An extension of the Geronimus transformation for orthogonal matrix polynomials on the real line

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

Abstract. We consider matrix polynomials orthogonal with respect to a sesquilinear form $\langle \cdot, \cdot \rangle_W$ such that

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

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 N. We deduce the integral representation of such sesquilinear forms in such a way a Sobolev type inner product appears. We obtain a connection formula between the sequences of matrix polynomials orthogonal with respect to μ and $\langle \cdot, \cdot \rangle_W$, as well as a relation between the corresponding block Jacobi and Hessenberg type matrices.

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

Keywords. Matrix orthogonal polynomials, Block Jacobi matrices, Matrix Geronimus transformation, Block Cholesky decomposition, Block LU decomposition, quasi-determinants.

1. Introduction

Matrix polynomials

Recall that if R is a ring, then a left module over R is a set M together with two operations

$$+ : M \times M \to M \text{ and } \cdot : R \times M \to M$$

such that for $m, n \in M$ and $a, b \in \mathbb{R}$ we have

i) $(M.+)$ is an Abelian group.

The work of Luis E. Garza was supported by Conacyt (México), grant 156668. The work of Juan Carlos García-Ardila and Francisco Marcellán has been supported by Dirección General de Investigación, Científica y Técnica , Ministerio de Economía y Competitividad of Spain, grant MTM2015-65888-C4-2-P.

ii) $(a+b)\cdot m = a\cdot m + b\cdot m$ and $a\cdot(m+n) = a\cdot m + a\cdot n$.
iii) $(a\cdot b)\cdot m = a\cdot(b\cdot m)$.

In a similar way, one defines a right module on R. If M is a left and right module over R, then M is said to be a bi-module [27, 30].
M is said to be a free left (or right) module over R if M admits a basis, that is, there exists a subset S of M such that S is not empty, S generates M, ($M =< S >= span(S)$) and S is linearly independent.

Let $\mathbb{R}$ (resp. $\mathbb{C}$) be the set of real (resp. complex) numbers and denote by $\mathbb{R}^{p\times p}$ (resp. $\mathbb{C}^{p\times p}$) the ring of $p\times p$ matrices with real (resp. complex) entries. Recall that for any matrices $A_k \in \mathbb{R}^{p\times p}$, $0 \leqslant k \leqslant n$, with $\det(A_n) \neq 0$, the matrix $P(t) = A_n t^n + A_{n-1}t^{n-1} + \cdots + A_1 t + A_0$ is said to be a matrix polynomial of degree n. In particular, if $A_n = I_p$, the identity $p\times p$ matrix, then the polynomial is said to be monic. The set of matrix polynomials with coefficients in $\mathbb{R}^{p\times p}$ will be denoted by $\mathbb{R}^{p\times p}[t]$. $t_0 \in \mathbb{C}$ is said to be a zero of $P(t)$ if $\det[P(t_0)] = 0$. Clearly, from the above definition, $P(t)$ has at most np zeros.

Observe that $\mathbb{R}^{p\times p}[t]$ is a free bi-module (and, in particular, a left module) on the ring $\mathbb{R}^{p\times p}$ with basis $\{I_p, tI_p, t^2 I_p \ldots\}$. Important submodules of $\mathbb{R}^{p\times p}[t]$ are the sets $\mathbb{R}_n^{p\times p}[t]$ of matrix polynomials of degree less than or equal to n with the basis $\{I_p, tI_p, \ldots t^n I_p\}$ of cardinality $n+1$. Since $\mathbb{R}_n^{p\times p}[t]$ has an invariant basis number [30], then any other basis has the same cardinality. If $(r_n(t))_{n\in\mathbb{N}}$ is a sequence of monic matrix polynomials where each $r_n(t)$ has degree n, then $< (r_n(t))_{n\in\mathbb{N}} >$ is a free left module over $\mathbb{R}^{p\times p}$ with basis precisely $(r_n(t))_{n\in\mathbb{N}}$. Notice that $< (r_n(t))_{n\in\mathbb{N}} >$ is a submodule of $\mathbb{R}^{p\times p}[t]$. Furthermore, for each $n \in \{0,1,2,\ldots\}$ there exist elements $b_{n,k} \in \mathbb{R}^{p\times p}$, $k = 0,\ldots,n-1$, such that

$$r_n(t) = t^n I_p + \sum_{k=0}^{n-1} b_{n,k} t^k.$$

The above relation is equivalent to

$$\begin{pmatrix} r_0(t) \\ r_1(t) \\ r_2(t) \\ \vdots \end{pmatrix} = L \begin{pmatrix} I_p \\ tI_p \\ t^2 I_p \\ \vdots \end{pmatrix},$$

where L is a semi-infinite lower matrix with 1's in the main diagonal. Due to the structure of the matrix L there exists a unique semi-infinite matrix L^{-1} such that $LL^{-1} = L^{-1}L = \mathbb{I}_p$, where $\mathbb{I}_p = diag(I_p, I_p \ldots)$ is the block semi-infinite identity matrix [10]. The above implies that there exists an isomorphism between $\mathbb{R}^{p\times p}[t]$ and $< (r_n(t))_{n\in\mathbb{N}} >$, and therefore $\mathbb{R}^{p\times p}[t] = < (r_n(t))_{n\in\mathbb{N}} >$ and $(r_n(t))_{n\in\mathbb{N}}$ is a basis of $\mathbb{R}^{p\times p}[t]$. Using a similar procedure we get that $\mathbb{R}_n^{p\times p}[t] = < (r_k(t))_{k=0}^n >$ for every $n \in \mathbb{N}$.

Given a matrix polynomial $P(t)$, we can define a polynomial operator $P : \mathbb{C}^{p\times p} \longrightarrow \mathbb{C}^{p\times p}$ such that

$$P(Z) = A_n Z^n + A_{n-1} Z^{n-1} + \cdots + A_1 Z + A_0 I_p, \quad Z \in \mathbb{C}^{p\times p}, \tag{1}$$

i.e. the evaluation of the polynomial $P(t)$ at the matrix Z. It is worth noting the importance of the order of the factors in (1) due to the non-commutativity of the product of matrices. Given $\Omega \in \mathbb{C}^{p\times p}$, it is easily verified that $P(t)$ can be written as $P(t) = P_\Omega(t)(tI_p - \Omega)$, where $P_\Omega(t)$ is a matrix polynomial of degree $n-1$, if and only if the operator P satisfies $P(\Omega) = 0_{p\times p}$ (the null $p\times p$ matrix).

Remark 1. *The following notation will be used in the sequel. If B is the block semi-infinite matrix*

$$B = \begin{pmatrix} B_{1,1} & B_{1,2} & B_{1,3} & \cdots \\ B_{2,1} & B_{2,2} & & \\ B_{3,1} & & \ddots & \\ \vdots & & & \end{pmatrix},$$

where $B_{i,j}$ is a $p\times p$ matrix, and A is a $p\times p$ matrix, then the product $A \circledast B$ will be understood as

$$A \circledast B = \begin{pmatrix} A & & \\ & A & \\ & & \ddots \end{pmatrix} \begin{pmatrix} B_{1,1} & B_{1,2} & B_{1,3} & \cdots \\ B_{2,1} & B_{2,2} & & \\ B_{3,1} & & \ddots & \\ \vdots & & & \end{pmatrix} = \begin{pmatrix} AB_{1,1} & AB_{1,2} & AB_{1,3} & \cdots \\ AB_{2,1} & AB_{2,2} & & \\ AB_{3,1} & & \ddots & \\ \vdots & & & \end{pmatrix}.$$

Similarly, $B \circledast A$ means

$$B \circledast A = \begin{pmatrix} B_{1,1}A & B_{1,2}A & B_{1,3}A & \cdots \\ B_{2,1}A & B_{2,2}A & & \\ B_{3,1}A & & \ddots & \\ \vdots & & & \end{pmatrix}.$$

In particular if B and C are block semi-infinite Hessenberg matrices with blocks of size $p\times p$, then $B(A \circledast C) = (B \circledast A)C$ ([7] *Proposition 2.3). Hereinafter we will always work with semi-infinite Hessenberg matrices. If $P(t) = t^n + A_{n-1}t^{n-1} + \cdots + A_0$ is a matrix polynomial with $A_i \in \mathbb{R}^{p\times p}$ and B is the above block semi-infinite matrix, then we understand $P(B)$ as*

$$P(B) = B^n + A_{n-1} \circledast B^{n-1} + \cdots + A_0 \circledast \mathbb{I}_p.$$

Definition 1. *The block semi-infinite matrix*

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

is said to be the shift matrix.

Notice that if $\chi(t) = (I_p, tI_p, t^2 I_p, \dots)^T$, then $\Lambda\chi(t) = t\chi(t)$. Moreover, if $W(t)$ is an $N-th$ degree matrix polynomial $W(t) = \sum_{k=0}^{N} c_k t^k$, with $c_k \in \mathbb{R}^{p\times p}$, then $\chi(t)W(t) = W(\Lambda)\chi(t)$. Notice also that $\Lambda\Lambda^T = \mathbb{I}_p$, but $\Lambda^T\Lambda = \mathrm{diag}(0_{p\times p}, I_p \dots)$.

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

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

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

$$\begin{vmatrix} \boxed{a_{1,1}} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} = \begin{vmatrix} \boxed{a_{1,1}} & a_{1,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} - \begin{vmatrix} \boxed{a_{1,2}} & a_{1,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} \begin{vmatrix} \boxed{a_{2,2}} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix}^{-1} \begin{vmatrix} \boxed{a_{2,1}} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{vmatrix}, \quad (2)$$

when the right side expression makes sense. For more information on quasi-determinants, we refer the reader to [21].

Spectral theory of matrix polynomials

One of the goals of this contribution is to get an explicit integral representation of a symmetric sesquilinear that is defined in terms of some matrix polynomial. For this purpose we need to introduce the concept of canonical Jordan chain. Let $W(t)$ be a monic matrix polynomial of degree N, let $\lambda_1, \dots, \lambda_q$ be their zeros and let $\alpha_1, \dots, \alpha_q$ be their corresponding multiplicities. Since $W(t)$ is a monic polynomial, we have $\sum_{i=1}^{q} \alpha_i = Np$. For a given λ_k, if there exists a nonzero vector $v_{0,k}$ such that

$$W(\lambda_k)v_{0,k} = 0_p,$$

then $v_{0,k}$ is said to be an eigenvector of $W(t)$ associated with λ_k.

Definition 2. *A sequence of vectors $\{v_{0,k}, v_{1,k} \dots, v_{m_k-1,k}\}$ is said to be a Jordan chain of length m_k associated with λ_k if $v_{0,k}$ is an eigenvector of $W(t)$ corresponding to λ_k and*

$$\sum_{i=0}^{j} \frac{1}{i!} W^{(i)}(\lambda_k) v_{j-i,k} = 0_p, \quad j = 0, \dots, m_k - 1.$$

The maximal length of a Jordan chain corresponding with the zero λ_k is called the multiplicity of the eigenvector $v_{0,k}$ and is denoted by $m(v_{0,k})$. Hereafter we deal only with Jordan chains of maximal length.

Definition 3. *Given a basis* $\{v_{0,k}^{[1]}, v_{0,k}^{[2]} \ldots, v_{0,k}^{[d_k]}\}$ *of the linear subspace* $\operatorname{Ker}(W(\lambda_k))$ *with* $\dim(\operatorname{Ker}(W(\lambda_k))) = d_k$*, a canonical Jordan chain associated with the zero* λ_k *is defined as a system of Jordan chains with maximal length*

$$v_{0,k}^{[i]}, v_{1,k}^{[i]}, \cdots, v_{m_i-1,k}^{[i]}, \quad i = 1, \ldots, d_k.$$

If $m(v_{0,k}^{[i]}) = m_k$*, then*

$$m(\lambda_k, W(t)) =: \sum_{i=1}^{d_k} m_i$$

is said to be the Jordan multiplicity of λ_k*.*

The following proposition, which is a direct consequence of Lemma 12.5 of [28] (see also [22]), will be the main tool in the sequel.

Proposition 2. *For each zero* λ_k *of* $W(t)$ *with multiplicity* α_k*, there exists a canonical maximal Jordan chain*

$$v_{0,k}^{[i]}, v_{1,k}^{[i]}, \cdots, v_{m_i-1,k}^{[i]}, \quad i = 1, \ldots d_k,$$

such that $m(\lambda_k, W(t)) = \alpha_k$ *if and only if* $\{v_{0,k}^{[1]}, \cdots, v_{0,k}^{[d_k]}\}$ *is a basis of* $\operatorname{Ker}(W(\lambda_k))$.

Matrix orthogonal polynomials

Recall that a sesquilinear form $\langle \cdot, \cdot \rangle$ from the bi-module $\mathbb{R}^{p\times p}[t]$ to the ring of the $p \times p$ matrices is a map

$$\langle \cdot, \cdot \rangle : \mathbb{R}^{p\times p}[t] \times \mathbb{R}^{p\times p}[t] \longrightarrow \mathbb{R}^{p\times p}$$

satisfying

i) $\langle AP(t) + BR(t), Q(t) \rangle = A \langle P(t), Q(t) \rangle + B \langle R(t), Q(t) \rangle$,
ii) $\langle P(t), AQ(t) + BR(t) \rangle = \langle P(t), Q(t) \rangle A^T + \langle P(t), R(t) \rangle B^T$,

for any matrices $A, B \in \mathbb{R}^{p\times p}$ and any matrix polynomials $P, Q, R \in \mathbb{R}^{p\times p}[t]$. If additionally $\langle P(t), Q(t) \rangle = \langle Q(t), P(t) \rangle^T$ holds for every polynomials P, Q, then $\langle \cdot, \cdot \rangle$ is said to be a symmetric sesquilinear form.

Let $M = (\mu_{i,j})_{i,j=0}^{p-1}$ be a positive definite symmetric matrix of measures supported on $\mathfrak{I} \subset \mathbb{R}$, and let us introduce the following symmetric sesquilinear form

$$\langle P(t), Q(t) \rangle_L = \int_{\mathfrak{I}} P(t) dM Q^T(t), \quad P(t), Q(t) \in \mathbb{R}^{p\times p}[t]. \tag{3}$$

This is known in the literature (see [31, 32]) as a left inner product. Similarly, we can define the right inner product by

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

Since $\langle R, Q\rangle_L = \left\langle R^T, Q^T\right\rangle_R$ for every $R, Q \in \mathbb{R}^{p\times p}[t]$, in the sequel we only deal with the left inner product.

A generalization of the Gram-Schmidt orthogonalization process for the canonical basis $(t^n I_p)_{n\in\mathbb{N}}$ of $\mathbb{R}^{p\times p}[t]$ allows us to construct a sequence of matrix polynomials $(P_n(t))_{n\in\mathbb{N}}$ such that its leading coefficient is a nonsingular matrix and

$$\int_{\mathfrak{I}} P_n(t)dMP_m^T(t) = \delta_{n,m}S_n, \quad n, m \geqslant 0, \quad deg(P_n(t)) = n,$$

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 $\langle\cdot,\cdot\rangle_L$. Notice that we can always assume that $(P_n(t))_{n\in\mathbb{N}}$ is a sequence of monic matrix orthogonal polynomials. In this situation, it satisfies the three term recurrence relation [12, 32].

$$\begin{aligned} tP_n(t) &= P_{n+1}(t) + B_nP_n(t) + A_nP_{n-1}(t), \quad n \geq 0, \\ P_{-1}(t) &= 0_{p\times p}, \quad P_0(t) = I_p, \end{aligned} \tag{4}$$

where $A_n, B_n \in \mathbb{R}^{p\times p}$ are nonsingular and hermitian matrices, respectively. The above relation can be written in matrix form as

$$tP = J_{mon}P \quad \text{with} \quad J_{mon} = \begin{pmatrix} B_0 & I_p & & & \\ A_1 & B_1 & I_p & & \\ & A_2 & B_2 & I_p & \\ & & \ddots & \ddots & \ddots \end{pmatrix},$$

where $P = [P_0^T(t), P_1^T(t)\cdots]^T$. J_{mon} is called the block Jacobi matrix associated with the sequence $(P_n(t))_{n\in\mathbb{N}}$.

On the other hand, let us define the i, j matrix moment associated with $\langle\cdot,\cdot\rangle_L$ with respect to the basis $(t^n I_p)_{n\in\mathbb{N}}$ by

$$m_{i,j} = \left\langle t^i I_p, t^j I_p\right\rangle_L = \int_{\mathfrak{I}} t^{i+j}dM, \quad i, j \geq 0.$$

The semi-infinite block matrix $H = (m_{i,j})_{i,j=0}^{\infty}$ is called the block Hankel matrix associated with $\langle\cdot,\cdot\rangle_L$. Notice that H can also be written as

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

where $\chi(t) = (I_p, tI_p, t^2I_p, \ldots)^T$. Since M is a positive definite matrix of measures, there exists a semi-infinite lower triangular block matrix T, with blocks I_p in its diagonal, and a semi-infinite block diagonal matrix $D = diag\{S_0, S_1, \ldots\}$ such that $H = T^{-1}DT^{-T}$. This is known as Cholesky block factorization (see [6]). As a consequence, $P = T\chi(t)$, and the orthogonality can also be expressed as

$$\int_{\mathfrak{I}} PdMP^T = T\left(\int_{\mathfrak{I}} \chi(x)dM\chi^T(x)\right)T^T = D.$$

Remark 3. *The sequence of monic matrix polynomials* $(P_n(t))_{n\in\mathbb{N}}$ *can be expressed in terms of the moments* $(m_{i,j})_{i,j=0}^{\infty}$ *as follows. Let us consider the block matrix*

$$H_n = \begin{pmatrix} m_{0,0} & \cdots & m_{0,n-1} \\ \vdots & \cdots & \vdots \\ m_{n-10,} & \cdots & m_{n-1,n-1} \end{pmatrix}.$$

If the block matrices H_n *are nonsingular for every* $n \in \mathbb{N}$*, then the sequence of matrix polynomials* $(P_n(t))_{n\in\mathbb{N}}$ *given by (see* [29]*)*

$$P_n(t) =: t^n I_p - \begin{pmatrix} m_{n,0} & \cdots & m_{n,n-1} \end{pmatrix} H_n^{-1} \begin{pmatrix} I_p \\ tI_p \\ \vdots \\ t^{n-1} I_p \end{pmatrix},$$

is orthogonal with respect to $\langle\cdot,\cdot\rangle_L$*. Furthermore, if* $(r_n(t))_{n\in\mathbb{N}}$ *is another ordered basis of monic polynomials of the left module* $\mathbb{R}^{p\times p}[t]$*, deg*$(r_n(t)) = n$*, and* H_r *is the matrix of moments associated with the new basis, i.e* $(H_r)_{i,j} =: \mu_{i,j} = \left\langle r_i(t), r_j(t)\right\rangle_L$*, then due to the uniqueness of* $(P_n(t))_{n\in\mathbb{N}}$*, we have*

$$P_n(t) = r_n(t) - \begin{pmatrix} \mu_{n,0} & \mu_{n,1} & \cdots & \mu_{n,n-1} \end{pmatrix} (H_r)_n^{-1} \begin{pmatrix} r_0(t) \\ r_1(t) \\ \vdots \\ r_{n-1}(t) \end{pmatrix}. \tag{5}$$

Given a sequence of matrix monic orthogonal polynomials $(P_n(t))_{n\in\mathbb{N}}$, with respect to $\langle\cdot,\cdot\rangle_L$, we define the n-th Christoffel–Darboux kernel matrix polynomial

$$K_n(x,y) := \sum_{k=0}^{n} (P_k(y))^T S_k^{-1} P_k(x).$$

In the same way as the matrix orthogonal polynomials, the kernel matrix polynomial has a representation in terms of the moments associated with the basis $(r_n(t))_{n\in\mathbb{N}}$ as follows

$$K_n(x,y) = \begin{pmatrix} r_0^T(y) & \ldots & r_n^T(y) \end{pmatrix} (H_r)_{n+1}^{-1} \begin{pmatrix} r_0(x) \\ \vdots \\ r_n(x) \end{pmatrix}. \tag{6}$$

Recall that since S_n are positive definite matrices, and thus 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 $\langle P_n(t), P_n(t)\rangle_L$. 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

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

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

$$tQ_n(t) = C_{n+1}Q_{n+1}(t) + E_n Q_n(t) + C_n^T Q_{n-1}(t), \quad n \geq 0, \quad Q_{-1} = 0_{p\times p}, \quad Q_0 = I_p, \tag{7}$$

where the matrices C_n are nonsingular and $E_n = E_n^T$. 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 $(U_n Q_n(t))_{n\in\mathbb{N}}$ is also a sequence of orthonormal polynomials. An analogue of the Favard's theorem for matrix polynomials was proved by A. I. Aptekarev and E. M. Nikishin in [5]. Indeed, they proved that if a sequence of matrix polynomials $(Q_n(t))_{n\in\mathbb{N}}$ satisfies (7), then there exists a symmetric matrix of measures M such that the sequence $(Q_n(t))_{n\in\mathbb{N}}$ is orthonormal with respect to M.

Given the symmetric sesquilinear form (3), the sequence of functions $(F_n(t))_{n\in\mathbb{N}}$ defined by

$$F_n(t) := \int \frac{1}{t-y} P_n(y) dM(y),$$

also satisfies the three term recurrence relation (4) but with different initial conditions. In particular, when $n = 0$ the function

$$F(t) =: F_0(t) = \int \frac{1}{t-y} dM(y) \tag{8}$$

is said to be the matrix Stieltjes function associated with the matrix of measures M. In the scalar case, the Stieltjes function has been extensively studied (see [1, 9, 33]) due to its close relation with the measure (and therefore with their associated orthogonal polynomials). In the matrix case there is an important result given by Durán in [15] where the Markov theorem in the matrix case (see [1] for the scalar case) is proved.
Notice that we can write the matrix Stieltjes function (8) as the formal series

$$F(t) = \sum_{n=0}^{\infty} \frac{1}{t^{n+1}} m_n,$$

where $m_n = m_{i,j}$ if $i + j = n$.

There is an exhaustive literature concerning the theory of matrix orthogonal polynomials that is focused on the extension of results which are known for the scalar case (see for example [11, 15, 16, 19, 29, 31, 32]). Some unexpected results have been obtained. For instance, the classification of sequencess of matrix orthogonal polynomials (far away from the classical diagonal cases) satisfying second order linear differential equations with polynomial coefficients which are independent of the degree of the polynomial eigenfunctions, i.e. the so-called Bochner problem, is still open (see [18]). There are also orthogonal families that do not satisfy a scalar-type Rodrigues' formula [17], or families satisfying a first order linear differential equation [8], a fact that does not hold in the scalar case.

A problem of great interest in recent years is the bispectral problem and the generation of new solutions using the Darboux process [23, 24]. In particular, for matrix polynomials satisfying the bispectral problem, new solutions can be generated using Christoffel and Geronimus transformations (see for example [25]). A first step was given in [2], where the authors study a perturbation of matrix of measures consisting in the multiplication by a matrix polynomial of an arbitrary degree such that its leading coefficient is a non-singular matrix (Matrix Christoffel Perturbation). In a very recent work [3], the same authors study the matrix analogue of the scalar Geronimus transformation as well as several extensions of them including left and right multiplication by different matrix polynomials of matrices whose entries are linear functionals on the linear space of of polynomials with complex coefficients. In both [2] and [3], the representation of the new families of bi-orthogonal polynomials with respect to the new perturbations are given in terms of quasi-determinants.

In this contribution, we focus our attention on symmetric Geronimus transformations for positive definite sesquilinear forms, extending some results obtained in the scalar case (see [13, 14]). In particular, we are interested in the analysis of sesquilinear forms $\langle \cdot, \cdot \rangle_W$ such that

$$\langle PW, QW \rangle_W = \int_{\Im} P dM Q^T, \tag{9}$$

where M is a positive definite matrix of measures and $W(t)$ is a fixed matrix polynomial of arbitrary degree. A first example of a matrix Geronimus transformation was given in [13] (see also [20]), where the authors obtained a relation between a (non-diagonal) Sobolev inner product and a matrix Geronimus transformation. The above problem was studied in [3], as a composition of a Geronimus and Christoffel transformation (in this order), i.e. for a generalized function u_x, its symmetric Geronimus transformation $\check{u}_x$ is given by the transformations

$$\hat{u}_x \det(W(x)) W(x)^T = u_x \mapsto \check{u}_x = \operatorname{adj}(W(x)) \hat{u}_x.$$

We do not use this approach. Instead, we study the symmetric Geronimus transformation in a different way, which allow us to obtain other type of results.

The structure of the manuscript is as follows. In Section 2, we define a symmetric sesquilinear form that represents a Geronimus transformation of a matrix of measures. Thus we get an inner product of Sobolev type in the linear space of matrix polynomials. This is an inverse problem in the sense stated in [13] for the scalar case. In Section 3 a connection formula for the corresponding sequences of matrix orthogonal polynomials is obtained as well as a relation for the corresponding block Jacobi matrices. Finally, assuming some conditions on the structure of the moment matrix of the sesquilinear form, in Section 4 the relation between the corresponding Stieltjes functions is deduced. In such a case we get a spectral linear transform in the sense of [33].

2. A generalized Geronimus transformation for symmetric matrix sesquilinear forms

Let $W(t)$ be a matrix monic polynomial of degree N with Np zeros outside the interior of the convex hull of $\Im$, the support of the matrix of measures dM. Let $\mathcal{B}_W$ be the set $\mathcal{B}_W := \{t^i W^m : i = 0, \ldots, N-1, m > 0\}$. Since $\deg(t^i W^m) = i + Nm$, then $\mathcal{B}_W$

is a basis for the left module $\mathbb{R}^{p\times p}[t]$.

Now, we define a sesquilinear form $\langle f, g\rangle_W$ on $\mathbb{R}^{p\times p}[t]$ such that

$$\langle R(t)W(t), Q(t)W(t)\rangle_W = \int_{\Im} R(t)dMQ(t)^T. \tag{10}$$

Notice that $\langle\cdot,\cdot\rangle_W$ is not completely defined by (10). Indeed, if $\hat{\mu}_{k,j}$ are the moments associated with $\langle\cdot,\cdot\rangle_W$ with respect to the basis $\mathcal{B}_W$, i.e. $\hat{\mu}_{Nm+k,Nm'+k'} = \left\langle t^k W^m(t), t^{k'} W^{m'}(t)\right\rangle_W$, then for $0 \le k, k' \le N-1$, the moments $\hat{\mu}_{k,Nm'+k'}$ and $\hat{\mu}_{Nm+k,k'}$ (this is, the first N rows and columns on the matrix of moments) can be chosen arbitrarily. However, we require that $\hat{\mu}_{Nm+k,k'} = \hat{\mu}^T_{k',Nm+k}$ in order that $\langle\cdot,\cdot\rangle_W$ will be symmetric. If $\langle\cdot,\cdot\rangle_W$ is a non degenerate sesquilinear form and we denote by $H_{\mathcal{B}}$ and $\hat{H}_{\mathcal{B}}$ the block moment matrices associated with dM and $\langle\cdot,\cdot\rangle_W$, respectively, using this basis, then $\hat{H}_{\mathcal{B}}$ and $H_{\mathcal{B}}$ are related as follows

$$H_{\mathcal{B}} = \begin{pmatrix} \mu_{0,0} & \mu_{0,1} & \mu_{0,2} & \cdots \\ \mu_{1,0} & \mu_{1,1} & \mu_{1,2} & \cdots \\ \mu_{2,0} & \mu_{2,1} & \mu_{2,2} & \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad \hat{H}_{\mathcal{B}} = \left(\begin{array}{ccc|c} \hat{\mu}_{0,0} & \cdots & \hat{\mu}_{0,N-1} & \cdots \\ \vdots & \ddots & \vdots & \\ \hat{\mu}_{N-1,0} & \cdots & \hat{\mu}_{N-1,N-1} & \cdots \\ \hline \vdots & & \vdots & H_{\mathcal{B}} \end{array}\right).$$

On the other hand, the matrix of moments $\hat{H}$ associated with $\langle\cdot,\cdot\rangle_W$, computed in terms of the canonical basis $(t^n I_p)_{n\in\mathbb{N}}$ has a block Cholesky factorization $\hat{H} = \hat{T}^{-1}\hat{D}\hat{T}^{-T}$ with $\hat{T}$ a lower triangular block matrix with I_p in its main diagonal and $\hat{D}$ a block diagonal matrix. The block matrices H and $\hat{H}$ are related as follows

$$H = \langle\chi(t),\chi(t)\rangle_L = \langle\chi(t)W(t),\chi(t)W(t)\rangle_W = W(\Lambda)\langle\chi(t),\chi(t)\rangle_W W(\Lambda)^T = W(\Lambda)\hat{H}W(\Lambda)^T.$$

A condition for the existence of the sequence of monic matrix orthogonal polynomials $(\hat{P}_n(t))_{n\in N}$ with respect to $\langle\cdot,\cdot\rangle_W$ is that the matrices $(\hat{H}_{\mathcal{B}})_n$ must be nonsingular for every $n \in N$. Indeed, if we consider the quasi-determinant

$$|(\hat{H}_{\mathcal{B}})_{N+n}| = \begin{vmatrix} \boxed{(\hat{H}_{\mathcal{B}})_N} & E_n \\ E_n^T & (H_{\mathcal{B}})_n \end{vmatrix} \quad \text{with} \quad E_n = \begin{pmatrix} \hat{\mu}_{0,N} & \cdots & \cdots & \hat{\mu}_{0,N+n-1} \\ \vdots & & & \vdots \\ \hat{\mu}_{N-1,N} & \cdots & \cdots & \hat{\mu}_{N-1,N+n-1} \end{pmatrix},$$

then using the determinant formula $\det((\hat{H}_{\mathcal{B}})_{N+n}) = \det((H_{\mathcal{B}})_N)\det(|(\hat{H}_{\mathcal{B}})_{N+n}|)$ we conclude that the sequence of polynomials $(\hat{P}_n(t))_{n\in N}$ will exist if the matrices $(\hat{H}_{\mathcal{B}})_k$, $k = 1,\ldots N$ and $|(\hat{H}_{\mathcal{B}})_{N+n}|$, $n \in \mathbb{N}$ are non singular. Observe that for $n = N(l-1)+s$, with $s = 0,\ldots N-1$ and $l > 1$ we have

$$|(\hat{H}_{\mathcal{B}})_{N+n}| = \left|\begin{array}{c|c|c} \boxed{(\hat{H}_{\mathcal{B}})_N} & A_n & B_n \\ \hline A_n^T & (H_{\mathcal{B}})_{n-N} & C_n \\ \hline B_n^T & C_n^T & D_n \end{array}\right|, \tag{11}$$

where

$$A_n = \begin{pmatrix} \hat{\mu}_{0,N} & \cdots & \hat{\mu}_{0,n-1} \\ \vdots & & \vdots \\ \hat{\mu}_{N-1,N} & \cdots & \hat{\mu}_{N-1,n-1} \end{pmatrix}, \qquad B_n = \begin{pmatrix} \hat{\mu}_{0,n} & \cdots & \hat{\mu}_{0,n+N-1} \\ \vdots & & \vdots \\ \hat{\mu}_{N-1,n} & \cdots & \hat{\mu}_{N-1,n+N-1} \end{pmatrix},$$

$$C_n = \begin{pmatrix} \mu_{0,n-N} & \cdots & \mu_{0,n-1} \\ \vdots & & \vdots \\ \mu_{n-N-1,n-N} & \cdots & \hat{\mu}_{n-N-1,n-1} \end{pmatrix}, \qquad D_n = \begin{pmatrix} \mu_{n-N,n-N} & \cdots & \mu_{n-N,n-1} \\ \vdots & & \vdots \\ \mu_{n-1,n-N} & \cdots & \hat{\mu}_{n-1,n-1} \end{pmatrix}.$$

With this in mind, we get the following proposition.

Proposition 4. *For $\ell > 1$ and $0 \le s \le N-1$,*

$$|(\hat{H}_{\mathcal{B}})_{N\ell+s}| = (\hat{H}_{\mathcal{B}})_N + E_s (H_{\mathcal{B}})_s^{-1} E_s^T \tag{12}$$

$$- \sum_{j=1}^{\ell-1} \begin{vmatrix} \boxed{B_{Nj+s}} & A_{Nj+s} \\ C_{Nj+s} & (H_{\mathcal{B}})_{N(j-1)+s} \end{vmatrix} \begin{vmatrix} \boxed{D_{Nj+s}} & C^T_{Nj+s} \\ C_{Nj+s} & (H_{\mathcal{B}})_{N(j-1)+s} \end{vmatrix}^{-1} \begin{vmatrix} \boxed{B_{Nj+s}} & C^T_{Nj+s} \\ A^T_{Nj+s} & (H_{\mathcal{B}})_{N(j-1)+s} \end{vmatrix}.$$

Moreover

$$\begin{vmatrix} \boxed{B_n} & C_n^T \\ A_n^T & (H_{\mathcal{B}})_{n-N} \end{vmatrix} = \begin{vmatrix} \boxed{B_n} & A_n \\ C_n & (H_{\mathcal{B}})_{n-N} \end{vmatrix}^T =: \begin{pmatrix} d_{n,0} & \cdots & d_{n,n-1} \\ \vdots & & \vdots \\ d_{n+N-1,0} & \cdots & d_{n+N-1,N-1} \end{pmatrix},$$

where

$$d_{m+N,k} = \left\langle P_m(t) W(t) + \left[\left\langle r_m(y), K^T_{m-1}(t,y) - K^T_{n-N-1}(t,y) \right\rangle_L \right] W(t), r_k(t) \right\rangle_W,$$

for $n - N \le m \le n-1$, $0 \le k \le N-1$, and

$$\begin{vmatrix} \boxed{D_n} & C_n^T \\ C_n & (H_{\mathcal{B}})_{n-N} \end{vmatrix} =: \begin{pmatrix} h_{n-N,n-N} & \cdots & h_{n-N,n-1} \\ \vdots & & \vdots \\ h_{n-1,n-N} & \cdots & h_{n-1,n-1} \end{pmatrix},$$

where

$$h_{m,k} = \left\langle P_m(t) + \left[\left\langle r_m(y), K^T_{m-1}(t,y) - K^T_{n-N-1}(t,y) \right\rangle_L \right], r_k(t) \right\rangle_L,$$

for $n - N \le m, k \le n-1$.

Proof. Let $n \ge N$. From (11) and properties of the quasi-determinant [21]

$$|(\hat{H}_{\mathcal{B}})_{N+n}| = \left| \begin{array}{c|c|c} \boxed{(\hat{H}_{\mathcal{B}})_N} & B_n & A_n \\ \hline B_n^T & D_n & C_n^T \\ \hline A_n^T & C_n & (H_{\mathcal{B}})_{n-N} \end{array} \right|.$$

Since B_n and D_n are square matrices, then using Sylvester's theorem (see (2)), we get

$$|(\hat{H}_{\mathcal{B}})_{N+n}| =$$

$$|(\hat{H}_{\mathcal{B}})_n| - \begin{vmatrix} \boxed{B_n} & A_n \\ C_n & (H_{\mathcal{B}})_{n-N} \end{vmatrix} \begin{vmatrix} \boxed{D_n} & C_n^T \\ C_n & (H_{\mathcal{B}})_{n-N} \end{vmatrix}^{-1} \begin{vmatrix} \boxed{B_n} & C_n^T \\ A_n^T & (H_{\mathcal{B}})_{n-N} \end{vmatrix}.$$

Thus, (12) follows in a recursive way. On the other hand, for $n-N \le m \le n-1$ and $0 \le k \le n-1$,

$$d_{m+N,k} = \hat{\mu}_{m+N,k} - (\mu_{m,0}, \cdots, \mu_{m,n-N-1})\,[(H_{\mathcal{B}})_{n-N}]^{-1} \begin{pmatrix} \hat{\mu}_{N,k} \\ \vdots \\ \hat{\mu}_{n-1,k} \end{pmatrix}$$

and using (5) and (6) we get

$$\hat{\mu}_{m+N,k} = \left\langle \left[P_m(t) + \left\langle r_m(y), K_{m-1}^T(t,y) \right\rangle_L \right] W(t), r_k(t) \right\rangle_W,$$

$$(\mu_{m,0}, \cdots, \mu_{m,n-N-1})\,[(H_{\mathcal{B}})_{n-N}]^{-1} \begin{pmatrix} \hat{\mu}_{N,k} \\ \vdots \\ \hat{\mu}_{n-1,k} \end{pmatrix} = \left\langle \left[\left\langle r_m(y), K_{m-1}^T(t,y) - K_{n-N-1}^T(t,y) \right\rangle_L \right] W(t), r_k(t) \right\rangle_W.$$

In the same way, for $n-N \le m, k \le n-1$,

$$h_{m,k} = \mu_{m,k} - (\mu_{m,0}, \cdots, \mu_{m,n-N-1})\,[(H_{\mathcal{B}})_{n-N}]^{-1} \begin{pmatrix} \mu_{0,k} \\ \vdots \\ \mu_{n-N-1,k} \end{pmatrix},$$

and thus we get the result. □

Recall that since $\mathcal{B}_W$ is a basis of left module $\mathbb{R}^{p\times p}[t]$, every matrix polynomial $f(t)$ of degree $n = sN + j$ can be written as

$$f(t) = \sum_{l=0}^{N-1} \sum_{m=0}^{s} a_{l,m} t^l W^m(t), \tag{13}$$

where $a_{l,m} = 0_{p\times p}$ if $m = s$ and $l > j$. Let λ_k be a zero of $W(t)$ with multiplicity α_k, $k = 1, \cdots, q$. If $\{v_{0,k}^{[1]} \ldots, v_{0,k}^{[d_k]}\}$ is a basis for $\mathrm{Ker}(W(\lambda_k))$ with dimension d_k, then from Proposition 2 there exists a canonical Jordan chain

$$v_{0,k}^{[i]}, v_{1,k}^{[i]}, \cdots, v_{m_i-1,k}^{[i]} \quad i = 1, \ldots, d_k,$$

such that $\sum_{i=1}^{d_k} m_i = \alpha_k$. For each $i = 1, \ldots, d_k$, we define the following root vector polynomials

$$v_{k,i}(t) = \sum_{r=0}^{m_i-1} (t-\lambda_k)^r v_{r,k}^{[i]}.$$

Definition 4. *Let λ_k be a zero of $W(t)$ and let $v_{i,k}(t)$, $0 \le i \le d_k$, be its associated root vector polynomial defined as above. For $r \in \mathbb{N}$, we define the matrix linear operator $J_{k,i}^{(r)}(\cdot) : \mathbb{R}^{p\times p}[t] \longrightarrow \mathbb{C}^p$ as follows*

$$J_{k,i}^{(r)}(f(t)) := \frac{1}{r!} \left(f(t) v_{k,i}(t)\right)^{(r)} \Big|_{t=\lambda_k} = \sum_{j=0}^{r} \frac{1}{j!} f^{(j)}(\lambda_k) v_{r-j,k}^{[i]}.$$

From Definition 2 and Proposition 2, we get

$$J_{k,i}^{(r)}(W(t)) = 0_p, \quad r = 0\ldots, m_i - 1,$$

and thus, using the representation (13), we obtain

$$\begin{aligned}
J_{k,1}^{(0)}(f(t)) &= a_{0,0}J_{k,1}^{(0)}(1) + \cdots + a_{N-1,0}J_{k,1}^{(0)}\left(t^{N-1}\right),\\
&\vdots\\
J_{k,1}^{(m_1-1)}(f(t)) &= a_{0,0}J_{k,1}^{(m_1-1)}(1) + \cdots + a_{N-1,0}J_{k,1}^{(m_1-1)}\left(t^{N-1}\right),\\
&\vdots\\
J_{k,d_k}^{(0)}(f(t)) &= a_{0,0}J_{k,d_k}^{(0)}(1) + \cdots + a_{N-1,0}J_{k,d_k}^{(0)}\left(t^{N-1}\right),\\
&\vdots\\
J_{k,1}^{(m_{d_k}-1)}(f(t)) &= a_{0,0}J_{k,d_k}^{(m_{d_k}-1)}(1) + \cdots + a_{N-1,0}J_{k,d_k}^{(m_{d_k}-1)}\left(t^{N-1}\right).
\end{aligned}$$

As a consequence, if we define

$$J_k(f(t)) = \begin{pmatrix} J_{k,1}^{(0)}(f(t)) & \cdots & J_{k,1}^{(m_1-1)}(f(t)) & \cdots & J_{k,d_k}^{(0)}(f(t)) & \cdots & J_{k,d_k}^{(m_{d_k}-1)}(f(t)) \end{pmatrix}_{p\times\alpha_k},$$

then for each $k = 1, \cdots, q$, we get

$$\begin{aligned}
J_k(f(t)) &= \begin{pmatrix} a_{0,0} & \cdots & a_{N-1,0} \end{pmatrix} \begin{pmatrix} J_{k,1}^{(0)}(1) & \cdots & J_{k,d_k}^{(m_{d_k}-1)}(1) \\ \vdots & \cdots & \vdots \\ J_{k,1}^{(0)}(t^{N-1}) & \cdots & J_{k,d_k}^{(m_{d_k}-1)}(t^{N-1}) \end{pmatrix}\\
&= \begin{pmatrix} a_{0,0} & \cdots & a_{N-1,0} \end{pmatrix} \begin{pmatrix} J_k(1) \\ \vdots \\ J_k(t^{N-1}) \end{pmatrix}_{Np\times\alpha_k}.
\end{aligned}$$

With this in mind, we can write

$$\begin{pmatrix} J_1(f(t)) & \cdots & J_q(f(t)) \end{pmatrix} = \begin{pmatrix} a_{0,0} & \cdots & a_{N-1,0} \end{pmatrix}\mathbb{T}, \quad \text{where} \quad \mathbb{T} = \begin{pmatrix} J_1(1) & \cdots & J_q(1) \\ \vdots & & \vdots \\ J_1(t^{N-1}) & \cdots & J_q(t^{N-1}) \end{pmatrix}.$$

Since $\mathbb{T}$ is a nonsingular matrix (see [22], Theorem 1.20.), then

$$\begin{pmatrix} J_1(f(t)) & \cdots & J_q(f(t)) \end{pmatrix}\mathbb{T}^{-1} = \begin{pmatrix} a_{0,0} & \cdots & a_{N-1,0} \end{pmatrix}.$$

On the other hand, let $f^{[1]}(t)$ be the matrix polynomial defined by

$$f^{[1]}(t) = \left(f(t) - \sum_{l=0}^{N-1} a_{l,0}t^l \right) W^{-1}(t).$$

Similarly, we define recursively the following sequence of matrix polynomials $(f^{[i]}(t))_{i=2}^{s}$

$$f^{[i]}(t) = \left(f^{[i-1]}(t) - \sum_{l=0}^{N-1} a_{l,i-1}t^l \right) W^{-1}(t) = \sum_{m=i}^{s}\sum_{l=0}^{N-1} a_{l,m}t^l W^{m-i}(t).$$

Proceeding as above, for the sequence of matrix polynomials $(f^{[i]}(t))_{i=1}^{s}$ we get

$$\left(J_0\left(f^{[i]}(t)\right) \quad \cdots \quad J_q\left(f^{[i]}(t)\right)\right)_{p\times Np}\mathbb{T}^{-1} = \left(a_{0,i} \quad a_{1,i} \quad \cdots \quad a_{N-1,i}\right)_{p\times Np}. \tag{14}$$

We are now ready to obtain an explicit representation for $\langle\cdot,\cdot\rangle_W$.

Proposition 5. *Let $d\hat{M}$ be a positive definite matrix of measures such that $W(t)d\hat{M}W(t)^T = dM$. Let $f = \sum_{l=0}^{N-1}\sum_{m=0}^{s} a_{l,m}t^l W^m(t)$ and $g = \sum_{l'=0}^{N-1}\sum_{m'=0}^{s'} a_{l',m'}t^{l'} W^{m'}(t)$ be arbitrary matrix polynomials. Then $\langle\cdot,\cdot\rangle_W$ can be represented as follows*

$$\begin{aligned}\langle f,g\rangle_W &= \int f d\hat{M} g^T \\ &+ \sum_{m=1}^{r}\left(J_1(f^{[m]}) \quad \cdots \quad J_q(f^{[m]})\right)\mathbb{T}^{-1}\begin{pmatrix}\hat{\Omega}_{Nm,0} & \cdots & \hat{\Omega}_{Nm,N-1}\\ \vdots & & \\ \hat{\Omega}_{N-1+Nm,0} & \cdots & \hat{\Omega}_{N-1+Nm,N-1}\end{pmatrix}\mathbb{T}^{-T}\begin{pmatrix}J_1(g)^T\\ \vdots\\ J_q(g)^T\end{pmatrix}\\ &+ \sum_{m=1}^{r}\left(J_1(f) \quad \cdots \quad J_q(f)\right)\mathbb{T}^{-1}\begin{pmatrix}\hat{\Omega}_{0,Nm} & \cdots & \hat{\Omega}_{0,N-1+Nm}\\ \vdots & & \\ \hat{\Omega}_{N-1,Nm} & \cdots & \hat{\Omega}_{N-1,N-1+Nm}\end{pmatrix}\mathbb{T}^{-T}\begin{pmatrix}J_1(g^{[m]})^T\\ \vdots\\ J_q(g^{[m]})^T\end{pmatrix}\\ &+ \left(J_1(f) \quad \cdots \quad J_q(f)\right)\mathbb{T}^{-1}\begin{pmatrix}\hat{\Omega}_{0,0} & \cdots & \hat{\Omega}_{0,N-1}\\ \vdots & & \\ \hat{\Omega}_{N-1,0} & \cdots & \hat{\Omega}_{N-1,N-1}\end{pmatrix}\mathbb{T}^{-T}\begin{pmatrix}J_1(g)^T\\ \vdots\\ J_q(g)^T\end{pmatrix},\end{aligned}$$

where $r = \max\{s, s'\}$, and

$$\hat{\Omega}_{i+Nj,i'+Nj'} = \left\langle t^i W^j(t), t^{i'} W^{j'}(t)\right\rangle_W - \int t^i W^j(t) d\hat{M}\left(t^{i'} W^{j'}(t)\right)^T, \tag{15}$$

i.e., the difference between the moments associated with the bilinear form $\langle\cdot,\cdot\rangle_W$ and the moments associated with $d\hat{M}$.

Remark 6. *For $m = 0,\dots,r$, the matrices*

$$\begin{pmatrix}\hat{\Omega}_{Nm,0} & \cdots & \hat{\Omega}_{Nm,N-1}\\ \vdots & & \\ \hat{\Omega}_{N-1+Nm,0} & \cdots & \hat{\Omega}_{N-1+Nm,N-1}\end{pmatrix}, \quad \begin{pmatrix}\hat{\Omega}_{0,Nm} & \cdots & \hat{\Omega}_{0,Nm+N-1}\\ \vdots & & \\ \hat{\Omega}_{N-1,Nm} & \cdots & \hat{\Omega}_{N-1,Nm+N-1}\end{pmatrix}, \quad \text{and} \quad \begin{pmatrix}\hat{\Omega}_{0,0} & \cdots & \hat{\Omega}_{0,N-1}\\ \vdots & & \\ \hat{\Omega}_{N-1,0} & \cdots & \hat{\Omega}_{N-1,N-1}\end{pmatrix}$$

depend on the the moments $\hat{\mu}_{k,Nm'+k'}$ and $\hat{\mu}_{Nm+k,k'}$, which can be chosen arbitrarily for $0 \le k, k' \le N-1$.

Proof. Let us write

$$\langle f,g\rangle_W = \left\langle \sum_{l=0}^{N-1}\sum_{m=1}^{s} a_{l,m}t^l W^m(t), \sum_{l'=0}^{N-1}\sum_{m'=1}^{s'} b_{l',m'}t^{l'}W^{m'}(t)\right\rangle_W + \left\langle \sum_{l=0}^{N-1}\sum_{m=1}^{s} a_{l,m}t^l W^m(t), \sum_{l'=0}^{N-1} b_{l',0}t^{l'}\right\rangle_W$$
$$+ \left\langle \sum_{l=0}^{N-1} a_{l,0}t^l, \sum_{l'=0}^{N-1}\sum_{m'=1}^{s'} a_{l',m'}t^{l'}W^{m'}(t)\right\rangle_W + \left\langle \sum_{l=0}^{N-1} a_{l,0}t^l W^m(t), \sum_{l'=0}^{N-1} b_{l',0}t^{l'}\right\rangle_W$$
$$= \int f d\hat{M} g^T + \left\langle \sum_{l=0}^{N-1}\sum_{m=1}^{s} a_{l,m}t^l W^m(t), \sum_{l'=0}^{N-1} b_{l',0}t^{l'}\right\rangle_W - \int \sum_{l=0}^{N-1}\sum_{m=1}^{s} a_{l,m}t^l W^m(t) d\hat{M}\left(\sum_{l'=0}^{N-1} b_{l',0}t^{l'}\right)^T$$
$$+ \left\langle \sum_{l=0}^{N-1} a_{l,0}t^l, \sum_{l'=0}^{N-1}\sum_{m'=1}^{s'} b_{l',m'}t^{l'}W^{m'}(t)\right\rangle_W - \int \sum_{l=0}^{N-1} a_{l,0}t^l d\hat{M}\left(\sum_{l'=0}^{N-1}\sum_{m'=1}^{s'} b_{l',m'}t^{l'}W^{m'}(t)\right)^T$$
$$+ \left\langle \sum_{l=0}^{N-1} a_{l,0}t^l W^m(t), \sum_{l'=0}^{N-1} b_{l',0}t^{l'}\right\rangle_W - \int \sum_{l=0}^{N-1} a_{l,0}t^l d\hat{M}\left(\sum_{l'=0}^{N-1} b_{l',0}t^{l'}\right)^T.$$

Defining $\hat{\Omega}_{i+Nj,i'+Nj'}$ as in (15), we get

$$\sum_{l'=0}^{N-1}\sum_{l=0}^{N-1}\sum_{m=1}^{s} a_{l,m}\hat{\Omega}_{l+Nm,l'}b_{l',0}^T =$$

$$= \begin{pmatrix} \sum_{l=0}^{N-1}\sum_{m=1}^{s} a_{l,m}\hat{\Omega}_{l+Nm,0} & \cdots & \sum_{l=0}^{N-1}\sum_{m=1}^{s} a_{l,m}\hat{\Omega}_{l+Nm,N-1}\end{pmatrix}\begin{pmatrix} b_{0,0}^T \\ \vdots \\ b_{N-1,0}^T\end{pmatrix}$$

$$= \sum_{m=1}^{s}\left(\begin{pmatrix} a_{0,m} & \cdots & a_{N-1,m}\end{pmatrix}\begin{pmatrix}\hat{\Omega}_{Nm,0} \\ \vdots \\ \hat{\Omega}_{N-1+Nm,0}\end{pmatrix}\cdots\begin{pmatrix} a_{0,m} & \cdots & a_{N-1,m}\end{pmatrix}\begin{pmatrix}\hat{\Omega}_{Nm,N-1} \\ \vdots \\ \hat{\Omega}_{N-1+Nm,N-1}\end{pmatrix}\right)\begin{pmatrix} b_{0,0}^T \\ \vdots \\ b_{N-1,0}^T\end{pmatrix}$$

$$= \sum_{m=1}^{s}\begin{pmatrix} a_{0,m} & \cdots & a_{N-1,m}\end{pmatrix}\begin{pmatrix}\hat{\Omega}_{Nm,0} & \cdots & \hat{\Omega}_{Nm,N-1} \\ \vdots & & \\ \hat{\Omega}_{N-1+Nm,0} & \cdots & \hat{\Omega}_{N-1+Nm,N-1}\end{pmatrix}\begin{pmatrix} b_{0,0}^T \\ \vdots \\ b_{N-1,0}^T\end{pmatrix}.$$

Similarly,

$$\sum_{l'=0}^{N-1}\sum_{l=0}^{N-1}\sum_{m'=1}^{s} a_{l,0}\hat{\Omega}_{l,Nm'+l'}b_{l',m'}^T$$
$$= \sum_{m'=1}^{s'}\begin{pmatrix} a_{0,0} & \cdots & a_{N-1,0}\end{pmatrix}\begin{pmatrix}\hat{\Omega}_{0,Nm'} & \cdots & \hat{\Omega}_{0,Nm'+N-1} \\ \vdots & & \\ \hat{\Omega}_{N-1,Nm'} & \cdots & \hat{\Omega}_{N-1,Nm'+N-1}\end{pmatrix}\begin{pmatrix} b_{0,m'}^T \\ \vdots \\ b_{N-1,m'}^T\end{pmatrix},$$

and taking into account the previous equations and (14), the result follows. The limit on the sums can be taken as $r = \max\{s, s'\}$ since the additional terms vanish. □

Corollary 1. *If* $W(t) = (tI_p - A)$, *then*

$$\langle f,g\rangle_W = \int f d\hat{M} g^T + \sum_{i=1}^{s} f^{(i)}(A)\frac{1}{i!}\hat{\Omega}_{i,0}[g(A)]^T + \sum_{i=1}^{s} f(A)\frac{1}{i!}\hat{\Omega}_{0,i}[g^{(i)}(A)]^T + f(A)\hat{\Omega}_{0,0}[g(A)]^T.$$

Proof. Given a polynomial $f(t)$ of degree s, it can be written as $\sum_{m=0}^{s} a_m(tI_p - A)^m$, where $a_m \in \mathbb{R}^{p\times p}$. Assume without loss of generality that

$$A = \begin{pmatrix} N_{p_1}(\lambda_1) & & \\ & \ddots & \\ & & N_{p_r}(\lambda_r) \end{pmatrix}, \quad \text{with} \quad N_{p_k}(\lambda_k) = \begin{pmatrix} \lambda_k & 1 & & & \\ & \lambda_k & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_k & 1 \\ & & & & \lambda_k \end{pmatrix}_{p_k\times p_k},$$

for $p_1+\cdots+p_r = p$. Given a zero λ_k of $W(t)$, define the sequence of vectors $(v_{j,k}^{[1]})_{j=0}^{p_{k-1}}$ as follows: $v_{j,k}^{[1]} = e_{p_0+\cdots p_{k-1+j+1}}$ where $e_i = (0\cdots,1_i,\cdots 0)_p^T$. Since

$$v_{j-1,k}^{[1]} + (\lambda_k I_p - A)v_{j,k}^{[1]} = 0, \quad j = 0,\cdots,p_{k-1},$$

then $(v_{j,k}^{[1]})_{j=0}^{p_{k-1}}$ is a Jordan chain corresponding to λ_k. Thus, if $v_{1,k}(t) = \sum_{j=0}^{p_k-1}(t-\lambda_k)^j v_{j,m}^{[1]}$, then (see Definition 2)

$$J_{k,1}(f^{[m]}) = a_m v_{j,m}^{[1]}.$$

Thus, for $k = 1,\ldots,r$, we have

$$\begin{aligned}\left(J_1(f^{[m]}) \quad \cdots \quad J_r(f^{[m]})\right) &= a_m\left(v_{0,1}^{[1]} \quad \cdots \quad v_{p_1-1,1}^{[1]} \cdots \quad v_{0,r}^{[1]} \cdots \quad v_{p_r-1,r}^{[1]}\right) \\ &= a_m I_p = \frac{1}{m!}f^{(m)}(A).\end{aligned}$$

The above yields the result. Notice the connection between the Jordan chain and the evaluation of a polynomial at a matrix. □

3. Connection formulas

Let $(\hat{P}_n(t))_{n\in\mathbb{N}}$ be the sequence of monic orthogonal matrix polynomials with respect to $\langle\cdot,\cdot\rangle_W$. Using the basis $\mathcal{B}_W$ we can write $\hat{P}_{n+N}(t)$ as

$$\hat{P}_{n+N}(t) = \sum_{m\geq 0}\sum_{l=0}^{N-1} a_{l,m}^{[n+N]} t^l W^m(t).$$

Let $H_{\mathcal{B}}$ and $\hat{H}_{\mathcal{B}}$ be the moment matrices with respect $\langle\cdot,\cdot\rangle_L$ and $\langle\cdot,\cdot\rangle_W$, respectively, computed in terms of the basis $\mathcal{B}_W$. Denoting $r_{Nm+l}(t) =: t^l W^m(t)$, $0 \leq l \leq N-1$, and using the ideas described in Remark 3, we have

$$\hat{P}_{n+N}(t) = r_{n+N}(t) - \begin{pmatrix}\hat{\mu}_{n+N,0} & \hat{\mu}_{n+N,1} & \cdots & \hat{\mu}_{n+N,N+n-1}\end{pmatrix} (\hat{H}_{\mathcal{B}})^{-1}_{n+N} \begin{pmatrix} r_0(t) \\ r_1(t) \\ \vdots \\ r_{n+N-1}(t) \end{pmatrix}.$$

Furthermore, we have

$$(\hat{H}_{\mathcal{B}})_{n+N} = \begin{pmatrix} (\hat{H}_{\mathcal{B}})_N & \tilde{E}_n \\ \tilde{E}_n^T & (H_{\mathcal{B}})_n \end{pmatrix}, \quad \text{where} \quad \tilde{E}_n = \begin{pmatrix} \hat{\mu}_{0,N} & \cdots & \cdots & \hat{\mu}_{0,N+n-1} \\ \vdots & & & \vdots \\ \hat{\mu}_{N-1,N} & \cdots & \cdots & \hat{\mu}_{N-1,N+n-1} \end{pmatrix}.$$

With this in mind, and using the inverse formula for 2×2 block matrices obtained from the Schur complement (see [26]), we obtain

$$\begin{aligned}\left(J_1(\hat{P}_{n+N}), \cdots, J_q(\hat{P}_{n+N})\right) = \Big[&\begin{pmatrix}\hat{\mu}_{n+N,0} & \cdots & \hat{\mu}_{n+N,N-1}\end{pmatrix} \\ &- \begin{pmatrix}\mu_{n,0} & \cdots & \mu_{n,n-1}\end{pmatrix} (H_{\mathcal{B}})_n^{-1} \tilde{E}_n^T \Big] \left((\hat{H}_{\mathcal{B}})_N - \tilde{E}_n (H_{\mathcal{B}})_n^{-1} \tilde{E}_n^T\right)^{-1} \mathbb{T}.\end{aligned}$$

As consequence,

$$\left(J_1(\hat{P}_{n+N}), \cdots, J_q(\hat{P}_{n+N})\right) \mathbb{T}^{-1} = \left(\left\langle P_n(t), I_p \right\rangle_W \quad \cdots \quad \left\langle P_n(t), t^{N-1} I_p \right\rangle_W\right) |(\hat{H}_W)_{n+N}|^{-1}.$$

Let $\varepsilon_{n+N} =: \left(\langle P_n(t), I_p \rangle_W \cdots \langle P_n(t), t^{N-1} I_p \rangle_W \right) |(\hat{H}_W)_{n+N}|^{-1}$. Therefore, we can establish the following connection formula.

Proposition 7. *Assuming that $\langle \cdot, \cdot \rangle_W$ is nontrivial, the following connection formula holds*

$$\hat{P}_{n+N}(t) = P_n(t) W(t) + \varepsilon_{n+N} \left[\begin{pmatrix} I_p \\ \vdots \\ t^{N-1} I_p \end{pmatrix} - \sum_{k=0}^{n-1} \begin{pmatrix} E_{0,k} \\ \vdots \\ E_{N-1,k} \end{pmatrix} P_k(t) W(t) \right],$$

where

$$E_{l,k} = \left\langle t^l I_p, P_k W(t) \right\rangle_W ||P_k||_L^{-2}.$$

Proof. Let $\hat{P}_{n+N}(t) = \sum_{m \geqslant 0} \sum_{l=0}^{N-1} \hat{a}_{l,m}^{[n+N]} t^l W^m(t)$ be the matrix orthogonal polynomial of degree $N + n$ with respect to $\langle \cdot, \cdot \rangle_W$ given in terms of the basis $\mathcal{B}_w$. Since $(P_n(t) W(t))_{n \in \mathbb{N}}$ is a basis of the left module $\mathbb{R}^{p \times p}[t] W(t)$, then there exist matrices $(\gamma_{n,k})_{k=0}^{n-1}$ such that

$$\hat{P}_{n+N}(t) - \sum_{l=0}^{N-1} \hat{a}_{l,0}^{[n+N]} t^l = P_n(t) W(t) + \sum_{k=0}^{n-1} \gamma_{n,k} P_k(t) W(t).$$

Since for $k = 1, \ldots, n-1$, we have

$$\left\langle \hat{P}_{n+N}(t) - \sum_{l=0}^{N-1} \hat{a}_{l,0}^{[n+N]} t^l, P_k(t) W(t) \right\rangle_W = - \sum_{l=0}^{N-1} \hat{a}_{l,0}^{[n+N]} \left\langle t^l I_p, P_k(t) W(t) \right\rangle_W, \tag{16}$$

and thus,

$$\gamma_{n,k} = -\sum_{l=0}^{N-1} \hat{a}_{l,0}^{[n+N]} \left\langle t^l I_p, P_k(t)W(t)\right\rangle_W \|P_k\|_L^{-2} = -\sum_{l=0}^{N-1} \hat{a}_{l,0}^{[n+N]} E_{l,k}. \tag{17}$$

Using (16) and (17) we get

$$\hat{P}_{n+N}(t) = P_n(t)W(t) + \sum_{l=0}^{N-1} \hat{a}_{l,0}^{[n+N]} \left(t^l I_p - \sum_{k=0}^{n-1} E_{l,k} P_k(t)W(t) \right), \tag{18}$$

and since $\left(J_1\left(\hat{P}_{n+N}(t)\right) \quad \cdots \quad J_q\left(\hat{P}_{n+N}(t)\right)\right)\mathbb{T}^{-1} = \left(\hat{a}_{0,0}^{[n+N]} \quad \hat{a}_{1,0}^{[n+N]} \quad \cdots \quad \hat{a}_{N-1,0}^{[n+N]}\right)$ we get the result. □

Proposition 8. *The sequences $(P_n(t))_{n\in\mathbb{N}}$ and $(\hat{P}_n(t))_{n\in\mathbb{N}}$ satisfy the following inverse connection formula*

$$PW(t) = M\hat{P}, \tag{19}$$

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

$$\beta_{n,k} = \begin{cases} I_p, & \text{if} \quad k = n+N, \\ \left\langle P_n(t)W(t), \sum_{l=0}^{N-1} \hat{a}_{l,0}^{[k]} t^l \right\rangle_W \|\hat{P}_k\|_W^{-2}, & \text{if} \quad 0 \le k \le (N-1)+n, \\ 0_{p\times p}, & \text{otherwise.} \end{cases}$$

Proof. Since $(\hat{P}_n(t))_{n\in\mathbb{N}}$ is a basis of the left module $\mathbb{R}^{p\times p}[t]$, then there exist matrices $(\beta_{n,k})_{k=0}^{n+N-1}$ such that

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

If $k < N$, then

$$\beta_{n,k} = \left\langle P_n(t)W(t), \sum_{l=0}^{N-1} \hat{a}_{l,0}^{[k]} t^l \right\rangle_W \|\hat{P}_k\|_W^{-2}.$$

On the other hand, if $k \ge N$, using (18) we get

$$\begin{aligned} \left\langle P_n W(t), \hat{P}_k(t)\right\rangle_W &= \left\langle P_n W(t), P_{k-N}(t)W(t) + \sum_{l=0}^{N-1} \hat{a}_{l,0}^{[k]} \left(t^l I_p - \sum_{i=0}^{k-N-1} E_{l,i} P_i(t)W(t) \right) \right\rangle_W \\ &= \sum_{l=0}^{N-1} \left(\left\langle P_n W(t), t^l I_p \right\rangle_W - \sum_{i=0}^{k-N-1} \langle P_n W(t), P_i(t)W(t)\rangle_W E_{l,i}^T \right) \hat{a}_{l,0}^{[k]T} \\ &= \left\langle P_n W(t), \sum_{l=0}^{N-1} \hat{a}_{l,0}^{[k]} t^l \right\rangle_W . \end{aligned}$$

The above implies that for $k = 0, \ldots, N+n-1$,

$$\beta_{n,k} = \left\langle P_n(t)W(t), \sum_{l=0}^{N-1} \hat{a}_{l,0}^{[k]} t^l \right\rangle_W \|\hat{P}_k\|_W^{-2}.$$

□

Remark 9. *The quantities ε_{n+N}, $E_{l,k}$, and $\beta_{n,k}$ appearing in Propositions* 7 *and* 8, *which depend on the sesquilinear form $\langle\cdot,\cdot\rangle_W$, can be obtained using only "non-perturbed" data using Proposition* 5.

Since $\langle\cdot,\cdot\rangle_W$ is a sesquilinear form that not necessarily satisfies $\langle tP, Q\rangle_W = \langle P, tQ\rangle_W$ for every $P, Q \in \mathbb{R}^{p\times p}[t]$, then the semi-infinite block matrix $\hat{J}$ associated with the multiplication operator by t with respect to the sequence of polynomials $(\hat{P}_n(t))_{n\in\mathbb{N}}$ (i.e., $\hat{P}t = \hat{J}\hat{P}$) is a block Hessenberg matrix.

Proposition 10. *Let $W(t) = \sum_{j=0}^{N} c_j t^j$, $c_j \in \mathbb{R}^{p\times p}$. If J and $\hat{J}$ are the block Jacobi and block Hessenberg matrices with respect to the sequences of monic matrix orthogonal polynomials $(P_n(t))_{n\in\mathbb{N}}$, $(\hat{P}_n(t))_{n\in\mathbb{N}}$, respectively, then*

$$W_T(J) = ML, \quad W_{\hat{T}}(\hat{J}) = LM,$$

where $W_T(t) = \sum_{j=0}^{N}(T \circledast \beta_j)T^{-1}t^j$, $W_{\hat{T}}(t) = \sum_{j=0}^{N}(\hat{T} \circledast \beta_j)\hat{T}^{-1}t^j$ and L is the lower triangular block matrix with I_p in the diagonal and such that $\hat{P} = LP$.

Proof. From the hypothesis and (19), we get $PW(t) = (ML)P$ and $\hat{P}W(t) = (LM)\hat{P}$. On the other hand, taking into account that $P = T\chi(t)$ and the properties of the shift matrix, we have

$$\begin{aligned} PW(t) &= \sum_{j=0}^{N}\left((T \circledast c_j)T^{-1}\right)Pt^j \\ &= \sum_{j=0}^{N}\left((T \circledast c_j)T^{-1}\right)JPt^{j-1} = \cdots = \sum_{j=0}^{N}\left((T \circledast c_j)T^{-1}\right)J^jP \\ &= W_T(J)P. \end{aligned}$$

Thus $(W_T(J) - ML)P = \mathbf{0}$ where $\mathbf{0}$ is the semi-infinite matrix of zeros. Due to J has Jacobi block structure, then it is easy to see that both $W_T(J)$ and ML are block Hessenberg matrices with shape

$$\begin{pmatrix} \overbrace{* \quad * \quad \cdots \quad I_p}^{N+n} & & & \\ * \quad * \quad \cdots \quad * & I_p & & \\ * \quad * \quad \cdots \quad * & * & I_p & \\ \cdots \quad \cdots \quad \cdots \quad \cdots & \cdots & \cdots & \cdots \end{pmatrix}.$$

Since $(P_n(t))_{n\in\mathbb{N}}$ is a basis of left module $\mathbb{R}^{p\times p}[t]$ we conclude that $W_T(J) - ML = \mathbf{0}$. The other equation can be obtained in a similar way. □

4. Stieltjes function for the perturbed sesquilinear form

Let us assume that $\langle\cdot,\cdot\rangle_W$ satisfies

$$\hat{m}_{n,k} = \left\langle t^n I_p, t^k I_p\right\rangle_W = \left\langle t^{n+k} I_p, I_p\right\rangle_W = \left\langle I_p, t^{n+k} I_p\right\rangle_W =: \hat{m}_{n+k}, \quad n, m \in \mathbb{N}. \quad (20)$$

On the other hand, if $W(t)$ has q zeros $\lambda_1, \ldots \lambda_q$ such that $\dim[\mathrm{Ker}(\lambda_i)] = \alpha_i$ and $\sum_{i=1}^{q}\alpha_i = Np$, then each λ_i has α_i associated linearly independent eigenvectors, so we will assume that $W(t)$ has Np zeros $\lambda_1 \ldots \lambda_{Np}$ with associated eigenvectors

$v_1 \dots v_{Np}$. Assume also that the matrices $\hat{\Omega}_{i+Nj,i'+Nj'}$ defined in (15) are the zero matrix if j or j' are different from zero and

$$\mathbb{T}^{-1}\begin{pmatrix} \hat{\Omega}_{0,0} & \cdots & \hat{\Omega}_{0,N-1} \\ \vdots & & \\ \hat{\Omega}_{N-1,0} & \cdots & \hat{\Omega}_{N-1,N-1} \end{pmatrix}\mathbb{T}^{-T} = diag(l_1, \cdots l_{Np}),$$

where $l_1 \dots, l_{Np} \in \mathbb{R}$. With these assumptions, it is easy to see that the representation for $\langle \cdot, \cdot \rangle_W$ becomes

$$\langle P, Q \rangle_W = \int P d\hat{M} Q^T + \sum_{k=0}^{Np} P(\lambda_k) v_k l_k v_k^T Q^T(\lambda_k),$$

and satisfies (20). In general, if we assume that $\langle \cdot, \cdot \rangle_W$ satisfies (20) and $W(t) = \sum_{k=0}^{N} c_k t^k$, then from (10)

$$m_n = \langle t^n W(t), W(t) \rangle_W = \sum_{k,i=0}^{N} c_k \left\langle t^{n+k} I_p, t^i I_p \right\rangle_W c_i^T = \sum_{k,i=0} c_k \hat{m}_{n+k+i} c_i^T.$$

As a consequence, the relation between the corresponding matrix Stieltjes functions can be obtained from

$$\begin{aligned} F(t) &= \sum_{i=0}^{\infty} \frac{m_n}{t^{n+1}} = \sum_{n=0}^{\infty} \sum_{i,k=0}^{N} c_k \frac{\hat{m}_{n+k+i}}{t^{n+k+i+1}} t^{k+i} c_i^T \\ &= \sum_{n=0}^{\infty} \sum_{i,k=0}^{N} c_k \left[\frac{\hat{m}_{n+k+i}}{t^{n+k+i+1}} t^{k+i} + \sum_{s=0}^{k+i-1} \frac{\hat{m}_s}{t^{s+1}} t^{k+i} - \sum_{s=0}^{k+i-1} \frac{\hat{m}_s}{t^{s+1}} t^{k+i} \right] c_i^T \\ &= \sum_{i,k=0}^{N} c_k \left[\hat{F}(t) t^{k+i} - \sum_{s=0}^{k+i-1} \frac{\hat{m}_s}{t^{s+1}} t^{k+i} \right] c_i^T \\ &= W(t) \hat{F}(t) W^T(t) - B(t), \end{aligned}$$

where $B(t) = \sum_{i,k=0}^{N} c_k \left[\sum_{s=0}^{k+i-1} \frac{\hat{m}_s}{t^{s+1}} t^{k+i} \right] c_i^T$ is a matrix polynomial of degree $2N-1$. In other words, the Stieltjes functions associated with the original and the perturbed sesquilinear forms are related by

$$\hat{F}(t) = W(t)^{-1} \left[F(t) + B(t) \right] W^{-T}(t).$$

Acknowledgements

We thank the valuable suggestions, comments and criticism by the referees. They significantly improved the contents and presentation of the manuscript. The work of Luis E. Garza was supported by Conacyt (México), grant 156668. The work of Juan Carlos García-Ardila and Francisco Marcellán has been supported by Dirección General de Investigación, Científica y Técnica , Ministerio de Economía y Competitividad of Spain, grant MTM2015-65888-C4-2-P.

References

[1] N. I. Akhiezer, *Classical moment problem and related questions in analysis,* Oliver & Boyd, Edinburgh 1965.

[2] C. Álvarez-Fernández, G. Ariznabarreta, J. C. García-Ardila, M. Mañas, F. Marcellán, *Christoffel transformations for matrix orthogonal polynomials in the real line and the non-Abelian 2D Toda lattice hierarchy*, Internat. Math. Res. Notices, doi: 10.1093/imrn/rnw027, 2016.

[3] C. Álvarez-Fernández, G. Ariznabarreta, J. C. García-Ardila, M. Mañas, F. Marcellán, *Transformation theory and Christoffel formulas for matrix biorthogonal polynomials on the real line* arXiv:1605.04617 [math.CA].

[4] G. E. Andrews, R. Askey, R. Roy, *Special Functions*, Cambridge Univ. Press, Cambridge, 1999.

[5] 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.

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

[7] M. J. Cantero, F. Marcellán, L. Moral, L. Velázquez, *Darboux transformations for CMV matrices* Adv. Math. **208** (2016), 122-206.

[8] M. M. Castro, F. A. Grünbaum, *Orthogonal matrix polynomials satisfying first order differential equations: a collection of instructive examples,* J. Nonlinear Math. Phys. **12** (2005), suppl. 2, 63-76.

[9] 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.

[10] R.G. Cooke, *Infinite Matrices and Sequence Spaces*, Macmillan & Co., Ltd., London, 1950.

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

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

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

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

[15] A. J. Durán, *Markov's theorem for orthogonal matrix polynomials,* Canad. J. Math. **48** (1996), 1180-1195.

[16] A. J. Durán, *Ratio asymptotic for orthogonal matrix polynomials,* J. Approx. Theory **100** (1999), 304-344.

[17] A. J. Durán, F. A. Grünbaum, *Orthogonal matrix polynomials, scalar-type Rodrigues' formulas and Pearson equations,* J. Approx. Theory **134** (2005), 267-280.

[18] A. J. Durán, F. A. Grünbaum, *A survey on orthogonal matrix polynomials satisfying second order differential equations,* J. Comput. Appl. Math. **178** (2005), 169-190.

[19] A. J. Durán, P. López-Rodríguez, *Orthogonal matrix polynomials: Zeros and Blumenthal's theorem,* J. Approx. Theory **84** (1996), 96-118.

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

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

[22] I. Gohberg, P. Lancaster, L. Rodman, *Matrix Polynomials*, Computer Science and Applied Mathematics. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1982

[23] F. A. Grünbaum, *The Darboux process and a noncommutative bispectral problem: some explorations and challenges,* in *Geometric aspects of analysis and mechanics* 161-177, Progr. Math., *292*, Birkhäuser/Springer, New York 2011.

[24] F. A. Grünbaum, L. Haine, *Bispectral Darboux transformations: an extension of the Krall polynomials,* Internat. Math. Res. Notices **8** (1997), 359-392.

[25] F. A. Grünbaum and M. Yakimov, *Discrete bispectral Darboux transformations from Jacobi operators*, Pacific J. Math. **204** (2002), 395-431.

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

[27] S. Lang, *Algebra*, , Third Edition, Springer-Verlag, New York, 2005.

[28] A. S. Markus, *Introduction to the spectral theory of polynomials operator pencil,* Translated from the Russian by H. H. McFaden. Translation edited by Ben Silver. With an appendix by M. V. Keldysh. Translations of Mathematical Monographs, **71**. Amer. Math. Soc., Providence, RI, 1988.

[29] L. Miranian, *Matrix valued orthogonal polynomials on the real line: some extensions of the classical theory*, J. Phys. A, **38** (2005), 5731-5749.

[30] L. Rowen, *Ring Theory,* Vol. I, Academic Press, San Diego, CA, 1988.

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

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

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

Juan Carlos García-Ardila
Departamento de Matemáticas, Universidad Carlos III de Madrid,
Avenida Universidad 30, Leganés,
Spain
e-mail: jugarcia@math.uc3m.es

Luis E. Garza
Facultad de Ciencias, Universidad de Colima,
Bernal Díaz del Castillo 340, Colima
México
e-mail: garzaleg@gmail.com

Francisco Marcellán
Departamento de Matemáticas, Universidad Carlos III de Madrid,
Avenida Universidad 30, Leganés,
Instituto de Ciencias Matemáticas (ICMAT),
C/ Nicolás Cabrera, 13-15,
Campus de Cantoblanco UAM, 28049 Madrid,
Spain
e-mail: pacomarc@ing.uc3m.es