

MARKOV'S THEOREM FOR ORTHOGONAL MATRIX POLYNOMIALS

ANTONIO J. DURAN

ABSTRACT. Markov's Theorem shows asymptotic behavior of the ratio between the n -th orthonormal polynomial with respect to a positive measure and the n -th polynomial of the second kind. In this paper we extend Markov's Theorem for orthogonal matrix polynomials.

1. Introduction. A. Markov established in 1895 (see [M]) the following result, which is now known as Markov's theorem:

$$\lim_{n \rightarrow \infty} \frac{q_n(z)}{p_n(z)} = \int_a^b \frac{d\mu(t)}{z - t} \quad \text{for } z \in \mathbb{C} \setminus [a, b],$$

where μ is a probability measure on the finite interval $[a, b]$, $(p_n)_n$ is the sequence of orthonormal polynomials with respect to μ and $(q_n)_n$ is the corresponding sequence of polynomials of the second kind, defined by

$$q_n(x) = \int \frac{p_n(x) - p_n(t)}{x - t} d\mu(t), \quad n \geq 0.$$

The hypothesis μ having compact support is too restrictive, and actually the determinacy of the measure μ is a sufficient condition. Even for some families of indeterminate measures Markov's theorem holds (see the recent survey about Markov's theorem [B]).

The purpose of this paper is to extend Markov's theorem for orthogonal matrix polynomials.

We consider a $N \times N$ positive definite matrix of measures W (for any Borel set $A \subset \mathbb{R}$, $W(A)$ is a positive semidefinite numerical matrix), having moments of every order, *i.e.*, the matrix integral

$$\int_{\mathbb{R}} t^n dW(t)$$

exists for any nonnegative integer n .

The matrix inner product defined in the usual way by W in the space of matrix polynomials has associated a sequence of orthonormal matrix polynomials $(P_n)_n$, satisfying

$$\int P_n(t) dW(t) P_m^*(t) = \delta_{n,m} I, \quad n, m \geq 0.$$

$P_n(t)$ is a matrix polynomial of degree n , with a non-singular leading coefficient and is defined upon a multiplication on the left by a unitary matrix.

This work has been partially supported by DGICYT ref. PB93-0926.

Received by the editors March 2, 1995; revised July 5, 1995.

AMS subject classification: Primary: 42C05, 47A56, 65D32.

© Canadian Mathematical Society 1996.

As in the scalar case, the sequence of orthonormal matrix polynomials $(P_n)_n$ satisfies a three-term recurrence relation

$$tP_n(t) = A_{n+1}P_{n+1}(t) + B_nP_n(t) + A_n^*P_{n-1}(t), \quad n \geq 0, \quad P_{-1}(t) = \theta,$$

where A_n are non-singular matrices and B_n are hermitian. Without loss of generality, we assume $P_0(t) = I$. (Here and in the rest of this paper, we write θ for the null matrix, the dimension of which can be determined from the context.)

This three term recurrence relation characterizes the orthonormality of a sequence of matrix polynomials with respect to a positive definite matrix of measures (see, for instance, [AN] or [DL]). In [D1], [D2] and [DV] a very close relationship between orthogonal matrix polynomials and scalar polynomials satisfying a higher order recurrence relation has been established.

The corresponding matrix polynomials of the second kind are defined by

$$Q_n(x) = \int \frac{P_n(x) - P_n(t)}{x - t} dW(t).$$

We say that the positive definite matrix of measures W is determinate if no other positive definite matrix of measures has the same moments as those of W , i.e., the positive definite matrix of measures W is uniquely determined by the moments $\int t^n dW(t)$ ($n \in \mathbb{N}$).

To establish Markov's theorem, we need the following definitions: Δ_n stands for the set of zeros of the matrix polynomial P_n , i.e., the zeros of $\det(P_n)$. In [DL], it is proved that these zeros are real and have multiplicity at most N . We finally put

$$(1.1) \quad \Gamma = \bigcap_{N \geq 0} M_N, \quad \text{where } M_N = \overline{\bigcup_{n \geq N} \Delta_n}.$$

It is proved in [DL] that orthogonalizing matrix of measures μ for the matrix polynomials $(P_n)_n$ can be found as weak accumulation points of a sequence of discrete measures μ_n with support precisely Δ_n . Therefore, for a determinate matrix of measures W , we have $\text{supp}(W) \subset \Gamma$.

The main result of this paper is the following matrix extension of Markov's Theorem.

THEOREM 1.1. *Assume that W is determinate. Then*

$$\lim_{n \rightarrow \infty} P_n^{-1}(z)Q_n(z) = \int \frac{dW(t)}{z - t} \quad \text{for } z \in \mathbb{C} \setminus \Gamma,$$

and the convergence is uniform for z in compact subsets of $\mathbb{C} \setminus \Gamma$.

We will show that in the matrix version of Markov's theorem, the matrices P_n^{-1} and Q_n must be multiplied in the order $P_n^{-1}Q_n$, otherwise the result could be false.

To prove Markov's theorem we find a quadrature formula (Section 3) for the inner product defined by the matrix of measures W in the space of matrix polynomials. The matrix coefficients, which appear in that quadrature formula, will be given in a closed expression, and this expression will be the key to establish the matrix version of Markov's Theorem (see [SV, and DL] for other versions of the quadrature formula).

In the proof of this quadrature formula we use a property of the zeros (Section 2) of the n -th orthonormal matrix polynomial P_n which will play a fundamental role in the whole theory of orthogonal matrix polynomials:

If α is a zero of $P_n(t)$ of multiplicity p , then $\text{rank}(P_n(\alpha)) = N - p$ (hence $p \leq N$), and the matrices $(\text{Adj}(P_n(t)))^{(p-1)}(\alpha)$ and $P_{n-1}^*(\alpha)A_n$ define two isomorphisms from the space of right eigenvectors of $P_n(\alpha)$ associated to the eigenvalue 0 into the space of left eigenvectors of $P_n(\alpha)$ associated to the eigenvalue 0. The inverse mapping of $P_{n-1}^*(\alpha)A_n$ is that defined by the matrix $Q_n(\alpha)$.

For a given matrix A , we denote by $\text{Adj}(A)$ the classical adjoint, *i.e.*, the matrix uniquely defined by the property

$$A \text{Adj}(A) = \text{Adj}(A)A = \det(A)I.$$

That the zeros of orthogonal polynomials are simple (in the scalar case) is a well-known property, which is not true in the matrix case. Precisely, the property we have noted above will play in the matrix case the same role as that of the non-multiplicity of the zeros in the scalar case.

To complete this paper, we give a generalization of Markov’s theorem showing the asymptotic behavior between the sequence of orthonormal matrix polynomials and the sequence of k -th associated polynomials (Section 5)

2. Zeros. We start with the following lemma which contains the matrix version of some classical formulae for orthonormal scalar polynomials. The proofs work as in the scalar case and so are omitted.

LEMMA 2.1. (1) *The Christoffel-Darboux formula and some special cases*

$$(2.1) \quad \begin{aligned} P_{n-1}^*(z)A_nP_n(w) - P_n^*(z)A_n^*P_{n-1}(w) \\ = (w - z) \sum_{k=0}^{n-1} P_k^*(z)P_k(w), \quad z, w \in \mathbb{C}. \end{aligned}$$

$$(2.2) \quad P_{n-1}^*(z)A_nP_n(z) - P_n^*(z)A_n^*P_{n-1}(z) = \theta, \quad z \in \mathbb{C}.$$

$$(2.3) \quad P_{n-1}^*(z)A_nP'_n(z) - P_n^*(z)A_n^*P'_{n-1}(z) = \sum_{k=0}^{n-1} P_k^*(z)P_k(z), \quad z \in \mathbb{C}.$$

(2) *Some particular cases of the Green formula:*

$$(2.4) \quad \begin{aligned} P_{n-1}^*(z)A_nQ_n(w) - P_n^*(z)A_n^*Q_{n-1}(w) \\ = I + (w - z) \sum_{k=0}^{n-1} P_k^*(z)Q_k(w), \quad z, w \in \mathbb{C}. \end{aligned}$$

$$(2.5) \quad P_{n-1}^*(z)A_nQ_n(z) - P_n^*(z)A_n^*Q_{n-1}(z) = I, \quad z \in \mathbb{C}.$$

(3) *The Liouville-Ostrogradski formula*

$$(2.6) \quad Q_n(z)P_{n-1}^*(z) - P_n(z)Q_{n-1}^*(z) = A_n^{-1}.$$

To prove the remarkable property about the zeros of P_n stressed in the introduction, and the quadrature formula, we need the following lemma:

LEMMA 2.2. *Let $A(t)$ be a $N \times N$ matrix polynomial and let a be a zero of $A(t)$ of multiplicity p , i.e., a zero of multiplicity p of the scalar polynomial $\det(A(t))$. We put*

$$L(a, A) = \{v \in \mathbb{C}^N : vA(a) = \theta\}, \quad R(a, A) = \{v \in \mathbb{C}^N : A(a)v^* = \theta\}.$$

If $\dim(L(a, A)) = \dim(R(a, A)) = p$, then $(\text{Adj}(A(t)))^{(l)}(a) = \theta$, for $l = 0, \dots, p - 2$ and $(\text{Adj}(A(t)))^{(p-1)}(a) \neq \theta$. Moreover $\text{rank}((\text{Adj}(A(t)))^{(p-1)}(a)) = p$ and

$$(\text{Adj}(A(t)))^{(p-1)}(a)$$

defines a linear mapping from \mathbb{C}^N onto $L(a, A)$ which is an isomorphism from $R(a, A)$ into $L(a, A)$.

PROOF. It follows straightforward from Lemma 2.2 of [DL] that

$$(\text{Adj}(A(t)))^{(l)}(a) = \theta, \quad \text{for } l = 0, \dots, p - 2.$$

We now prove that $(\text{Adj}(A(t)))^{(p-1)}(a) \neq \theta$. Since a is a zero of $A(t)$ of multiplicity p , by differentiating the formula $\text{Adj}(A(t))A(t) = \det A(t)I$ and taking into account what we have already proved, we obtain that

$$(2.7) \quad (\text{Adj}(A(t)))^{(p-1)}(a)A(a) = \theta, \quad \text{and}$$

$$(2.8) \quad (\text{Adj}(A(t)))^{(p)}(a)A(a) + (\text{Adj}(A(t)))^{(p-1)}(a)A'(a) = (\det A(t))^{(p)}(a)I.$$

If $(\text{Adj}(A(t)))^{(p-1)}(a) = \theta$, (2.8) gives that

$$(\text{Adj}(A(t)))^{(p)}(a)A(a) = (\det A(t))^{(p)}(a)I.$$

But this is impossible because $A(a)$ is singular but $(\det A(t))^{(p)}(a)I$ is non-singular (a is a zero of multiplicity just p of $\det(A(t))$).

The formula (2.7) gives that the matrix

$$(\text{Adj}(A(t)))^{(p-1)}(a)$$

defines a linear mapping

$$\begin{aligned} & \left(\text{Adj}(A(t))\right)^{(p-1)}(a): \mathbb{C}^N \rightarrow L(a, A) \\ & v \mapsto v\left(\text{Adj}(A(t))\right)^{(p-1)}(a). \end{aligned}$$

For $v \in R(a, A)$, $v \neq \theta$, (2.8) gives that

$$v\left(\text{Adj}(A(t))\right)^{(p-1)}(a)A'(a)v^* \neq \theta.$$

So, the linear mapping defined by $\left(\text{Adj}(A(t))\right)^{(p-1)}(a)$ is injective in $R(a, A)$. Hence,

$$\text{rank}\left(\text{Adj}(A(t))\right)^{(p-1)}(a) \geq \dim(R(a, A)) = p.$$

But (2.7) gives that the rows of $\left(\text{Adj}(A(t))\right)^{(p-1)}(a)$ are vectors of $L(a, A)$, hence

$$\text{rank}\left(\text{Adj}(A(t))\right)^{(p-1)}(a) \leq \dim(L(a, A)) = p.$$

And the proof is finished. ■

Some properties of zeros of orthogonal matrix polynomials on the real line have been established recently in [SV] and [DL] (see also [Z]). We complete here those results by proving the property we noted in the introduction of this paper.

THEOREM 2.3. *If a is a zero of $P_n(t)$ of multiplicity p , then*

- (1) *a is real, $p \leq N$, $\text{rank}(P_n(a)) = N - p$, and $\dim(R(a, P_n)) = \dim(L(a, P_n)) = p$. $R(a, P_n)$ and $L(a, P_n)$ are, respectively, the spaces of right and left eigenvectors of $P_n(a)$ associated to the eigenvalue 0.*
- (2) *The matrix $P_{n-1}^*(a)A_n$ defines an isomorphism from $R(a, P_n)$ into $L(a, P_n)$. Its inverse mapping is the isomorphism defined by the matrix $Q_n(a)$.*
- (3) *The matrix $\left(\text{Adj}(P_n(t))\right)^{(p-1)}(a)$ defines a linear mapping from \mathbb{C}^N onto $L(a, P_n)$ which is an isomorphism from $R(a, P_n)$ into $L(a, P_n)$.*
- (4) *$\left(\text{Adj}(P_n(t))\right)^{(p-1)}(a)P_n(a) = P_n(a)\left(\text{Adj}(P_n(t))\right)^{(p-1)}(a) = \theta$ and*

$$\text{rank}\left(\text{Adj}(P_n(t))\right)^{(p-1)}(a) = p,$$

being the p linearly independent rows of $\left(\text{Adj}(P_n(t))\right)^{(p-1)}(a)$ a basis of the linear space of left eigenvectors of $P_n(a)$ associated to 0.

PROOF. (1) It is contained in parts (1) and (2) of Theorem 1.1 in [DL].

(2) First of all, we prove that if $v \in R(a, P_n)$ then $vP_{n-1}^*(a)A_n \in L(a, P_n)$. Indeed, if $P_n(a)v^* = \theta$, the formula (2.2) gives that $vP_{n-1}^*(a)A_nP_n(a) = \theta$, that is, $vP_{n-1}^*(a)A_n \in L(a, P_n)$.

For a vector $v \in R(a, P_n)$, the formula (2.5) gives

$$(2.9) \quad vP_{n-1}^*(a)A_nQ_n(a) = v,$$

and so, $vP_{n-1}^*(a)A_{n-1} \neq \theta$. This proves that the mapping is injective. Since the dimension of $R(a, P_n)$ and $L(a, P_n)$ is the same, the mapping is automatically an isomorphism.

(2.9) shows that its inverse mapping is just that defined by $Q_n(a)$.

(3) and (4) From (1) we have that $\dim(R(a, P_n)) = \dim(L(a, P_n)) = p$. Then (3) and (4) of Theorem 2.3 are straightforwardly deduced from Lemma 2.2. ■

Theorem 2.3 can be extended for perturbations of the orthonormal polynomials. Indeed, let A be a matrix satisfying that A_nA is hermitian, and let us consider the polynomial $P_n(t) - AP_{n-1}(t)$. Then, Theorem 2.3 holds for this polynomial, *i.e.*:

THEOREM 2.4. *If a is a zero of $P_n(t) - AP_{n-1}(t)$ of multiplicity p , where $A_nA = A^*A_n^*$, then*

- (1) *a is real, $p \leq N$, $\text{rank}(P_n(a) - AP_{n-1}(a)) = N - p$, and $\dim(R(a, P_n - AP_{n-1})) = \dim(L(a, P_n - AP_{n-1})) = p$.*
- (2) *The matrix $P_{n-1}^*(a)A_n$ defines an isomorphism from $R(a, P_n - AP_{n-1})$ into $L(a, P_n - AP_{n-1})$. Its inverse mapping is the isomorphism defined by the matrix $Q_n(a) - AQ_{n-1}(a)$.*
- (3) *The matrix $(\text{Adj}(P_n(t) - AP_{n-1}(t)))^{(p-1)}(a)$ defines a linear mapping from \mathbb{C}^N onto $L(a, P_n - AP_{n-1})$ which is an isomorphism from $R(a, P_n - AP_{n-1})$ into $L(a, P_n - AP_{n-1})$.*
- (4)

$$\begin{aligned} & (\text{Adj}(P_n(t) - AP_{n-1}(t)))^{(p-1)}(a)(P_n(a) - AP_{n-1}(a)) \\ &= (P_n(a) - AP_{n-1}(a))(\text{Adj}(P_n(t) - AP_{n-1}(t)))^{(p-1)}(a) = \theta \end{aligned}$$

and

$$\text{rank}(\text{Adj}(P_n(t) - AP_{n-1}(t)))^{(p-1)}(a) = p,$$

being the p linearly independent rows of $(\text{Adj}(P_n(t) - AP_{n-1}(t)))^{(p-1)}(a)$ a basis of the linear space of left eigenvectors of $P_n(a) - AP_{n-1}(a)$ associated to 0.

By modifying in an appropriate way the formulae showed in Lemma 2.1 (this can be easily done by using that the matrix A_nA is hermitian), the proof of Theorem 2.4 exactly works as that of Theorem 2.3.

3. The quadrature formula revisited. Quadrature formulae have been found recently for orthonormal matrix polynomials by Sinap and Van Assche (see [SV]) and Duran and Lopez-Rodriguez (see [DL]). Here and using other different approach, we improve these quadrature formulae by giving a closed expression for the matrix coefficients which appear in the formula. This expression will be the key to prove the matrix version of Markov's Theorem.

THEOREM 3.1. *Let n be a nonnegative integer. We write $x_{n,k}$ ($k = 1, \dots, m$) for the different zeros of the matrix polynomial P_n ordered in increasing size (hence $m \leq nN$) and $\Gamma_{n,k}$ for the matrices*

$$(3.1) \quad \Gamma_{n,k} = \frac{1}{\left(\det(P_n(t))\right)^{(l_k)}(x_{n,k})} \left(\text{Adj}(P_n(t))\right)^{(l_k-1)}(x_{n,k}) Q_n(x_{n,k}), \quad k = 1, \dots, m,$$

where l_k is the multiplicity of $x_{n,k}$.

(1) *For any polynomial P with $\text{dgr}(P) \leq 2n - 1$ the following formula holds*

$$(3.2) \quad \int P(t) dW(t) = \sum_{k=1}^m P(x_{n,k}) \Gamma_{n,k}.$$

(2) *The matrices $\Gamma_{n,k}$ are positive semidefinite matrices of rank l_k , $k = 1, \dots, m$.*

It is worth noting that the smaller the multiplicity of a zero $x_{n,k}$ is, the bigger the singularity of the matrix $\Gamma_{n,k}$ is. For instance, if $x_{n,k}$ is simple, then $\Gamma_{n,k}$ has rank one.

PROOF. To prove (1) of Theorem 3.1 we use the following surprising fact:

Although the zeros of P_n (i.e., the zeros of $\det(P_n)$) can be of multiplicity bigger than one (at most N) the decomposition

$$(3.3) \quad P(z)P_n^{-1}(z) = \sum_{k=1}^m \frac{C_{n,k}}{z - x_{n,k}} \quad \text{if } \text{dgr } P \leq n - 1,$$

is always possible.

Part (4) of Theorem 2.3 and Lemma 2.2 show the reason for this remarkable property: indeed, if a is a zero of multiplicity p for P_n then

$$\begin{aligned} \text{Adj}(P_n(a)) &= \left(\text{Adj}(P_n(t))\right)'(a) = \dots = \left(\text{Adj}(P_n(t))\right)^{(p-2)}(a) = \theta, \quad \text{and} \\ &\left(\text{Adj}(P_n(t))\right)^{(p-1)}(a) \neq \theta. \end{aligned}$$

This means that a is a zero of multiplicity at least $p - 1$ of each entry of $\text{Adj}(P_n(z))$. Hence a is a zero of multiplicity at least $p - 1$ of $P(z) \text{Adj}(P_n(z))$ for any matrix polynomial P , and so, if $\text{dgr}(P) \leq n - 1$, (i.e., $\text{dgr}(P \text{Adj}(P_n)) \leq nN - 1 < \text{dgr}(\det(P_n))$), the expression (3.3) holds for

$$P(z)P_n^{-1}(z) = \frac{P(z) \text{Adj}(P_n(z))}{\det(P_n(z))}.$$

We now prove (1) of Theorem 3.1. Let P be a matrix polynomial of degree less than or equal to $2n - 1$. Since P_n is a polynomial with non-singular leading coefficient, we can write ([G] p. 78)

$$(3.4) \quad P(t) = C(t)P_n(t) + R(t),$$

where $C(t)$ and $R(t)$ are matrix polynomials with $\text{dgr}(R) \leq n - 1$.

Then we have

$$(3.5) \quad P(t)P_n^{-1}(t) = C(t) + R(t)P_n^{-1}(t),$$

when t is not a zero of P_n . Since $\text{dgr}(R(t)) \leq n - 1$ we can write

$$R(t)P_n^{-1}(t) = \sum_{k=1}^m C_{n,k} \frac{1}{t - x_{n,k}},$$

where the matrices $C_{n,k}$ are given by

$$C_{n,k} = \frac{1}{\left(\det(P_n(t))\right)^{(l_k)}(x_{n,k})} R(x_{n,k}) \left(\text{Adj}(P_n(t))\right)^{(l_k-1)}(x_{n,k}).$$

Since $P_n(x_{n,k}) \left(\text{Adj}(P_n(t))\right)^{(l_k-1)}(x_{n,k}) = \theta$ (see part (4) of Theorem 2.3), (3.4) gives that the matrices $C_{n,k}$ can be written in the following way

$$(3.6) \quad C_{n,k} = \frac{1}{\left(\det(P_n(t))\right)^{(l_k)}(x_{n,k})} P(x_{n,k}) \left(\text{Adj}(P_n(t))\right)^{(l_k-1)}(x_{n,k}).$$

Then (3.5) gives

$$P(t) = C(t)P_n(t) + \sum_{k=1}^m C_{n,k} \frac{P_n(t)}{t - x_{n,k}}.$$

From (3.6) and part (4) of Theorem 2.3, we have

$$P(t) = C(t)P_n(t) + \sum_{k=1}^m C_{n,k} \frac{P_n(t) - P_n(x_{n,k})}{t - x_{n,k}}.$$

From the definition of the polynomial Q_n it follows that

$$\int P(t) dW(t) = \int C(t)P_n(t) dW(t) + \sum_{k=1}^m C_{n,k} Q_n(x_{n,k}).$$

But $\text{dgr}(C) < n$, so the orthonormality of P_n gives that $\int C(t)P_n(t) dW(t) = \theta$. Hence, we have

$$\int P(t) dW(t) = \sum_{k=1}^m C_{n,k} Q_n(x_{n,k}).$$

Then, part (1) of Theorem 3.1 follows from the definition of the matrices $\Gamma_{n,k}$ and $C_{n,k}$ (see (3.1) and (3.6)).

We now prove part (2) of Theorem 3.1:

We proceed in several steps:

STEP 1. The matrices $\Gamma_{n,k}$ ($k = 1, \dots, m$) are hermitian.

PROOF. If we prove that the matrix $P_n^{-1}(t)Q_n(t)$ is hermitian when $t \in \mathbb{R}$ is not a zero of P_n , Step 1 would follow from the decomposition

$$P_n^{-1}(t)Q_n(t) = \sum_{k=1}^m \Gamma_{n,k} \frac{1}{t - x_{n,k}}.$$

The orthogonality of P_n with respect to any other polynomial of degree less than n gives that

$$\begin{aligned} P_n^{-1}(t)Q_n(t) &= \int P_n^{-1}(t) \frac{P_n(t) - P_n(x)}{t - x} dW(x) \\ &= \int P_n^{-1}(t) \frac{P_n(t) - P_n(x)}{t - x} dW(x) (P_n^*(t) - P_n^*(x)) (P_n^*)^{-1}(t) \\ &= \int P_n^{-1}(t) (P_n(t) - P_n(x)) dW(x) \frac{P_n^*(t) - P_n^*(x)}{t - x} (P_n^*)^{-1}(t) \\ &= \int dW(x) \frac{P_n^*(t) - P_n^*(x)}{t - x} (P_n^*)^{-1}(t) \\ &= Q_n^*(t) (P_n^*)^{-1}(t), \end{aligned}$$

that is, $P_n^{-1}(t)Q_n(t)$ is hermitian when $t \in \mathbb{R}$ is not a zero of P_n . ■

STEP 2. $\text{rank}(\Gamma_{n,k}) = l_k, k = 1, \dots, m.$

PROOF. Theorem 2.3 gives that $\text{rank}\left(\text{Adj}(P_n(t))\right)^{(l_k-1)}(x_{n,k}) = l_k$. Since the left eigenvectors of $\left(\text{Adj}(P_n(t))\right)^{(l_k-1)}(x_{n,k})$ associated to 0 are left eigenvectors of $\Gamma_{n,k}$ associated to 0 (see (3.1)), it will be enough to prove that left eigenvectors of $\Gamma_{n,k}$ associated to 0 are left eigenvectors of $\left(\text{Adj}(P_n(t))\right)^{(l_k-1)}(x_{n,k})$ associated to 0.

Let u be for which $v = u\left(\text{Adj}(P_n(t))\right)^{(l_k-1)}(x_{n,k}) \neq \theta$. Part (3) of Theorem 2.3 shows that $v \in L(x_{n,k}, P_n)$. Then the formula (2.6) gives

$$v(Q_n(x_{n,k})P_{n-1}^*(x_{n,k}) - P_n(x_{n,k})Q_{n-1}^*(x_{n,k})) = vQ_n(x_{n,k})P_{n-1}^*(x_{n,k}) = vA_n^{-1}.$$

Since the matrix A_n^{-1} is non-singular, we have that $vQ_n(x_{n,k}) \neq \theta$, and taking into account the definition of v and $\Gamma_{n,k}$, we obtain $u\Gamma_{n,k} \neq \theta$. And Step 2 is proved. ■

We take vectors v_1, \dots, v_{l_k} satisfying the following properties:

- (1) v_i is a left eigenvector of the matrix $\Gamma_{n,k}$ associated to an eigenvalue $\alpha_i \neq 0, i = 1, \dots, l_k.$
- (2) $v_i v_j^* = \delta_{ij}, i, j = 1, \dots, l_k.$

STEP 3. For $i = 1, \dots, l_k$, constants $\beta_i \neq 0$ exist such that

$$v_i \left(\text{Adj}(P_n(t))\right)^{(l_k-1)}(x_{n,k}) = \beta_i v_i P_{n-1}^*(x_{n,k}) A_n.$$

PROOF. From part (3) of Theorem 2.3, we have that $v_i \left(\text{Adj}(P_n(t))\right)^{(l_k-1)}(x_{n,k}) \in L(x_{n,k}, P_n)$. Since the matrix $P_{n-1}^*(x_{n,k}) A_n$ defines an isomorphism from $R(x_{n,k}, P_n)$ into $L(x_{n,k}, P_n)$ (Theorem 2.3 (2)), we have that

$$(3.7) \quad v_i \left(\text{Adj}(P_n(t))\right)^{(l_k-1)}(x_{n,k}) = u_i P_{n-1}^*(x_{n,k}) A_n,$$

for certain $u_i \in R(x_{n,k}, P_n)$. Since $Q_n(x_{n,k})$ is the inverse mapping of $P_{n-1}^*(x_{n,k})A_n$, we have from (3.7) that

$$v_i \left(\text{Adj}(P_n(t)) \right)^{(l_{k-1})} (x_{n,k}) Q_n(x_{n,k}) = u_i.$$

Then, from the definition of the matrix $\Gamma_{n,k}$ (see (3.1)), we deduce that

$$v_i \Gamma_{n,k} = \frac{1}{\left(\det P_n(t) \right)^{(l_k)} (x_{n,k})} u_i.$$

Since $v_i \Gamma_{n,k} = \alpha_i v_i$, it is enough to take $\beta_i = \alpha_i \left(\det P_n(t) \right)^{(l_k)} (x_{n,k})$. ■

Finally, we prove that $v_i \Gamma_{n,k} v_i^* > 0$, i.e., the matrix $\Gamma_{n,k}$ is positive semidefinite. In Step 3, we have proved that $v_i = \frac{1}{\beta_i} u_i$, for certain $u_i \in R(x_{n,k}, P_n)$, then we deduce that $v_i \in R(x_{n,k}, P_n)$. From the formula

$$\begin{aligned} & \left(\text{Adj}(P_n(t)) \right)^{(l_k)} (x_{n,k}) P_n(x_{n,k}) + \left(\text{Adj}(P_n(t)) \right)^{(l_{k-1})} (x_{n,k}) P_n'(x_{n,k}) \\ &= \left(\det P_n(t) \right)^{(l_k)} (x_{n,k}) I, \end{aligned}$$

and Step 3, we obtain

$$(3.8) \quad \beta_i v_i P_{n-1}^*(x_{n,k}) A_n P_n'(x_{n,k}) v_i^* = \left(\det P_n(t) \right)^{(l_k)} (x_{n,k}).$$

Step 3, (3.8), (2.3) and (2.5) give that

$$\begin{aligned} v_i \Gamma_{n,k} v_i^* &= \frac{v_i \left(\text{Adj}(P_n(x_{n,k})) \right)^{(l_{k-1})} Q_n(x_{n,k}) v_i^*}{\left(\det P_n(t) \right)^{(l_k)} (x_{n,k})} \\ &= \frac{\beta_i v_i P_{n-1}^*(x_{n,k}) A_n Q_n(x_{n,k}) v_i^*}{\beta_i v_i P_{n-1}^*(x_{n,k}) A_n P_n'(x_{n,k}) v_i^*} \\ &= \frac{v_i \left(P_{n-1}^*(x_{n,k}) A_n Q_n(x_{n,k}) - P_n^*(x_{n,k}) A_n^* Q_{n-1}(x_{n,k}) \right) v_i^*}{v_i \left(P_{n-1}^*(x_{n,k}) A_n P_n'(x_{n,k}) - P_n^*(x_{n,k}) A_n^* P_{n-1}'(x_{n,k}) \right) v_i^*} \\ &= \frac{1}{v_i \left(\sum_{j=0}^{n-1} P_j^*(x_{n,k}) P_j(x_{n,k}) \right) v_i^*} > 0. \end{aligned}$$

And the theorem is proved. ■

Theorem 3.1 can also be extended for perturbations of the orthonormal polynomials. Indeed, let A be a matrix satisfying that $A_n A$ is hermitian, and let us consider the polynomial $P_n(t) - AP_{n-1}(t)$. Then, Theorem 3.1 holds for this polynomial, i.e.:

THEOREM 3.2. *Let n be a nonnegative integer. We write $x_{n,k}$ ($k = 1, \dots, m$) for the different zeros of the matrix polynomial $P_n - AP_{n-1}$ (where $A_n A = A^* A_n^*$) ordered in increasing size (hence $m \leq nN$) and $\Gamma_{n,k}$ for the matrices*

$$\begin{aligned} \Gamma_{n,k} &= \frac{1}{\left(\det(P_n(t) - AP_{n-1}(t)) \right)^{(l_k)} (x_{n,k})} \left(\text{Adj}(P_n(t) - AP_{n-1}(t)) \right)^{(l_{k-1})} (x_{n,k}) \\ &\cdot \left(Q_n(x_{n,k}) - A Q_{n-1}(x_{n,k}) \right), \quad k = 1, \dots, m, \end{aligned}$$

where l_k is the multiplicity of $x_{n,k}$.

(1) For any polynomial P with $\text{dgr}(P) \leq 2n - 2$ the following formula holds

$$(3.9) \quad \int P(t) dW(t) = \sum_{k=1}^m P(x_{n,k})\Gamma_{n,k}.$$

(2) The matrices $\Gamma_{n,k}$ are positive semidefinite matrices of rank l_k , for $k = 1, \dots, m$.

Using Theorem 2.4 instead of Theorem 2.3 the proof of Theorem 3.2 exactly works as that of Theorem 3.1. It is worth noting that (3.9) can not be extended for $\text{dgr}(P) \leq 2n - 1$ because of

$$\int C(t)(P_n(t) - AP_{n-1}(t)) dW(t) = \theta$$

can only be guaranteed when $\text{dgr}(C) < n - 1$.

REMARK 3.3. In the scalar case, the coefficients of the quadrature formula satisfy a remarkable property called Markov-Stieltjes inequalities:

$$\sum_{j=1}^{k-1} \lambda_{n,j} < \int_a^{x_{n,k}} d\mu(x) < \sum_{j=1}^k \lambda_{n,j}, \quad k = 1, \dots, n,$$

where $x_{n,1}, \dots, x_{n,n}$ are the zeros of the orthonormal polynomial p_n , and $\lambda_{n,k}$ are the quadrature weights given by

$$\lambda_{n,k} = \frac{1}{\sum_{j=1}^{n-1} |p_j(x_{n,k})|^2}.$$

This property was conjectured by Chebyshev in 1874, and proved independently by Markov and Stieltjes ten years later. We present here a counterexample showing that these Markov-Stieltjes inequalities are no longer true for orthogonal matrix polynomials. The reason is the highly singular character of quadrature weights in the matrix case.

Indeed, let $\mu_1, \mu_2, (p_{1,n})_n$ and $(p_{2,n})_n$ be two positive measures and their sequences of orthonormal polynomials. We define the matrix of measures W by

$$W = \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}.$$

Then it is clear that its sequence of orthonormal matrix polynomials is

$$P_n(t) = \begin{pmatrix} p_{1,n} & 0 \\ 0 & p_{2,n} \end{pmatrix}, \quad n \geq 0.$$

We assume that the measures μ_1 and μ_2 have the same support $[a, b]$ and that $p_{1,n}$ and $p_{2,n}$ interlace their zeros. Then an easy computation gives that the coefficients in the matrix quadrature formula are

$$\Gamma_{n,k} = \begin{cases} \begin{pmatrix} \lambda_{1,n,\frac{k+1}{2}} & 0 \\ 0 & 0 \end{pmatrix} & \text{for } k = 1, 3, \dots, 2n - 1, \\ \begin{pmatrix} 0 & 0 \\ 0 & \lambda_{2,n,\frac{k}{2}} \end{pmatrix} & \text{for } k = 2, 4, \dots, 2n, \end{cases}$$

where $\lambda_{1,n,k}, \lambda_{2,n,k}$ ($k = 1, \dots, n$) are the quadrature weights for the measures μ_1, μ_2 respectively. Then, since

$$\Gamma_{n,1} = \begin{pmatrix} \lambda_{1,n,1} & 0 \\ 0 & 0 \end{pmatrix},$$

is a singular matrix, it is clear that

$$\int_a^{x_{1,n,1}} dW(x) \not\prec \Gamma_{n,1},$$

because of the non-singularity of $\int_a^{x_{1,n,1}} dW(x)$.

4. Markov's theorem for orthogonal matrix polynomials. We are now ready to prove the matrix version of Markov's theorem

PROOF OF THEOREM 1.1. First of all, we prove that

$$\lim_{n \rightarrow \infty} P_n^{-1}(z)Q_n(z) = \int \frac{dW(t)}{z-t}$$

for $z \in \mathbb{C} \setminus \Gamma$ (see (1.1) for the definition of Γ).

We consider the sequence of discrete matrices of measures $(\mu_n)_n$ defined by

$$\mu_n = \sum_{k=1}^m \delta_{x_{n,k}} \Gamma_{n,k}, \quad n \geq 0,$$

where $x_{n,k}$ ($k = 1, \dots, m$) are the different zeros of the matrix polynomial P_n ordered in increasing size, and $\Gamma_{n,k}$ are the positive semidefinite matrices defined by (3.1) and which appear in the quadrature formula (3.2).

From the definition of these matrices (see (3.1)), it follows straightforwardly that

$$(4.1) \quad P_n^{-1}(z)Q_n(z) = \sum_{k=1}^m \Gamma_{n,k} \frac{1}{z-x_{n,k}} = \int \frac{d\mu_n(t)}{z-t},$$

if z is not a zero of P_n . Taking into account (4.1), it will be enough to prove that

$$\lim_{n \rightarrow \infty} \int \frac{d\mu_n(t)}{z-t} = \int \frac{dW(t)}{z-t}$$

for $z \in \mathbb{C} \setminus \Gamma$. If not, we find a complex number $z \in \mathbb{C} \setminus \Gamma$, an increasing sequence of nonnegative integers $(n_m)_m$ and a positive constant C for which

$$(4.2) \quad \left\| \int \frac{dW(t)}{z-t} - \int \frac{d\mu_{n_m}(t)}{z-t} \right\|_2 \geq C > 0, \quad m \geq 0,$$

where we write $\| \cdot \|_2$ for the spectral norm of a matrix.

If we proceed as in Section 2 of [DL], taking two increasing sequences $(a_k)_k, (b_k)_k$ for which $a_k, b_k \rightarrow +\infty$, we can obtain (by using Banach-Alaoglu's theorem) a subsequence $(l_m)_m$ from $(n_m)_m$, and a positive definite matrix of measures ν such that for $k \geq 0$

$$(4.3) \quad \lim_{m \rightarrow \infty} \int_{-a_k}^{b_k} f(t) d\mu_{l_m}(t) = \int_{-a_k}^{b_k} f(t) d\nu(t),$$

for any continuous matrix function f defined in $[-a_k, b_k]$. Now it is easy to prove (see Section 2 of [DL] for more details) that the k -th moment of the matrix of measures ν is the limit of the k -th moment of the matrices of measures μ_{l_m} . But the quadrature formula (3.2) gives that the k -th moment of the matrix of measures μ_{l_m} ($k \leq 2l_m - 1$) is precisely the k -th moment of the matrix of measures W . Since W is a determinate matrix of measures, we conclude that $\nu = W$. Then (4.3) gives that for $k \geq 0$

$$(4.4) \quad \lim_{m \rightarrow \infty} \int_{-a_k}^{b_k} f(t) d\mu_{l_m}(t) = \int_{-a_k}^{b_k} f(t) dW(t).$$

By taking k and then l_m big enough, from (4.2) and (4.4) we obtain

$$(4.5) \quad \frac{C}{2} \leq \left\| \int_{-\infty}^{-a_k} \frac{d\mu_{l_m}(t)}{z-t} + \int_{b_k}^{+\infty} \frac{d\mu_{l_m}(t)}{z-t} \right\|_2.$$

We write S_0 for the first moment of the matrices of measures μ_{l_m} which is the first moment of the matrix of measures W . Then we have from (4.5) and the definition of the spectral norm that

$$\begin{aligned} \frac{C}{2} &\leq \max\left(\frac{1}{|z-a_k|}, \frac{1}{|z-b_k|}\right) \left\| \int_{-\infty}^{-a_k} d\mu_{l_m}(t) + \int_{b_k}^{+\infty} d\mu_{l_m}(t) \right\|_2 \\ &\leq \max\left(\frac{1}{|z-a_k|}, \frac{1}{|z-b_k|}\right) \left\| \int_{\mathbb{R}} d\mu_{l_m}(t) \right\|_2 \\ &\leq \max\left(\frac{1}{|z-a_k|}, \frac{1}{|z-b_k|}\right) \|S_0\|_2. \end{aligned}$$

But this implies $C = 0$, and therefore, (4.2) is not possible.

We now prove that the analytic functions which form the entries of the matrix $\int \frac{d\mu_n(t)}{z-t}$ are uniformly bounded in compact sets of $\mathbb{C} \setminus \Gamma$. Then, the uniform convergence in compact set of $\mathbb{C} \setminus \Gamma$ will follow from Stieltjes-Vitali's theorem.

Given a compact $K \subset \mathbb{C} \setminus \Gamma$, we notice that $K \cap M_N = \emptyset$, for N sufficiently big, and then there exists $C > 0$ such that

$$\left| \frac{1}{z-t} \right| \leq C \quad \text{for } z \in K \text{ and } t \in M_N.$$

Then, for $n \geq N$ and $v \in \mathbb{C}^N$, the positive definiteness of μ_n gives that

$$\left| v \int \frac{d\mu_n(t)}{z-t} v^* \right| = \int \frac{v d\mu_n(t) v^*}{|z-t|} \leq C \int v d\mu_n(t) v^* = v S_0 v^*,$$

and now it is easy to finish. ■

It is worth to note that the order in which the polynomials P_n^{-1} and Q_n appear multiplied in the matrix version of Markov's theorem, *i.e.*, $P_n^{-1}(z)Q_n(z)$, is essential to guarantee the validity of this result. Indeed, let W be a positive definite matrix of measures and $(P_n)_n, (Q_n)_n$ its orthonormal matrix polynomials and polynomials of the second kind, respectively. Let us consider a non-singular matrix C , and the positive definite matrix of measures defined by $R = CWC^*$. It is clear that $(P_n C^{-1})_n$ and $(Q_n C^*)$ are, respectively,

the sequence of orthonormal matrix polynomials and the sequence of polynomials of the second kind for R . Hence, if Markov's theorem held for $Q_n(z)P_n^{-1}(z)$, we would have

$$\lim_n Q_n(z)P_n^{-1}(z) = \int \frac{dW(t)}{z-t}$$

$$\lim_n Q_n(z)C^*CP_n^{-1}(z) = C \int \frac{dW(t)}{z-t} C^*,$$

which, in general, is clearly false.

Markov's theorem can be extended for perturbations of the orthonormal polynomials: Indeed, let A be a matrix satisfying that A_nA is hermitian, and let us consider the polynomial $P_n(t) - AP_{n-1}(t)$. We write $\Delta_{n,A}$ for the set of zeros of the matrix polynomial $P_n - AP_{n-1}$, and

$$\Gamma_A = \cap_{N \geq 0} M_{N,A}, \quad \text{where } M_{N,A} = \overline{\cup_{n \geq N} \Delta_{n,A}}.$$

Then

THEOREM 4.1. *Assume that W is determinate and that $A_nA = A^*A_n^*$ for $n \geq 0$. Then*

$$\lim_{n \rightarrow \infty} (P_n(z) - AP_{n-1}(z))^{-1} (Q_n(z) - AQ_{n-1}(z)) = \int \frac{dW(t)}{z-t} \quad \text{for } z \in \mathbb{C} \setminus \Gamma,$$

and the convergence is uniform for z in compact subsets of $\mathbb{C} \setminus \Gamma_A$.

Using Theorem 3.2 instead of Theorem 3.1 the proof of Theorem 4.1 exactly works as that of Theorem 1.1.

5. The k -th associated polynomials. For $k \geq 1$, the k -th associated polynomials $(P_n^{[k]})_n$ are defined by the formula

$$(5.1) \quad P_n^{[k]}(x) = \int \frac{P_{n+k}(x) - P_{n+k}(t)}{x-t} dW(t)P_{k-1}^*(t), \quad n \geq 0.$$

The orthogonality of the polynomials $(P_n)_n$ shows that the degree of $P_n^{[k]}$ is just n .

Associated polynomials already appear in Stieltjes' fundamental work [S]. (See the survey [V] for more details).

These k -th associated polynomials satisfy the following two recurrence formulae

$$(5.2) \quad xP_n^{[k]}(x) = A_{n+k+1}P_{n+1}^{[k]}(x) + B_{n+k}P_n^{[k]}(x) + A_{n+k}^*P_{n-1}^{[k]}(x), \quad n \geq 0,$$

and

$$(5.3) \quad xP_n^{[k]}(x) = P_{n-1}^{[k+1]}(x)A_k^* + P_n^{[k]}(x)B_{k-1}^* + P_{n+1}^{[k-1]}(x)A_{k-1}$$

with initial condition $A_0 = I$, $P_{-1}^{[k]}(x) = \theta$ and $P_0^{[k]}(x) = A_k^{-1}$. Indeed, we have

$$xP_n^{[k]}(x) = \int \frac{xP_{n+k}(x) - tP_{n+k}(t)}{x-t} dW(t)P_{k-1}^*(t)$$

$$- \int \frac{(x-t)P_{n+k}(t)}{x-t} dW(t)P_{k-1}^*(t),$$

but also

$$xP_n^{[k]}(x) = \int \frac{P_{n+k}(x) - P_{n+k}(t)}{x-t} dW(t) t P_{k-1}^*(t) \\ + \int \frac{P_{n+k}(x) - P_{n+k}(t)}{x-t} dW(t) (x-t) P_{k-1}^*(t).$$

Formulae (5.2) and (5.3) now follow from the three term recurrence formula for $(P_n)_n$ and the orthogonality of these polynomials.

From the formula (5.3) and the three term recurrence formula for $(P_n)_n$ one obtains the following *quasi* Christoffel-Darboux formula

THEOREM 5.1. *The following formulae are valid:*

$$(x-t) \sum_{k=1}^n P_{n-k}^{[k]}(x) P_{k-1}(t) = P_n(x) - P_n(t).$$

and

$$\sum_{k=1}^n P_{n-k}^{[k]}(x) P_{k-1}(x) = P_n'(x).$$

PROOF. Indeed, it is enough to write

$$\frac{P_n(x) - P_n(t)}{(x-t)} = \sum_{k=1}^n A_{k,n} P_{k-1}(t),$$

and compute the matrix coefficients $A_{k,n}$ by using the orthonormality of $(P_n)_n$. ■

The following generalization of Markov's Theorem is a consequence of Theorem 1.1, the formula (5.3) and the three term recurrence formula for $(P_n)_n$ (see [V] for the scalar version):

THEOREM 5.2. *Assume that W is determinate. Then for $k \geq 1$*

$$\lim_{n \rightarrow \infty} P_n^{-1}(z) P_{n-k}^{[k]}(z) = \int dW(t) \frac{P_{k-1}^*(t)}{z-t} \quad \text{for } z \in \mathbb{C} \setminus \Gamma,$$

and the convergence is uniform for z in compact subsets of $\mathbb{C} \setminus \Gamma$.

PROOF. It follows in a straightforward manner proceeding by induction on k and using the formula (5.3). ■

REFERENCES

- [AN] A. I. Aptekarev and E. M. Nikishin, *The scattering problem for a discrete Sturm-Liouville operator*, Mat. Sb. **121**(1983), 327–358; Math. USSR-Sb. **49**(1984), 325–355.
- [B] C. Berg, *Markov's theorem revisited*, J. Approx. Theory **78**(1994), 260–275.
- [D1] A. J. Duran, *A generalization of Favard's Theorem for polynomials satisfying a recurrence relation*, J. Approx. Theory **74**(1993), 83–109.
- [D2] ———, *On orthogonal polynomials with respect to a positive definite matrix of measures*, Canad. J. Math. **47**(1995), 88–112.
- [DL] A. J. Duran and P. Lopez-Rodriguez, *Orthogonal matrix polynomials: Zeros and Blumenthal's theorem*, J. Approx. Theory **84**(1996), 96–118.

- [DV] A. J. Duran and W. van Assche, *Orthogonal matrix polynomials and higher order recurrence relations*, *Linear Algebra Appl.* **219**(1995), 261–280.
- [G] F. R. Gantmacher, *The theory of matrices*, Chelsea Publishing Company, New York, **1**(1960).
- [M] A. Markov, *Deux démonstrations de la convergence de certaines fractions continues*, *Acta Math.* **19**(1895), 93–104.
- [SV] A. Sinap and W. van Assche, *Polynomial interpolation and Gaussian quadrature for matrix valued function*, *Linear Algebra Appl.* **207**(1994), 71–114.
- [S] T. J. Stieltjes, *Recherches sur les fractions continues*, *Ann. Fac. Sci. Toulouse Math.* **8**(1894), 1–22; [9] (1895), 5–47.
- [Z] D. Zhani, *Probleme des moments matriciels sur la droite: construction d'une famille de solutions et questions d'unicite*, *Publications du Departement de Mathematiques, Universite Claude-Bernard-Lyon I, Nouvelle serie D*(1984), 1–84.

Departamento de Análisis Matemático
Universidad de Sevilla, Apdo. 1160.
41080-Sevilla
Spain