\mathbb{N}}$ is that the $p\times p$ matrices $R_0 = \hat{\mu}_{0,0}$ and*

$$R_n = \hat{\mu}_{0,0} - \sum_{k=0}^{n-1}[P_k(t)(tI_p - A), I_p]_W^T S_{k-1}^{-1}[P_k(t)(tI_p - A), I_p]_W, \quad n \geqslant 1, \tag{7}$$

are nonsingular.

Proof. First, notice that

$$(\hat{H}_A)_n = \left(\begin{array}{c|cccc} \hat{\mu}_{0,0} & \hat{\mu}_{0,1} & \cdots & \hat{\mu}_{0,n-1} & \hat{\mu}_{0,n} \\ \hline \hat{\mu}_{1,0} & & & & \\ \vdots & & (H_A)_{n-1} & & \\ \hat{\mu}_{n-1,0} & & & & \\ \hat{\mu}_{n,0} & & & & \end{array}\right).$$

As stated before, $(\hat{P}_n)_{n\in\mathbb{N}}$ exists if and only if the matrices $(\hat{H}_A)_n$, $n \geqslant 0$, are nonsingular. Let $R_n = |(\hat{H}_A)_n|$ be the quasi-determinant of $(\hat{H}_A)_n$ with respect to the block $\hat{\mu}_{0,0}$, which is a $p\times p$ matrix. Then, $\det(\hat{H}_A)_n \neq 0$ if and only if R_n is nonsingular. On the other hand,

$$\begin{aligned} R_n &= \left|\begin{array}{cccc|c} & & (\hat{H}_A)_{n-1} & & \hat{\mu}_{0,n} \\ & & & & \hat{\mu}_{1,n} \\ & & & & \vdots \\ & & & & \hat{\mu}_{n-1,n} \\ \hline \hat{\mu}_{n,0} & \hat{\mu}_{n,1} & \cdots & \hat{\mu}_{n,n-1} & \hat{\mu}_{n,n} \end{array}\right| \\ &= \left|\begin{array}{c|ccc|c} \boxed{\hat{\mu}_{0,0}} & \hat{\mu}_{0,1} & \cdots & \hat{\mu}_{0,n-1} & \hat{\mu}_{0,n} \\ \hline \hat{\mu}_{1,0} & & & & \mu_{0,n-1} \\ \vdots & & (H_A)_{n-2} & & \vdots \\ \hat{\mu}_{n-1,0} & & & & \mu_{n-2,n-1} \\ \hline \hat{\mu}_{n,0} & \mu_{n-1,0} & \cdots & \mu_{n-1,n-2} & \mu_{n-1,n-1} \end{array}\right| \\ &= \left|\begin{array}{c|c|ccc} \boxed{\hat{\mu}_{0,0}} & \hat{\mu}_{0,n} & \hat{\mu}_{0,1} & \cdots & \hat{\mu}_{0,n-1} \\ \hline \hat{\mu}_{n,0} & \mu_{n-1,n-1} & \mu_{n-1,0} & \cdots & \mu_{n-1,n-2} \\ \hline \hat{\mu}_{1,0} & \mu_{0,n-1} & & & \\ \vdots & \vdots & & (H_A)_{n-2} & \\ \hat{\mu}_{n-1,0} & \mu_{n-2,n-1} & & & \end{array}\right|, \end{aligned}$$

By using the Sylvester's identity (4) for quasi-determinants (see [13]), we get

$$R_n = R_{n-1} - |D_{n-1}||C_{n-1}|^{-1}|D_{n-1}^T|,$$

where

$$|C_{n-1}| = \begin{vmatrix} \boxed{\mu_{n-1,n-1}} & \mu_{n-1,0} & \cdots & \mu_{n-1,n-2} \\ \mu_{0,n-1} & & & \\ \vdots & & (H_A)_{n-2} & \\ \mu_{n-2,n-1} & & & \end{vmatrix} =: |(H_A)_{n-1}|,$$

and

$$|D_{n-1}| = \begin{vmatrix} \boxed{\hat{\mu}_{0,n}} & \hat{\mu}_{0,1} & \cdots & \hat{\mu}_{0,n-1} \\ \mu_{0,n-1} & & & \\ \vdots & & (H_A)_{n-2} & \\ \mu_{n-2,n-1} & & & \end{vmatrix}.$$

Taking into account (3), we have

$$S_{n-1} = \langle P_{n-1}, P_{n-1}\rangle_L = \mu_{n-1,n-1} - Y_{n-1}^T (H_A)_{n-2}^{-1} Y_{n-1},$$

the Schur complement of $(H_A)_{n-2}$ of $(H_A)_{n-1}$, and therefore $|C_{n-1}| = |(H_A)_{n-1}| = S_{n-1}$. Similarly, we get $|D_{n-1}| = [P_{n-1}(t)(tI_p - A), I_p]_W^T$. As a consequence,

$$R_n = R_{n-1} - [P_{n-1}(t)(tI_P - A), I_p]_W^T S_{n-1}^{-1} [P_{n-1}(t)(tI_p - A), I_p]_W.$$

Since $R_0 = \hat{\mu}_{0,0}$, (7) follows in a recursive way. □

Remark 5. *If $P_n(t) = \sum_{k=0}^n \gamma_{n,k}(tI_p - A)^k$, where $\gamma_{n,k}$ are $p \times p$ matrices for $0 \leqslant k \leqslant n$, then*

$$[P_{n-1}(t)(tI_p - A), I_p]_W = [\sum_{k=0}^{n-1} \gamma_{n-1,k}(tI_p - A)^{k+1}, I_p]_W = \sum_{k=0}^{n-1} \gamma_{n-1,k}\hat{\mu}_{k+1,0}.$$

In other words, the existence condition given in Proposition 4 can be stated in terms of $(P_n)_{n\in\mathbb{N}}$ and the moments $\hat{\mu}_{0,0}, \ldots \hat{\mu}_{n,0}$.

Notice that we can express every polynomial $P(t) \in \mathbb{P}^{p\times p}(\mathbb{R})$ of degree n as

$$P(t) = \sum_{k=0}^{n} a_k (tI_p - A)^k,$$

where a_k are $p \times p$ matrices. The following result provides a representation of $[\cdot,\cdot]_W$ in terms of the matrix of measures μ.

Proposition 6. *Let $W(t) = tI_p - A$ be a matrix polynomial, and let $\hat{\mu}$ be a matrix of measures supported in $\Im \subset \mathbb{R}$ defined by $W(t)d\hat{\mu}W(t)^T = d\mu$. Assume that the matrix moments associated with $\hat{\mu}$ are finite. Let $P(t) = \sum_{k=0}^{m} a_k(tI_p - A)^k$ and $Q(t) = \sum_{k'=0}^{m'} b_{k'}(tI_p - A)^{k'}$ be polynomials of degree m and m', respectively. Then, the following representation of $[\cdot,\cdot]_W$ holds*

$$[P(t), Q(t)]_W = \int_{\Im} Pd\hat{\mu}Q^T + \sum_{i=1}^{s} P^{(i)}(A)W_i[Q(A)]^T + \sum_{i=1}^{s} P(A)W_i^T[Q^{(i)}(A)]^T + P(A)W_0[Q(A)]^T, \quad (8)$$

where $s = max\{deg(P), deg(Q)\}$,

$$W_k = \frac{1}{k!}\left(\hat{\mu}_{k,0} - \int_{\Im}(tI_p - A)^k d\hat{\mu}\right), \quad 0 \le k \le s,$$

$(\hat{\mu}_{k,0})_{k\in\mathbb{N}}$ are defined as in (6) *and $P^{(i)}(A)$ denotes the i-th derivative of the polynomial $P(t)$ evaluated at the matrix A (see Definition 3).*

Proof. We have

$$\begin{aligned}
[P, Q]_W &= \left[\sum_{k=0}^{m} a_k(tI_p - A)^k, \sum_{k'=0}^{m'} b_{k'}(tI_p - A)^{k'}\right]_W \\
&= \sum_{k=1}^{m}\sum_{k'=1}^{m'}[a_k(tI_p - A)^k, b_{k'}(tI_p - A)^{k'}]_W + \sum_{k=1}^{m} a_k[(tI_p - A)^k, I]_W b_0^T \\
&+ \sum_{k'=1}^{m'} a_0[I, (tI_p - A)^{k'}]_W b_k^T + a_0[I, I]_W b_0^T \\
&= \int_{\Im} Pd\hat{\mu}Q^T + \sum_{k=1}^{m} a_k\left(\hat{\mu}_{k,0} - \int_{\Im}(tI_p - A)^k d\hat{\mu}\right) b_0^T \\
&+ \sum_{k'=1}^{m'} a_0\left(\hat{\mu}_{0,k'} - \int_{\Im} d\hat{\mu}(tI_p - A)^{k'T}\right) b_{k'}^T + a_0\left(\hat{\mu}_{0,0} - \int_{\Im} d\hat{\mu}\right) b_0^T.
\end{aligned}$$

On the other hand, since

$$P^{(s)}(t) = \sum_{k=s}^{m} \frac{k!}{(k-s)!} a_k(tI_p - A)^{k-s} \quad Q^{(s)}(t) = \sum_{k=s}^{m'} \frac{k'!}{(k'-s)!} b'_k(tI_p - A)^{k'-s},$$

then

$$\frac{1}{s!}P^{(s)}(A) = a_s \quad b_s = \frac{1}{s!}Q^{(s)}(A).$$

As a consequence,

$$
\begin{aligned}
[P,Q]_W &= \int_{\mathfrak{I}} Pd\hat{\mu}Q^T + \sum_{k=1}^{m} a_k\left(\hat{\mu}_{k,0} - \int_{\mathfrak{I}} (tI_p - A)^k d\hat{\mu}\right) b_0^T \\
&\quad + \sum_{k'=1}^{m'} a_0\left(\hat{\mu}_{0,k'} - \int_{\mathfrak{I}} d\hat{\mu}(tI_p - A)^{k'T}\right) b_{k'}^T + a_0\left(\hat{\mu}_{0,0} - \int_{\mathfrak{I}} Id\hat{\mu}\right) b_0^T \\
&= \int_{\mathfrak{I}} Pd\hat{\mu}Q^T + \sum_{i=1}^{m} P^{(i)}(A)W_i[Q(A)]^T + \sum_{i=1}^{m'} P(A)W_i^T[Q^{(i)}(A)]^T + P(A)W_0[Q(A)]^T.
\end{aligned}
$$

□

Remark 7. *In the scalar case $p = 1$, with $W(t) = t - a$, the sesquilinear form* (8) *becomes*

$$
[P,Q]_W = \int PQd\hat{\mu} + \sum_{i=1}^{s} W_i\left(P^{(i)}(a)Q(a) + P(a)Q^{(i)}(a)\right) + W_0P(a)Q(a), \tag{9}
$$

where P, Q are scalar polynomials and $a, W_i \in \mathbb{R}$. If we assume that $W_i = 0$, for $i \geq 2$, then $[\cdot,\cdot]_w$ satisfies

$$
[W^2(t)P(t), Q(t)]_W = [P(t), W^2(t)Q(t)]_W = \int P(t)d\mu Q(t), \tag{10}
$$

which can be written as

$$
[P,Q]_w = \int P(t)Q(t)d\hat{\mu} + \begin{pmatrix} P(a) & P'(a) \end{pmatrix} M \begin{pmatrix} Q(a) \\ Q'(a) \end{pmatrix}, \quad M = \begin{pmatrix} w_0 & w_1 \\ w_1 & 0 \end{pmatrix}. \tag{11}
$$

The symmetric bilinear form defined by (11) *was studied in [10] (see also [9]) where the entry* $(2,2)$ *of the matrix M is not necessarily equal to zero. Notice that in our case the entry* $(2,2)$ *of the matrix M is zero due to the extra condition $[W(t)P(t), W(t)Q(t)]_W = \int Pd\mu Q^T$.*

Let us denote by $(\hat{P}_n)_{n\in\mathbb{N}}$ the sequence of monic matrix polynomials orthogonal with respect to $[\cdot,\cdot]_W$. The relation between $(\hat{P}_n)_{n\in\mathbb{N}}$ and $(P_n(t))_{n\in\mathbb{N}}$ is stated in the following Proposition.

Proposition 8. *Assuming $[\cdot,\cdot]_W$ is nontrivial, the following connection formula holds*

$$
\hat{P}_{n+1}(t) = P_n(t)(tI_p - A) + \hat{P}_{n+1}(A)\left(I_p - \sum_{k=0}^{n-1} E_kP_k(t)(tI_p - A)\right), \quad n \geqslant 0, \tag{12}
$$

where $E_k = \left(\int_{\mathfrak{I}} d\hat{\mu}(tI_p - A)^T P_k^T(t) + \sum_{i=1}^{k+1} iW_i^T[P_k^{(i-1)}(A)]^T\right) S_k^{-1}$, with the convention $\sum_{k=0}^{-1} \equiv 0$.

Proof. First, notice that $\hat{P}_{n+1}(t)-\hat{P}_{n+1}(A)$ has the form $q_n(t)(tI_p-A)$ (see (3)), where q_n is a monic matrix polynomial of degree n. Therefore, we can write

$$\hat{P}_{n+1}(t)-\hat{P}_{n+1}(A)=P_n(t)(tI_p-A)+\sum_{k=0}^{n-1}\lambda_{n+1,k}P_k(t)(tI_p-A),\quad n\geqslant 1,$$

where

$$\begin{aligned}\lambda_{n+1,k}&=[\hat{P}_{n+1}(t)-\hat{P}_{n+1}(A),P_k(t)(tI_p-A)]_W S_k^{-1}\\&=-\hat{P}_{n+1}(A)[I_p,P_k(t)(tI_p-A)]_W S_k{-1}\\&=-\hat{P}_{n+1}(A)\left(\int_{\Im}d\hat{\mu}(tI_p-A)^T P_k^T(t)+\sum_{i=1}^{k+1}iW_i^T P_k^{(i-1)}(A)^T\right)S_k^{-1}.\end{aligned}$$

Thus, if

$$E_k=\left(\int_{\Im}d\hat{\mu}(tI_p-A)^T P_k^T(t)+\sum_{i=1}^{k+1}iW_i^T P_k^{(i-1)}(A)^T\right)S_k^{-1},$$

then (12) holds for $n\geqslant 1$. Furthermore, since $\hat{P}_1(t)-\hat{P}_1(A)=tI_p-A$, the above formula holds also for $n=0$ if we define $\sum_{k=0}^{-1}E_kP_k(t)=0_{p\times p}$. □

Remark 9. *By using* (3) *for the perturbed polynomials, we see that*

$$\hat{P}_n(A)=(\hat{\mu}_{n,0},\hat{\mu}_{1,n},\dots,\hat{\mu}_{n,n-1})(\hat{H}_A)_{n-1}^{-1}(I_p,0_p,\dots,0_p)^T,$$

i.e $\hat{P}_n(A)$ *can be obtained by multiplying* $(\hat{\mu}_{n,0},\hat{\mu}_{n-1,0},\dots,\mu_{n-1,n-2})$ *by the first column of* $(\hat{H}_A)_{n-1}^{-1}$.

Proposition 10. *The sequences* $(P_n(t))_{n\in\mathbb{N}}$ *and* $(\hat{P}_n(t))_{n\in\mathbb{N}}$ *satisfy the following inverse connection formula*

$$P(tI_p-A)=M\hat{P},\tag{13}$$

where $\hat{P}=[\hat{P}_0^T(t),\hat{P}_1^T(t),\dots]^T$ *and* M *is a block lower Hessenberg matrix with block entries*

$$\beta_{i,j}=\begin{cases}I_p, & \textit{if}\quad j=i+1,\\ \left[P_i(t)(tI_p-A),I_p\right]_W\hat{P}_j(A)^T\hat{S}_j^{-1}, & \textit{if}\quad 0\leq j\leq i,\\ 0_{p\times p}, & \textit{otherwise}.\end{cases}$$

Proof. Since $(\hat{P}_n(t))_{n\in\mathbb{N}}$ is a basis of $\mathbb{P}^{p\times p}(\mathbb{R})$, for every $n\in\mathbb{N}$ there exists a sequence of $p\times p$ matrices $(\beta_{n,k})_{k=1}^n$ such that

$$P_n(t)(tI_p-A)=\hat{P}_{n+1}(t)+\sum_{k=0}^{n}\beta_{n,k}\hat{P}_k(t),$$

with

$$\beta_{n,k}=[P_n(t)(tI_p-A),\hat{P}_k(t)]_W[\hat{P}_k,\hat{P}_k]_W^{-1},\quad k\leqslant n.$$

By denoting $\hat{S}_k = [\hat{P}_k, \hat{P}_k]_W$, from (12) we get

$$[P_n(t)(tI_p - A), \hat{P}_k(t)]_W = \left[P_n(t)(tI_p - A), P_{k-1}(tI_p - A) + \hat{P}_k(A)\left(I_p - \sum_{j=0}^{k-2} E_j P_j(t)(tI_p - A)\right)\right]_W$$
$$= \langle P_n(t), P_{k-1}(t)\rangle_L + \left([P_n(t)(tI_p - A), I_p]_W - \sum_{j=0}^{k-2} \left\langle P_n(t), P_j(t)\right\rangle_L E_j^T\right)\hat{P}_j(A)^T.$$

As a consequence, we obtain

$$\beta_{n,k} = [P_n(t)(tI_P - A), I_p]_W \hat{P}_k(A)^T \hat{S}_k^{-1}, \quad k \leqslant n.$$

□

Consider the block matrix $\hat{J}$ such that $t\hat{P} = \hat{J}\hat{P}$. Since we are not assuming that the multiplication operator by t is symmetric with respect to $[\cdot,\cdot]_W$, $\hat{J}$ is not necessarily a block Jacobi matrix. In general, we will say that $\hat{J}$ is the lower block Hessenberg matrix associated with $(\hat{P}_n(t))_{n\in\mathbb{N}}$. Its relation with the block Jacobi matrix J is deduced in the following result.

Proposition 11. *Let J and $\hat{J}$ be the monic block Jacobi and Hessenberg matrices associated with the sequences $(P_n(t))_{n\in\mathbb{N}}$ and $(\hat{P}_n(t))_{n\in\mathbb{N}}$, respectively. If $H = T^{-1}DT^{-T}$ and $\hat{H} = \hat{T}^{-1}\hat{D}\hat{T}^{-T}$ are the moment matrices associated with $\langle\cdot,\cdot\rangle_L$ and $[\cdot,\cdot]_W$, respectively, computed in terms of the canonical basis, then*

$$\hat{J} - \hat{T}\mathbb{D}_A\hat{T}^{-1} = LM,$$
$$J - T\mathbb{D}_A T^{-1} = ML,$$

where L is the block lower triangular matrix with I_p on the diagonal such that

$$\hat{P} = LP \tag{14}$$

and $\mathbb{D}_A = diag\{A, A, \ldots\}$.

Proof. By using (13) and (14), we obtain

$$\hat{P}(tI_p - A) = LP(tI_p - A) = LM\hat{P}.$$

On the other hand, taking into account that $\hat{P} = \hat{T}\chi(t)$ and $\chi(t)A = \mathbb{D}_A\chi(t)$, we also have

$$\hat{P}(tI_p - A) = \hat{P}t - \hat{P}A = t\hat{P} - \hat{T}\mathbb{D}_A\hat{T}^{-1}\hat{P} = (\hat{J} - \hat{T}\mathbb{D}_A\hat{T}^{-1})\hat{P}.$$

thus $LM = \hat{J} - \hat{T}\mathbb{D}_A\hat{T}^{-1}$. The other equation follows similarly.

□

3 A non-symmetric Geronimus transformation

In this section, we will consider a non-symmetric perturbation of a bilinear form defined as follows. Let $W(t)$ be as in the previous section, and define the bilinear form $[f,g]_W^L$ on the linear space $\mathbb{P}^{p\times p}(\mathbb{R})$ by

$$[P(t)(tI_p - A), Q(t)]_W^L = \langle P(t), Q(t)\rangle_L = \int_{\Im} P(t)d\mu Q(t)^T, \quad P, Q \in \mathbb{P}^{p\times p}(\mathbb{R}). \tag{15}$$

Then, the sequence of moments associated with $[f,g]_W$ in terms of the basis $((tI_p - A)^n)_{n\in\mathbb{N}}$ is

$$\tilde{\mu}_{i,j} = [(tI_p - A)^i, (tI_p - A)^j]_W^L = \left\langle (tI_p - A)^{i-1}, (tI_p - A)^j \right\rangle_L = \mu_{i-1,j}, \quad i \geqslant 1, j \geqslant 0.$$

As above, the bilinear form $[\cdot,\cdot]_W^L$ is not completely defined by (15), since the moments $\tilde{\mu}_{0,j} = [I, (tI_p - A)^j]_W^L$, $j \geqslant 0$, are arbitrary. In other words, $[\cdot,\cdot]_W^L$ is not unique. On the other hand, if H and $\tilde{H}$ are the matrix of moments associated with $\langle\cdot,\cdot\rangle_L$ and $[\cdot,\cdot]_W^L$, respectively computed in terms of the canonical basis, then we have

$$H = (\Lambda - \mathbb{D}_A)\tilde{H}, \tag{16}$$

where Λ is the block shift matrix

$$\Lambda = \begin{pmatrix} 0_{p\times p} & I_p & & \\ & 0_{p\times p} & I_p & \\ & & \ddots & \ddots \end{pmatrix}.$$

Arguing as in the proof of Proposition 6, we obtain the following representation of $[\cdot,\cdot]_W^L$ in terms of an orthogonality matrix of measures.

Proposition 12. *Let $W(t) = tI_p - A$ be a matrix polynomial, and let $\tilde{\mu}$ be a matrix of measures supported in $\Im \subset \mathbb{R}$ such that $W(t)d\tilde{\mu} = d\mu$. Assume that the matrix moments associated with $\tilde{\mu}$ are finite. Let $P(t) = \sum_{k=0}^{m} a_k(tI_p - A)^k$ and $Q(t) = \sum_{k'=0}^{m'} b_{k'}(tI_p - A)^{k'}$ be polynomials of degree m and m', respectively. Then, we get the following representation of* $[\cdot,\cdot]_W$

$$[P(t), Q(t)]_W^L = \int_{\Im} P d\tilde{\mu} Q^T + \sum_{i=0}^{m'} P(A)\tilde{W}_i^T [Q^{(i)}(A)]^T, \tag{17}$$

where

$$\tilde{W}_i = \frac{1}{i!}\left(\tilde{\mu}_{0,i} - \int_{\Im} d\tilde{\mu}(tI_p - A)^{iT}\right).$$

Remark 13. *In the scalar case, assuming $W_i = 0$ for $i \geq 1$, we get the symmetric bilinear form*

$$[P, Q]_W^L = \int P(t)Q(t)d\hat{\mu} + \tilde{W}_0 P(a)Q(a), \tag{18}$$

which has been extensively studied in the literature (see [9], [10]). Notice that in this case we have $[WP, Q]_W^L = [P, WQ]_W^L$ for polynomials P, Q. For $p \geq 2$, symmetry would require $W(t)d\mu = d\mu W^T(t)$, which is a very restrictive condition on W.

3.1 Bi-orthogonal polynomials associated to the non-symmetric Geronimus transformation

Recall that we have the block Cholesky factorization $H = T^{-1}DT^{-T}$. Now, assume that $\tilde{H}$ has a block LU factorization, i.e. there exist a block lower triangular matrix $\tilde{L}$ with blocks I_p in its main diagonal and a nonsingular block upper triangular matrix $\tilde{U}$ such that $\tilde{H} = \tilde{L}\tilde{U}$. Then, the sequences of polynomials $(P_n^{[1]}(t))_{n\in\mathbb{N}}$ and $(P_n^{[2]}(t))_{n\in\mathbb{N}}$ defined by

$$P^{[1]} = \tilde{L}^{-1}\chi(t) \qquad P^{[2]} = \tilde{U}^{-T}\chi(t), \tag{19}$$

where $P^{[1]} = [P_0^{[1]}, P_1^{[1]}, \ldots]^T$, $P^{[2]} = [P_0^{[2]}, P_1^{[2]}, \ldots]^T$, are bi-orthonormal with respect to the bilinear form $[\cdot,\cdot]_W^L$, since

$$[P^{[1]}, P^{[2]}]_W = [\tilde{L}^{-1}\chi(t), \tilde{U}^{-T}\chi(t)]_W^L = \tilde{L}^{-1}[\chi(t), \chi(t)]_W^L \tilde{U}^{-1} = \tilde{L}^{-1}\tilde{H}\tilde{U}^{-1} = I_p.$$

Notice that $(P_n^{[1]}(t))_{n\in\mathbb{N}}$ is a monic matrix polynomial sequence and since the leading matrix coefficients of the sequence $(P_n^{[2]}(t))_{n\in\mathbb{N}}$ are non singular, we can denote its corresponding monic sequence by $(R_n^{[2]}(t))_{n\in\mathbb{N}}$. Furthermore, we also have

$$\begin{aligned} &[P_n^{[1]}(t), t^m I_p]_W^L = 0_{p\times p} \quad \text{if} \quad m < n, \\ &[t^m I_p, R_n^{[2]}(t)]_W^L = 0_{p\times p} \quad \text{if} \quad m < n. \end{aligned}$$

The next result states a connection formula between such bi-orthogonal sequences and $(P_n(t))_{n\in\mathbb{N}}$.

Proposition 14. *Assume that $\tilde{H}$ has a LU factorization $\tilde{H} = \tilde{L}\tilde{U}$. Then,*

$$\begin{aligned} U_1 P^{[1]} &= P(tI_p - A), \\ R^{[2]} &= L_2 P, \end{aligned}$$

where U_1 is a block upper bi-diagonal matrix and L_2 is a block lower bi-diagonal matrices with blocks I_p in its main diagonal.

Proof. Taking into account (16) and (19),

$$\begin{aligned} &H = (\Lambda - \mathbb{D}_A)\tilde{H} \\ &T_1^{-1}D_1T_1^{-T} = (\Lambda - \mathbb{D}_A)\tilde{L}\tilde{U} \\ &\tilde{U}^{-1}\tilde{L}^{-1}\chi(t) = T^T D^{-1} T(\Lambda - \mathbb{D}_A)\chi(t). \end{aligned}$$

Since $(\Lambda - \mathbb{D}_A)\chi(t) = \chi(t)(tI_p - A)$, we obtain

$$U_1 P^{[1]} = P(tI_p - A), \tag{20}$$

where U_1 is the block upper triangular matrix $U_1 = DT^{-T}\tilde{U}^{-1}$. By comparing the degrees of the polynomials, it is clear that U_1 is block bidiagonal. On the other hand, for $k \geq 1$

$$0_{p\times p} = [(tI_p - A)^k, R_n^{[2]}(t)]_W^L = \left\langle (tI_p - A)^{k-1}, R_n^{[2]}(t) \right\rangle_L, \quad n > k,$$

and thus

$$\left\langle P_k(t), R_n^{[2]}(t) \right\rangle_L = 0_{p\times p} \quad \text{if} \quad k < n-1.$$

As a consequence, we obtain

$$R_n^{[2]}(t) = P_n(t) + A_{n-1}^{[n]} P_{n-1}(t), \tag{21}$$

or, in matrix form,

$$R^{[2]} = L_2 P \quad \text{with} \quad L_2 = \begin{pmatrix} I_p & & & \\ A_0^{[1]} & I_p & & \\ & A_1^{[2]} & I_p & \\ & & \ddots & \ddots \end{pmatrix}. \tag{22}$$

Moreover, as $[I_p, R_n^{[2]}(t)]_W^L = 0_{p\times p}$, then

$$-[I_p, P_n(t)]_W^L = A_{n-1}^T [I_p, P_{n-1}(t)]_W^L.$$

□

From (21), the polynomials $R_n^{[2]}(t)$ can be expressed in terms of the quasi-determinant

$$R_n^{[2]}(t) = \begin{vmatrix} \boxed{P_n(t)} & ([I_p, P_n(t)]_W^L)^T \\ P_{n-1}(t) & ([I_p, P_{n-1}(t)]_W^L)^T \end{vmatrix}, \quad n \geqslant 1. \tag{23}$$

Furthermore, since $(P_n(t))_{n\in\mathbb{N}}$ is a basis for the space $\mathbb{P}^{p\times p}(\mathbb{R})$, there exists a block lower triangular matrix L_1 such that

$$P^{[1]} = L_1 P, \tag{24}$$

i.e. for any fixed n there exists a sequence of $p \times p$ matrices $(B_k^{[n]})_{k=0}^{n-1}$ such that

$$P_n^{[1]}(t) = P_n(t) + B_{n-1}^{[n]} P_{n-1}(t) + \cdots + B_0^{[n]} P_0(t).$$

Then, for $k = 0, \dots n-1$, we have

$$0_{p\times p} = [P_n^{[1]}(t), P_k^{[2]}]_W^L = [P_n(t), P_k^{[2]}]_W^L + B_{n-1}^{[n]}[P_{n-1}(t), P_k^{[2]}]_W^L + \cdots + B_k^{[n]}[P_k(t), P_k^{[2]}]_W^L,$$

and, as a consequence,

$$B_k^{[n]} = -[P_n(t), P_k^{[2]}(t)]_W^L - \sum_{j=k+1}^{n-1} B_j^{[n]}[P_j(t), P_k^{[2]}(t)]_W^L, \quad k \leqslant n-2,$$

and $B^{[n]}_{n-1} = -[P_n(t), P^{[2]}_{n-1}(t)]^L_W$. Thus, we obtain the quasi-determinant formula

$$P^{[1]}_n(t) = \begin{vmatrix} \boxed{P_n(t)} & [P_n(t), P^{[2]}_{n-1}(t)]^L_W & [P_n(t), P^{[2]}_{n-2}(t)]^L_W & \cdots & [P_n(t), P^{[2]}_{0}(t)]^L_W \\ P_{n-1}(t) & I_p & [P_{n-1}(t), P^{[2]}_{n-2}(t)]^L_W & \cdots & [P_{n-1}(t), P^{[2]}_{0}(t)]^L_W \\ \vdots & & \ddots & \ddots & \vdots \\ P_1(t) & & I_p & & [P_1(t), P^{[2]}_{0}(t)]^L_W \\ P_0(t) & & & & I_p \end{vmatrix}.$$

Remark 15. *In the scalar case (see Remark 13), we have*

$$P^{[1]}_n(t) = \det\begin{pmatrix} P_n(t) & [P_n(t), P^{[2]}_{n-1}(t)]^L_W & [P_n(t), P^{[2]}_{n-2}(t)]^L_W & \cdots & [P_n(t), P^{[2]}_{0}(t)]^L_W \\ P_{n-1}(t) & 1 & [P_{n-1}(t), P^{[2]}_{n-2}(t)]^L_W & \cdots & [P_{n-1}(t), P^{[2]}_{0}(t)]^L_W \\ \vdots & & \ddots & \ddots & \vdots \\ P_1(t) & & 1 & & [P_1(t), P^{[2]}_{0}(t)]^L_W \\ P_0(t) & & & & 1 \end{pmatrix}.$$

Moreover, if we assume $W_i = 0$ for $i \geq 1$ then $[\cdot,\cdot]_w$ is the symmetric bilinear form (18)*, and thus $\tilde{H}$ has a Cholesky factorization $\tilde{H} =: \tilde{L}\tilde{U} = \tilde{L}\tilde{D}\tilde{L}$. This implies that $P^{[1]}_n(t) = R^{[2]}_n(t)$ and using* (23) *we get*

$$R^{[2]}_n(t) = P^{[1]}_n(t) = \frac{1}{[1, P_{n-1}(t)]^L_w} \det\begin{pmatrix} P_n(t) & [1, P_n(t)]^L_w \\ P_{n-1}(t) & [1, P_{n-1}(t)]^L_w \end{pmatrix}, \quad n \geqslant 1. \tag{25}$$

An equivalent formula was obtained in [10] (see also [9]).

Finally, we obtain an alternative expression for the block Jacobi matrix associated with $(P_n(t))_{n\in\mathbb{N}}$ in terms of the block triangular matrices defined above.

Proposition 16. *Let J be the monic block Jacobi matrix associated with the sequence of polynomials $(P_n(t))_{n\in\mathbb{N}}$, and let U_1, L_1 and L_2 be the block triangular matrices defined in* (20), (24)*, and* (22)*, respectively. Then,*

$$P^{[1]}(tI_p - A) = L_1 U_1 P^{[1]}, \quad R^{[2]}(tI_p - A) = L_2 U_1 P^{[1]},$$

as well as

$$J - T\mathbb{D}_A T^{-1} = U_1 L_1.$$

Proof. By using (20) and (24), we have

$$L_1 U_1 P^{[1]} = L_1 P(tI_p - A) = P^{[1]}(tI_p - A)$$

and

$$R^{[2]}(tI_p - A) = L_2 P(tI_p - A) = L_2 U_1 P^{[1]}.$$

On the other hand,

$$P(tI_p - A) = Pt - PA = Pt - T\mathbb{D}_A T^{-1} P = (J - T\mathbb{D}_A T^{-1})P,$$

and also

$$P(tI_p - A) = U_1 P^{[1]} = U_1 L_1 P.$$

As a consequence, $J - T\mathbb{D}_A T^{-1} = U_1 L_1$. □

Remark 17. *If* $p = 1$, $W(t) = (t - a)$, *and* $W_i = 0$ *for* $i \geq 1$, *then from Remark 15 and* (25), *Proposition 11 becomes in the well known result [10, 26]*

$$\tilde{J} - aI = L_1 U_1, \quad J - aI = U_1 L_1,$$

where $\tilde{J}$ *is the Jacobi matrix associated with the sequence* $(P_n^{[1]}(t))_{n\in\mathbb{N}}$.

4 Final remarks and open problems

Another important perturbation of a sesquilinear form is the Christoffel transformation (see [3, 6] for linear functionals and [4] for bilinear forms). For the symmetric and non-symmetric cases are defined by

$$\langle P(t), Q(t)\rangle_W = \int P(t)W(t)d\mu W(t)^T Q^T(t), \quad \text{and} \quad \langle P(t), Q(t)\rangle_W^L = \int P(t)W(t)d\mu Q^T(t), \quad (26)$$

respectively, where $W(t) = (tI_p - A)$ and $d\mu$ is a matrix of measures. If $\langle P, Q\rangle_L$ is the sesquilinear form defined in (1), then using the symmetric and non-symmetric Geronimus transformation (6) and (17) respectively, it is easy to verify the following relations.

- Let us denote by $\mathcal{C}$ and $\mathcal{G}$ the symmetric Christoffel and Geronimus transformations, respectively. Then,

  $$\mathcal{C} \circ \mathcal{G}\left(\langle P, Q\rangle_L\right) = \langle P, Q\rangle_L.$$

  In other words, we obtain the identity transformation. The same fact holds for the composition of non-symmetric Christoffel and Geronimus transformations.

- Conversely, we get

  $$\mathcal{G} \circ \mathcal{C}\left(\langle P, Q\rangle_L\right) = \langle P, Q\rangle_L + \sum_{i=1}^{m} P^{(i)}(A) W_i [Q(A)]^T + \sum_{i=1}^{\prime m} P(A) W_i^T [Q^{(i)}(A)]^T + P(A) W_0^T [Q(A)]^T,$$

  which is an Uvarov type perturbation. If both transformations are non-symmetric, then the obtained discrete part added to the original sesquilinear form becomes

  $$\sum_{i=0}^{\prime m} P(A) W_i^T [Q^{(i)}(A)]^T.$$

Notice that the above compositions between Christoffel and Geronimus transformations are natural extensions of the scalar case. The matrix case is more interesting, since we can consider compositions for symmetric and non-symmetric transformations. For instance, the bilinear form $\langle P, Q\rangle_W^R =: \int P d\mu W^T(t) Q^T$ can be obtained from the composition of a symmetric Christoffel transformation and a non-symmetric Geronimus transformation.

Remark 18. *In this manuscript, we have studied the Geronimus transformation for a matrix polynomial $W(t)$ of degree* 1*, which draws some similarities with the scalar case. The generalization for a polynomial $W(t)$ with arbitrary degree remains an open problem. Notice that, unlike the scalar case, we can not approach this problem by using iterations of Geronimus transformations since, in general, a matrix polynomial can not be written as a product of polynomials of degree* 1*. As a consequence, different techniques are required. We will consider this problem in a further contribution.*

Acknowledgements. The work of Luis E. Garza was supported by Conacyt (México), grant 156668. The work of Francisco Marcellán has been supported by Dirección General de Investigación, Desarrollo e Innovación, Ministerio de Economía y Competitividad of Spain, grant MTM2012-36732-C03-01.

Bibliography

[1] A. I. Aptekarev, E. M. Nikishin, *The scattering problem for a discrete Sturm-Liouville operator*, Mat. Sb **121** (163): 327-358 (1983); Math. USSR Sb. **49** (1984), 325-355.

[2] G. Ariznabarreta, M. Mañas *Darboux transformations for multivariate orthogonal polynomials,* arXiv preprint arXiv:1503.04786.

[3] M. I. Bueno, F. Marcellán, *Darboux transformation and perturbation of linear functionals*, Linear Algebra Appl. **384** (2004), 215-242.

[4] M. I. Bueno, F. Marcellán, *Polynomial perturbations of bilinear functionals and Hessenberg matrices*, Linear Algebra Appl. **414** (2006), 64-83.

[5] M. I. Bueno, K. H. Kwon, F. Marcellán, *Discrete-continuous symmetrized Sobolev inner products* Acta Appl. Math. **82** (2004), 309-331.

[6] T. S. Chihara, *An Introduction to Orthogonal Polynomials*. In : *Mathematics and its Applications Series*, Vol. **13**. Gordon and Breach Science Publishers, New York-London-Paris, 1978.

[7] A. E. Choque Rivero, L. E. Garza, *Moment perturbation of matrix polynomials*, Integral Transforms and Spec. Funct. **26** (2015), 177-191.

[8] D. Damanik, A. Pushnitski, B. Simon *The analytic theory of matrix orthogonal polynomials,* Surv. Approx. Theory **4** (2008), 1-85.

[9] M. Derevyagin, J. C. García-Ardila, F. Marcellán, *Multiple Geronimus transformations* Linear Algebra Appl. **454** (2014), 158-183.

[10] M. Derevyagin, F. Marcellán, *A note on the Geronimus transformation and Sobolev orthogonal polynomials*, Numer. Algorithms **67** (2014), 271-287.

[11] H. Dette , W. J. Studden , *Matrix measures, moment spaces and Favard's theorem for the interval* $[0, 1]$ *and* $[0, \infty)$, Linear. Algebra Appl. (2002) 345:169–193.

[12] A. J. Durán, W. Van Assche, *Orthogonal matrix polynomials and higher-order recurrence relations.* Linear Algebra Appl. **219** (1995), 261-280.

[13] I. Gelfand, S. Gelfand, V. Retakh, R. L. Wilson, *Quasideterminants*, Adv. Math. **193** (2005), 56-141.

[14] Ya. Geronimus, *On polynomials orthogonal with regard to a given sequence of numbers*, Comm. Inst. Sci. Math. Mec. Univ. Kharkoff [Zapiski Inst. Mat. Mech.] (4) **17** (1940), 3-18 (Russian).

[15] R. A. Horn, C. R. Johnson, *Matrix Analysis*, Second Edition, Cambridge Univ. Press, Cambridge, 2013.

[16] I. V. Kovalishina, *Analytic theory of a class of interpolation problems*, Izv. Akad. Nauk SSSR Ser. Mat. **47** (1983), no. 3, 455-497.

[17] F. Marcellán, *Polinomios ortogonales no estándar. Aplicaciones en Análisis Numérico y Teoría de Aproximación*, Rev. Acad. Colomb. Ciencias Exactas, Físicas y Naturales, **30** (117) (2006), 563-579. (In Spanish).

[18] A. S Markus, V. Mereuca, *On the complete n-tuple of roots of the operator equation corresponding to a polynomial operator bundle,* Izv. Akad. Nauk. USSR Ser. Mat. **37** (1973) 1108-1131.

[19] F. Marcellán, P. Maroni, *Sur l'adjonction d'une Masse de Dirac à une Forme Régulière et Semiclassique*, Ann. Mat. Pura Appl. **162** (1992), 1–22.

[20] P. Maroni, *Sur la suite de polynômes orthogonaux associée à la forme* $u = \delta_c + \lambda(x-c)^{-1}$, Period. Math. Hungar. **21** (1990), 223-248 .

[21] P. Maroni, *Une Théorie Algébrique des Polynômes Orthogonaux. Application aux Polynôomes Orthogonaux Semiclassiques*, in: C. Brezinski, L. Gori, A Ronveaux Editors, Orthogonal Polynomials and their Applications, IMACS Ann. Comput. Appl., **9** (1991), Baltzer. Basel. 95–130.

[22] A. Sinap, W. Van Assche *Polynomial interpolation and Gaussian quadrature for matrix-valued functions* Linear Algebra Appl. **207** (1994), 71-114.

[23] A. Sinap, W. Van Assche *Orthogonal matrix polynomials and applications* J. Comput. Appl. Math. **66** (1996), 27-52.

[24] V. Spiridonov, L. Vinet, A. Zhedanov, *Spectral transformations, selfsimilar reductions and orthogonal polynomials*, J. Phys. A: Math. Gen. **30** (1997), 7621–7637.

[25] H. Yakhlef, F. Marcellán, M. Piñar, *Perturbations in the Nevai matrix class of orthogonal matrix polynomials*, Linear Algebra Appl. **336** (2001), 231–254.

[26] G. Yoon, *Darboux transforms and orthogonal polynomials*, Bull. Korean Math. Soc. **39** (2002), 359-376.

[27] A. Zhedanov, *Rational spectral transformations and orthogonal polynomials*, J. Comput. Appl. Math. **85** (1997), 67-86.