Can. J. Math. Vol. **49** (4), 1997 pp. 708–721

DENSITY QUESTIONS FOR THE TRUNCATED MATRIX MOMENT PROBLEM

ANTONIO J. DURAN AND PEDRO LOPEZ-RODRIGUEZ

ABSTRACT. For a truncated matrix moment problem, we describe in detail the set of positive definite matrices of measures μ in V_{2n} (this is the set of solutions of the problem of degree $2n$) for which the polynomials up to degree n are dense in the corresponding space $L^2(\mu)$. These matrices of measures are exactly the extremal measures of the set V_n.

1. **Introduction.** Let $\nu = (\nu_{i,j})_{1\leq i,j\leq N}$ be a positive definite matrix of measures with finite moments S_k of any order and having a sequence $(P_n)_{n=0}^{\infty}$ of orthonormal matrix polynomials, P_n of degree n and with non-singular leading coefficient. For the sake of simplicity we will assume $P_0(t) = I$.

These polynomials satisfy a three-term recurrence relation of the form

$$tP_n(t) = A_{n+1}P_{n+1}(t) + B_nP_n(t) + A_n^*P_{n-1}(t), \quad n \geq 0, \tag{1.1}$$

(A_n and B_n being $N \times N$ matrices such that $\det(A_n) \neq 0$ and $B_n^* = B_n$), with initial condition $P_{-1}(t) = \theta$ (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. For instance, here θ is the $N \times N$ null matrix). It is well-known that this recurrence relation is equivalent to the orthogonality with respect to a positive definite matrix of measures: this is the matrix version of Favard's Theorem (see [AN], [D1] and [DL1]).

We denote by $Q_n(t)$ the corresponding sequence of polynomials of the second kind,

$$Q_n(t) = \int_{\mathbb{R}} \frac{P_n(t) - P_n(x)}{t - x}\, d\nu(x), \quad n \geq 0,$$

which also satisfy the recurrence relation (1.1), with initial conditions $Q_0(t) = \theta$ and $Q_1(t) = A_1^{-1}$.

For $n \geq 0$ we denote by V_n the set of positive definite matrices of measures whose moments up to degree n are finite and coincide with those of ν, that is:

$$V_n = \Big\{ \mu = (\mu_{i,j})_{1\leq i,j\leq N} : \int_{\mathbb{R}} t^p d\mu_{i,j} = \int_{\mathbb{R}} t^p\, d\nu_{i,j}, \text{ for } 0 \leq p \leq n \text{ and } 1 \leq i,j \leq N \Big\}.$$

For μ a positive definite matrix of measures, the space $L^2(\mu)$ is defined as the set of $N \times N$ matrix functions $f\colon \mathbb{R} \longrightarrow M_{N\times N}(\mathbb{C})$ such that $\tau\big(f(t)M(t)f(t)^*\big) \in L^1(\tau\mu)$, where

This work has been partially supported by DGICYT ref. PB93-0926.
Received by the editors October 2, 1995; revised July 9, 1996.
AMS subject classification: 42C05, 44A60.

M is the Radon-Nikodym derivative of μ with respect to its trace $(\tau\mu)$ (for a matrix $A = (a_{i,j})_{1\le i,j\le N}$, we denote τA for its trace, *i.e.* $\tau A = \sum_{i=1}^{N} a_{i,i}$):

$$M = (m_{i,j})_{i,j=1}^{N} = \left(\frac{d\mu_{i,j}}{d\tau\mu}\right)_{1\le i,j\le N}.$$

The space $L^2(\mu)$ is endowed with the norm

$$\|f\|_{2,\mu} = \left\|\tau\big(f(t)M(t)f(t)^*\big)^{\frac{1}{2}}\right\|_{2,\tau\mu} = \left(\int_{\mathbb{R}} \tau\big(f(t)M(t)f(t)^*\big)\, d\tau\mu(t)\right)^{\frac{1}{2}}$$

and is a Hilbert space. The duality works as for the scalar case (see [R] or [DL2] for more details. For the definition of the L^p spaces associated to μ see [DL2]).

We stress that since we only impose the matrices of measures in V_{2n} to have finite moments up to degree $2n$, for $\mu \in V_{2n}$ we can guarantee only that the polynomials up to degree n belong to the corresponding space $L^2(\mu)$. In any case, the polynomials $(P_k)_{k=0,\dots,n}$ are orthonormal with respect to any measure in V_{2n}.

In this paper we characterize the matrices of measures μ of V_{2n-2} for which the polynomials up to degree $n-1$ are dense in $L^2(\mu)$: μ is an extremal matrix (in the sense of convexity) in V_{n-1}. Furthermore, this is the case if and only if μ is a discrete matrix of measures whose support (the support of a positive definite matrix of measures $\mu = (\mu_{i,j})_{1\le i,j\le N}$ is the support of its trace measure $\tau\mu = \sum_{i=1}^{N} \mu_{i,i}$) is the set of zeros of $\big(P_n(\lambda) - AP_{n-1}(\lambda)\big)$ (the zeros of a matrix polynomial $P(t)$ are the zeros of the determinant of $P(t)$), where the matrix A makes A_nA hermitian:

THEOREM 1. *For a matrix of measures μ in V_{2n-2} ($n \ge 1$) the following statements are equivalent:*

(1) μ is an extremal measure of the set V_{n-1}.

(2) The matrix polynomials of degree less than or equal to $n-1$ are dense in the space $L^2(\mu)$.

(3) There exists an $N \times N$ matrix A such that $A_nA = A^*A_n^*$ and for which $\mu = \sum_{i=1}^{q} G_i\delta_{x_i}$, where x_i, $i = 1,\dots,q$, are the different zeros of the polynomial $\det\big(P_n(\lambda) - AP_{n-1}(\lambda)\big)$ and G_i are the matrices which appear in the simple fraction decomposition

$$\big(P_n(\lambda) - AP_{n-1}(\lambda)\big)^{-1}\big(Q_n(\lambda) - AQ_{n-1}(\lambda)\big) = \sum_{i=1}^{q} \frac{G_i}{\lambda - x_i}. \tag{1.2}$$

The numbers x_i are real and the matrices G_i are positive semidefinite, $i = 1,\dots,q$.

The comparison between Theorem 1 and its scalar version points out the existence of important divergences created by the matrix structure. We devote the third section of this paper to study one of these divergences: in the scalar case, the points of the support of the extremal measures support the maximum mass in V_{2n-2}; indeed, if $\mu \in V_{2n-2}$ is extremal in V_{n-1}, it is well-known (see for example [A, p. 61]) that if $x_i \in \operatorname{supp}(\mu)$, then

$$\mu(\{x_i\}) = \frac{1}{\sum_{k=0}^{n-1} p_k^2(x_i)}$$

https://doi.org/10.4153/CJM-1997-034-5 Published online by Cambridge University Press

and $\mu(\{x_i\}) \geq \nu(\{x_i\})$, for any $\nu \in V_{2n-2}$.

This property is not true in the matrix case. Indeed, the matrices in V_{2n-2} extremal in V_{n-1} may not support the maximum mass in V_{2n-2}. This happens exactly when the mass is a non-singular matrix, in which case the polynomial $P_n(\lambda) - AP_{n-1}(\lambda)$ has a zero of maximum multiplicity at the corresponding point x_i. In that case A is given by $A = P_n(x_i)P_{n-1}^{-1}(x_i)$ (see Theorem 3.1).

2. **The structure Theorem.** We prove here Theorem 1. Before beginning on the proof we need some formulae and preliminary results.

LEMMA 2.1. *(1) The Christoffel-Darboux formula*

$$P_{n-1}^*(u)A_nP_n(v) - P_n^*(u)A_n^*P_{n-1}(v) = (v-u)\sum_{k=0}^{n-1} P_k^*(u)P_k(v), \quad u,v \in \mathbb{C}, \tag{2.1}$$

(2) with its particular case

$$P_{n-1}^*(z)A_nP_n'(z) - P_n^*(z)A_n^*P_{n-1}'(z) = \sum_{k=0}^{n-1} P_k^*(z)P_k(z), \quad z \in \mathbb{C}. \tag{2.2}$$

(3) The Green formula

$$P_{n-1}^*(u)A_nQ_n(v) - P_n^*(u)A_n^*Q_{n-1}(v) = I + (v-u)\sum_{k=0}^{n-1} P_k^*(u)Q_k(v), \quad u,v \in \mathbb{C}. \tag{2.3}$$

(4) If A is a $N \times N$ matrix such that $A_nA = A^*A_n^*$, and $\lambda \in \mathbb{C}$ is not a zero of $\det\big(P_n(\lambda) - AP_{n-1}(\lambda)\big)$ then

$$\begin{aligned}\big(Q_n^*(\lambda) - Q_{n-1}^*(\lambda)A^*\big)&\big(P_n^*(\lambda) - P_{n-1}^*(\lambda)A^*\big)^{-1}\\ &= \big(P_n(\lambda) - AP_{n-1}(\lambda)\big)^{-1}\big(Q_n(\lambda) - AQ_{n-1}(\lambda)\big).\end{aligned}$$

In particular, if $\lambda \in \mathbb{R}$ is not a zero of $\det\big(P_n(\lambda) - AP_{n-1}(\lambda)\big)$ then the matrix $\big(P_n(\lambda) - AP_{n-1}(\lambda)\big)^{-1}\big(Q_n(\lambda) - AQ_{n-1}(\lambda)\big)$ is hermitian.

(5) If μ_0 is an extremal matrix of measures of the set V_{n-1}, then μ_0 has at most nN^2 points in its support.

PROOF. (1), (2) and (3) can be found in Lemma 2.1 of [D2].

(4) can be proved as the Step 1 of [D2, Theorem 3.1].

(5) This result is taken from [K]. Taking into account this article is written in Ukrainian, we include here the translation of the proof, to make this paper more complete. We seize the opportunity to express Professor Alexander Aptekarev our gratitude for his assistance to translate this result.

Suppose on the contrary that μ_0 has a number of points in its support bigger than or equal to nN^2+1. Then it is possible to choose nN^2+1 disjoint intervals $\Delta_0, \ldots, \Delta_{nN^2}$ such that $\mu_0(\Delta_i) \neq \theta$, for $0 \leq i \leq nN^2$. Let's call

$$\Delta_{nN^2+1} = \mathbb{R} \setminus \bigcup_{i=1}^{nN^2} \Delta_i.$$

https://doi.org/10.4153/CJM-1997-034-5 Published online by Cambridge University Press

Since $\Delta_0 \subseteq \Delta_{nN^2+1}$ we have that $\mu_0(\Delta_{nN^2+1}) \neq \theta$.

Given a vector $(a_1, \ldots, a_{nN^2+1})$ in $\mathbb{R}^{nN^2+1}$ with $a_i \geq 0$ for $1 \leq i \leq nN^2+1$, it is possible to define a matrix of measures ν by

$$\nu(A) = \sum_{i=1}^{nN^2+1} a_i \mu_0(A \cap \Delta_i).$$

Let's call W_{n-1} the set of vectors $(a_1, \ldots, a_{nN^2+1})$ in $\mathbb{R}^{nN^2+1}$ with non-negative coordinates for which the corresponding matrix of measures ν belongs to V_{n-1}.

We then have that the matrix of measures associated to the vector $(1, \ldots, 1)$ is μ_0 and since this matrix of measures is extremal in V_{n-1} we have that $(1, \ldots, 1)$ is extremal in W_{n-1}.

The set W_{n-1} is characterized by the equations

$$\begin{cases} a_i \geq 0, & \text{for } 1 \leq i \leq nN^2+1 \\ \sum_{i=1}^{nN^2+1} a_i \int_{\Delta_i} t^k \, d\mu_0(t) = S_k, & \text{for } 0 \leq k \leq n-1, \end{cases} \tag{2.4}$$

which determine a system of nN^2 equations in the nN^2+1 unknowns a_i. Thus it is possible to choose a solution $(h_1, \ldots, h_{nN^2+1})$ for the homogeneous system (that is, replacing S_k by θ) such that $|h_i| < 1$ for $1 \leq i \leq nN^2+1$.

Defining $H_+ = (1+h_1, \ldots, 1+h_{nN^2+1})$ and $H_- = (1-h_1, \ldots, 1-h_{nN^2+1})$, it is clear that H_+ and H_- satisfy (2.4), and so does $H = \frac{1}{2}(H_+ + H_-)$, so H is not extremal in W_{n-1}, which is a contradiction. ∎

We are now ready to proceed with the proof of Theorem 1:

PROOF OF THEOREM 1. (1) $\Rightarrow$ (2) If μ is extremal in V_{n-1}, from Lemma 2.1(5) follows that μ is discrete and then of the form

$$\mu = \sum_{i=1}^{q} G_i \delta_{x_i}$$

where q is certain natural number, x_i are real numbers and G_i are positive semidefinite numerical matrices.

If the polynomials up to degree $n-1$ are not dense in the space $L^2(\mu)$ then by the Hahn-Banach Theorem there exists a non zero operator Λ acting on $L^2(\mu)$ such that it vanishes on any polynomial of degree lower than n. By the duality Theorem we can represent this operator with a unique function $g = \sum_{i=1}^{q} A_i \delta_{x_i}$ in $L^2(\mu)$. For any function f in $L^2(\mu)$ the operator Λ is defined by

$$\Lambda(f) = \int_{\mathbb{R}} \tau\big(f(t)\, d\mu(t) g^*(t)\big) = \sum_{i=1}^{q} \tau\big(f(x_i) G_i A_i^*\big).$$

Since $\Lambda(p) = 0$ for any polynomial p up to degree $n-1$ we have that the not necessarily positive definite matrix of measures

$$\mu_0 = \sum_{i=1}^{q} G_i A_i^* \delta_{x_i}$$

https://doi.org/10.4153/CJM-1997-034-5 Published online by Cambridge University Press

has null moments up to degree $n-1$.

By considering

$$\mu_0^H = \sum_{i=1}^q (G_i A_i^* + A_i G_i)\delta_{x_i} = \sum_{i=1}^q H_i \delta_{x_i}$$

we obtain a hermitian matrix of measures with null moments up to degree $n-1$. Let's put a_i for the smallest non zero eigenvalue of the matrix G_i (we recall that G_i are positive semidefinite) and choose a positive number C such that

$$\frac{1}{C}\max_{1\le i\le q}\|H_i\|_2 < \min_{1\le i\le q} a_i.$$

We decompose the matrix of measures μ in the following way

$$\mu = \frac{1}{2}\Big(\mu + \frac{1}{C}\mu_0^H\Big) + \frac{1}{2}\Big(\mu - \frac{1}{C}\mu_0^H\Big),$$

and we next prove that the matrices of measures $\mu \pm \frac{1}{C}\mu_0^H$ are positive definite, for which it is enough to prove that the numerical matrices $G_i \pm \frac{1}{C}H_i$ are positive semidefinite. For it, if v is a vector in $\mathbb{C}^N$ and belongs to $\operatorname{Ker} G_i$, then

$$v\Big(G_i \pm \frac{1}{C}H_i\Big)v^* = v\Big(G_i \pm \frac{1}{C}(G_i A_i^* + A_i G_i)\Big)v^* = 0$$

and if v is orthogonal to $\operatorname{Ker} G_i$ then

$$\begin{aligned} v\Big(G_i \pm \frac{1}{C}H_i\Big)v^* &= vG_iv^* \pm \frac{1}{C}vH_iv^* \\ &\ge \|v\|^2 a_i - \frac{1}{C}\|H_i\|\,\|v\|^2 \\ &\ge \Big(\min_{1\le i\le q} a_i - \frac{1}{C}\max_{1\le i\le q}\|H_i\|\Big)\|v\|^2 > 0. \end{aligned}$$

Hence we conclude that the matrix of measures μ can not be extremal in V_{n-1}, which is a contradiction.

(2) $\Rightarrow$ (1) Suppose now that the polynomials up to degree $n-1$ are dense in $L^2(\mu)$ and also that μ is not extremal in the set V_{n-1}, that is μ can be written as $\mu = \alpha\mu_1 + (1-\alpha)\mu_2$, for certain $0 < \alpha < 1$ and μ_1 and μ_2 being two different matrices of measures in V_{n-1}. We then define the operators T and T_1 on $L^2(\mu)$ by:

$$T(f) = \int_{\mathbb{R}} \tau(f\,d\mu I) \quad \text{and} \quad T_1(f) = \int_{\mathbb{R}} \tau(f\,d\mu_1 I).$$

Both are clearly linear and T is continuous because I belongs to $L^2(\mu)$. For T_1 we have:

$$\begin{aligned} \Big|\int_{\mathbb{R}} \tau(f\,d\mu_1 I)\Big| &= \frac{1}{\alpha}\Big|\int_{\mathbb{R}} \tau\Big(f\,d\big(\mu - (1-\alpha)\mu_2\big)I\Big)\Big| \\ &\le \frac{1}{\alpha}\Big(\int_{\mathbb{R}} \tau\Big(f\,d\big(\mu - (1-\alpha)\mu_2\big)f^*\Big)\Big)^{\frac12}\Big(\int_{\mathbb{R}} \tau\Big(I\,d\big(\mu - (1-\alpha)\mu_2\big)I\Big)\Big)^{\frac12} \\ &\le \frac{1}{\alpha}\Big(\int_{\mathbb{R}} \tau(f\,d\mu f^*)\Big)^{\frac12}\Big(\int_{\mathbb{R}} d\tau\mu\Big)^{\frac12}, \end{aligned}$$

https://doi.org/10.4153/CJM-1997-034-5 Published online by Cambridge University Press

so if we take $c = \frac{1}{\alpha}(\int_{\mathbb{R}} d\tau\mu)^{\frac{1}{2}}$ we have $|T_1(f)| \leq c\|f\|_{2,\mu}$, that is T_1 is a continuous operator. We define now $U = T - T_1$, U is a linear and continuous operator defined on $L^2(\mu)$, U is not the zero operator but vanishes on any polynomial up to degree $n-1$, so we conclude that the polynomials up to degree $n-1$ are not dense in the space $L^2(\mu)$. ∎

REMARK. Observe that the proof given for (2) ⇒ (1) also works for a non-truncated matrix moment problem. In that case, if μ is a solution of the moment problem and the polynomials are dense in the space $L^2(\mu)$, then μ is extremal in the set of solutions of that matrix moment problem.

To prove (3) ⇒ (2) we need the following lemma:

LEMMA 2.2. *(1) If $\mu = M(t)\,d\tau\mu$ is a positive definite matrix of measures and $(P_k)_{k=0}^{\infty}$ is a sequence of associated orthonormal matrix polynomials, with $\operatorname{dgr}(P_k) = k$ and with non-singular leading coefficient, then for any f in $L^2(\mu)$ and for any n natural, the best approximation of f in $L^2(\mu)$ by a polynomial of degree at most n is the given by the Fourier series*

$$\sum_{k=0}^{n}(f, P_k)P_k$$

where

$$(f, P_k) = \int_{\mathbb{R}} f(t)M(t)P_k^*(t)\,d\tau\mu(t).$$

(2) For a given positive definite matrix of measures of the form

$$\mu = \sum_{i=1}^{q} G_i\delta_{x_i}$$

where q is a natural number, x_i are real numbers and G_i are positive semidefinite numerical matrices, the following statements are equivalent:

(a) $\overline{\mathbb{C}_{n-1}^{N\times N}[x]} = L^2(\mu)$.

(b) $\sum_{i=1}^{q}\operatorname{rank}(G_i) = nN$.

PROOF. (1) It is enough to proceed as in the scalar case.

(2) For a matrix of measures like this (discrete and with finite support), it is clear that any function f in $L^2(\mu)$ can be represented as

$$f = \sum_{i=1}^{q} F_i\delta_{x_i}$$

where F_i are numerical matrices and x_i are the points in the support of μ. We identify f with θ if and only if $F_iG_i = \theta$, for $i = 1, \ldots, q$. Since $\dim\big(\operatorname{Im}(F \longrightarrow G_iF)\big) = N\operatorname{rank}(G_i)$ it is not difficult to find

$$\dim\big(L^2(\mu)\big) = \sum_{i=1}^{q} N\operatorname{rank}(G_i) = N\sum_{i=1}^{q}\operatorname{rank}(G_i). \tag{2.5}$$

(a) ⇒ (b) If $\mathbb{C}_{n-1}^{N\times N}[x]$ is dense in $L^2(\mu)$, we can represent any function f in $L^2(\mu)$ in the form

$$f = \sum_{k=0}^{n-1} A_kP_k(t)$$

A_k being $N \times N$ numerical matrices; if

$$\sum_{k=0}^{n-1} A_k P_k(t) = \sum_{k=0}^{n-1} A'_k P_k(t),$$

it is enough to use the orthonormality of $P_0, \ldots, P_{n-1}$ to obtain that $A_k = A'_k$, for $k = 0, \ldots, n-1$. This way, ranging the matrices A_k in $M_{N\times N}(\mathbb{C})$ we obtain the whole space $L^2(\mu)$ and hence we deduce

$$\dim\bigl(L^2(\mu)\bigr) = nN^2. \tag{2.6}$$

Taking now into account (2.5) we have

$$\sum_{i=1}^{q} \operatorname{rank}(G_i) = nN.$$

(b) $\Rightarrow$ (a) Given P, Q in $\mathbb{C}_{n-1}^{N\times N}[x]$, $P \neq Q$, we can write

$$P = \sum_{k=0}^{n-1} A_k P_k, \quad Q = \sum_{k=0}^{n-1} B_k P_k$$

with $B_k \neq A_k$ for some k between 0 and $n-1$, hence $P \neq Q$ in $L^2(\mu)$. Since $\mathbb{C}_{n-1}^{N\times N}[x] \subseteq L^2(\mu)$, we have $\dim\bigl(L^2(\mu)\bigr) \geq nN^2$.

Furthermore, from (2.5) we obtain

$$\dim(L^2(\mu)) = N\Bigl(\sum_{i=1}^{q} \operatorname{rank}(G_i)\Bigr) = nN^2,$$

so $\mathbb{C}_{n-1}^{N\times N}[x] = L^2(\mu)$, and hence the polynomials up to degree $n-1$ are dense in $L^2(\mu)$. ∎

We return to the proof of the Theorem, proving (3) $\Rightarrow$ (2)

In [D2, Section 3] it is proved that for a sequence of orthonormal matrix polynomials $(P_n)_n$ satisfying the three-term recurrence relation

$$tP_n(t) = A_{n+1}P_{n+1}(t) + B_n P_n(t) + A_n^* P_{n-1}(t),$$

if A is a $N \times N$ matrix such that $A_n A = A^* A_n^*$, then the simple fraction decomposition

$$\bigl(P_n(\lambda) - AP_{n-1}(\lambda)\bigr)^{-1} R(\lambda) = \sum_{k=1}^{m} \frac{G_{n,k}}{\lambda - x_{n,k}} \tag{2.7}$$

is always possible, for R any matrix polynomial of degree lower than or equal to $n-1$. This is possible even though the zeros $x_{n,k}$ of $\det\bigl(P_n(\lambda) - AP_{n-1}(\lambda)\bigr) = 0$ can have multiplicity bigger than 1. $G_{n,k}$ are certain numerical matrices explicitly given by:

$$G_{n,k} = \frac{1}{\Bigl(\det\bigl(P_n(t) - AP_{n-1}(t)\bigr)\Bigr)^{(l_k)}(x_{n,k})} \cdot \Bigl(\operatorname{Adj}\bigl(P_n(t) - AP_{n-1}(t)\bigr)\Bigr)^{(l_k-1)}(x_{n,k}) R(x_{n,k}), \quad k = 1, \ldots, m,$$

https://doi.org/10.4153/CJM-1997-034-5 Published online by Cambridge University Press

where l_k is the order of $x_{n,k}$ as a zero of the polynomial $\det\big(P_n(\lambda)-AP_{n-1}(\lambda)\big)$. For the particular case $R(\lambda)=Q_n(\lambda)-AQ_{n-1}(\lambda)$, the matrices $G_{n,k}$ ($k=1,\dots,m$) are positive semidefinite and the rank of $G_{n,k}$ coincides with the multiplicity of the zero $x_{n,k}$.

In consequence, $\sum_{i=1}^{q}\operatorname{rank}(G_i)=\sum_{i=1}^{q}$ multiplicity $(x_i)=nN$. Now it is enough to apply Lemma 2.2(2).

(2) $\Rightarrow$ (3) Suppose now $\mu\in V_{2n-2}$, and $\mathbb{C}_{n-1}^{N\times N}[x]$ is dense in $L^2(\mu)$. Since this condition is equivalent to the extremality of μ in V_{n-1}, from Lemma 2.1(5), we can express $\mu=\sum_{i=1}^{q}G_i\delta_{x_i}$, with G_i positive semidefinite numerical matrices and x_i real numbers we are going to determine. We proceed in several steps.

STEP ONE. For $1\le i,j\le q$, the following formula holds

$$G_j\big(P_{n-1}^*(x_j)A_nP_n(x_i)-P_n^*(x_j)A_n^*P_{n-1}(x_i)\big)G_i=\theta.$$

From Lemma 2.2(2), and since by hypothesis $\overline{\mathbb{C}_n^{N\times N}[x]}=L^2(\mu)$, for any function f in $L^2(\mu)$ we have

$$f=\sum_{k=0}^{n-1}(f,P_k)P_k$$

in $L^2(\mu)$, where (f,P_k) is the Fourier coefficient of f associated to P_k. Taking into account that we identify f with θ if and only if $F_iG_i=\theta$, for $i=1,\dots,q$, this is equivalent to

$$f(x_i)G_i=\sum_{k=0}^{n-1}(f,P_k)P_k(x_i)G_i,\qquad \text{for } 1\le i\le q.$$

For $\lambda\in\mathbb{C}\setminus\mathbb{R}$, we call $f_\lambda(t)=\frac{I}{t-\lambda}$. Since f_λ is bounded, $f_\lambda\in L^2(\mu)$. We compute its Fourier coefficients:

$$\begin{aligned}(f_\lambda,P_k)&=\int_{\mathbb{R}}\frac{I}{t-\lambda}\,d\mu(t)P_k^*(t)\\&=\int_{\mathbb{R}}d\mu(t)\frac{P_k^*(t)-P_k^*(\lambda)}{(t-\lambda)}+\int_{\mathbb{R}}\frac{d\mu(t)}{t-\lambda}P_k^*(\lambda)\\&=Q_k^*(\lambda)+\omega(\lambda)P_k^*(\lambda),\end{aligned}$$

where $\omega(\lambda)$ is the Hilbert transform of f_λ, that is

$$\omega(\lambda)=\int_{\mathbb{R}}\frac{d\mu(t)}{t-\lambda}=\sum_{i=1}^{q}\frac{G_i}{x_i-\lambda}.$$

Then, for $1\le i\le q$ and $\lambda\in\mathbb{C}\setminus\mathbb{R}$, we have

$$\begin{aligned}\frac{G_i}{x_i-\lambda}&=\sum_{k=0}^{n-1}\big(Q_k^*(\lambda)+\omega(\lambda)P_k^*(\lambda)\big)P_k(x_i)G_i\\&=\sum_{k=0}^{n-1}Q_k^*(\lambda)P_k(x_i)G_i+\omega(\lambda)\sum_{k=0}^{n-1}P_k^*(\lambda)P_k(x_i)G_i,\end{aligned}$$

https://doi.org/10.4153/CJM-1997-034-5 Published online by Cambridge University Press

which gives

$$\theta = \Big[-I + (x_i - \lambda)\sum_{k=0}^{n-1} Q_k^*(\lambda)P_k(x_i)\Big]G_i + \omega(\lambda)\Big[(x_i - \lambda)\sum_{k=0}^{n-1} P_k^*(\lambda)P_k(x_i)\Big]G_i$$

for $1 \le i \le q$ and $\lambda \in \mathbb{C} \setminus \mathbb{R}$.

By using the formulae (2.1) and (2.3) we obtain

$$\begin{aligned}\theta &= [Q_{n-1}^*(\lambda)A_nP_n(x_i) - Q_n^*(\lambda)A_n^*P_{n-1}(x_i)]G_i \\ &\quad + \omega(\lambda)[P_{n-1}^*(\lambda)A_nP_n(x_i) - P_n^*(\lambda)A_n^*P_{n-1}(x_i)]G_i\end{aligned} \tag{2.8}$$

for $1 \le i \le q$ and $\lambda \in \mathbb{C} \setminus \mathbb{R}$.

Multiplying (2.8) by $(x_j - \lambda)$ and taking limit for λ tending to x_j, for $j \ne i$, and taking into account that $\omega(\lambda) = \sum_{i=1}^q \frac{G_i}{x_i - \lambda}$ we obtain

$$G_j\big(P_{n-1}^*(x_j)A_nP_n(x_i) - P_n^*(x_j)A_n^*P_{n-1}(x_i)\big)G_i = \theta \tag{2.9}$$

for $1 \le i, j \le q$, $i \ne j$.

Since $P_{n-1}^*(x_i)A_nP_n(x_i) - P_n^*(x_i)A_n^*P_{n-1}(x_i) = \theta$ (formula (2.1)), formula (2.9) holds for every $1 \le i, j \le q$.

STEP TWO. Definition of the matrix A.

To define the matrix A we need to prove that the sum of the images of the mappings represented by the matrices $P_{n-1}(x_i)G_i$, $i = 1, \dots, q$, is all $\mathbb{C}^N$. To prove this it will be enough to establish that the sum of the images of the mappings represented by the matrices $P_{n-1}(x_i)G_iP_{n-1}^*(x_i)$, $i = 1, \dots, q$, is all $\mathbb{C}^N$.

Let $A_i = (v_{i,j})_{j=1,\dots,k_i}$ be a basis of $\mathrm{Im}\big(P_{n-1}(x_i)G_iP_{n-1}^*(x_i)\big)$. If the space spanned by the elements of A_i $(1 \le i \le q)$, that is, $\langle A_1, \dots, A_q\rangle$ has dimension lower than N, then there exists a non zero vector v orthogonal to $\langle A_1, \dots, A_q\rangle$. For this vector v and for any i we can write

$$P_{n-1}(x_i)G_iP_{n-1}^*(x_i)v = \sum_{j=1}^{k_i} \alpha_{i,j}v_{i,j},$$

where $\alpha_{i,j}$ are complex numbers. Taking now into account that v is orthogonal to A_i we have

$$v^*P_{n-1}(x_i)G_iP_{n-1}^*(x_i)v = \sum_{j=1}^{k_i} \alpha_{i,j}v^*v_{i,j} = 0.$$

Since

$$\sum_{i=1}^q P_{n-1}(x_i)G_iP_{n-1}^*(x_i) = \int_{\mathbb{R}} P_{n-1}(t)\,d\mu(t)P_{n-1}^*(t) = I, \tag{2.10}$$

summing in the former equality for i we get $v^*v = 0$, which implies $v = 0$, which is a contradiction.

We are now ready to define the matrix A.

A is the matrix representing a linear mapping $A\colon \mathbb{C}^N \longrightarrow \mathbb{C}^N$ we define as follows: for a vector v in $\mathbb{C}^N$, we can write $v = \sum_{i=1}^q P_{n-1}(x_i)G_i v_i$, where v_i are certain vectors in $\mathbb{C}^N$. Then we put $Av = \sum_{i=1}^q P_n(x_i)G_i v_i$.

Let's see now that A is well defined. Suppose that v admits the two representations

$$v = \sum_{i=1}^q P_{n-1}(x_i)G_i w_i = \sum_{i=1}^q P_{n-1}(x_i)G_i z_i,$$

where w_i and z_i are vectors in $\mathbb{C}^N$. We have to prove that

$$\sum_{i=1}^q P_n(x_i)G_i w_i = \sum_{i=1}^q P_n(x_i)G_i z_i.$$

By multiplying in (2.9) to the right by w_i and summing for i we get

$$G_j P_{n-1}^*(x_j)A_n \sum_{i=1}^q P_n(x_i)G_i w_i - G_j P_n^*(x_j)A_n^* v = \theta, \quad \text{for } 1 \le j \le q. \tag{2.11}$$

By multiplying to the right by z_i and summing for i we get

$$G_j P_{n-1}^*(x_j)A_n \sum_{i=1}^q P_n(x_i)G_i z_i - G_j P_n^*(x_j)A_n^* v = \theta, \quad \text{for } 1 \le j \le q. \tag{2.12}$$

Subtracting (2.11) and (2.12) yields

$$G_j P_{n-1}^*(x_j)A_n \Big(\sum_{i=1}^q P_n(x_i)G_i w_i - \sum_{i=1}^q P_n(x_i)G_i z_i\Big) = \theta, \quad \text{for } 1 \le j \le q. \tag{2.13}$$

Multiplying in (2.13) by $P_{n-1}(x_j)$ to the left and summing we get

$$\Big(\sum_{j=1}^q P_{n-1}(x_j)G_j P_{n-1}^*(x_j)\Big)A_n\Big(\sum_{i=1}^q P_n(x_i)G_i w_i - \sum_{i=1}^q P_n(x_i)G_i z_i\Big) = \theta.$$

Again taking into account (2.10) and that A_n is non-singular we deduce $\sum_{i=1}^q P_n(x_i)G_i w_i = \sum_{i=1}^q P_n(x_i)G_i z_i$ and thus A is well defined.

STEP THREE. $A_n A = A^* A_n^*$.

Given an arbitrary vector v in $\mathbb{C}^N$, we write it as $v = \sum_{i=1}^q P_{n-1}(x_i)G_i v_i$, where v_i are vectors in $\mathbb{C}^N$. Multiplying in (2.9) to the right by v_i and summing for i yields

$$G_j\Big(P_{n-1}^*(x_j)A_n A - P_n^*(x_j)A_n^*\Big)v = \theta,$$

hence we deduce

$$G_j P_{n-1}^*(x_j)A_n A - G_j P_n^*(x_j)A_n^* = \theta.$$

Multiplying now in this formula to the left by v_j^* and summing yields

$$v^* A_n A - v^* A^* A_n^* = \theta$$

for any vector v in $\mathbb{C}^N$, so we have $A_n A = A^* A_n^*$.

https://doi.org/10.4153/CJM-1997-034-5 Published online by Cambridge University Press

STEP FOUR. For $\lambda \in \mathbb{C} \setminus \mathbb{R}$,

$$\omega(\lambda) = -\big(P_n(\lambda) - AP_{n-1}(\lambda)\big)^{-1}\big(Q_n(\lambda) - AQ_{n-1}(\lambda)\big).$$

We retake now formula (2.8):

$$\begin{aligned}\theta = {} & [Q^*_{n-1}(\lambda)A_nP_n(x_i) - Q^*_n(\lambda)A^*_nP_{n-1}(x_i)]G_i \\ & + \omega(\lambda)[P^*_{n-1}(\lambda)A_nP_n(x_i) - P^*_n(\lambda)A^*_nP_{n-1}(x_i)]G_i,\end{aligned}$$

for $1 \le i \le q$. Let $v \in \mathbb{C}^N$ and write it as

$$v = \sum_{i=1}^{q} P_{n-1}(x_i)G_i\omega_i.$$

Then $Av = \sum_{i=1}^{q} P_n(x_i)G_i\omega_i$. Multiplication by ω_i in the above formula (2.8) for each i followed by summation over i yields

$$\theta = [Q^*_{n-1}(\lambda)A_nA - Q^*_n(\lambda)A^*_n]v + \omega(\lambda)[P^*_{n-1}(\lambda)A_nA - P^*_n(\lambda)A^*_n]v$$

for every v in $\mathbb{C}^N$ and every $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Since A_n is non-singular we deduce

$$\theta = [Q^*_{n-1}(\lambda)A_nAA_n^{*-1} - Q^*_n(\lambda)]A^*_nv + \omega(\lambda)[P^*_{n-1}(\lambda)A_nAA_n^{*-1} - P^*_n(\lambda)]A^*_nv$$

for every vector v in $\mathbb{C}^N$ and $\lambda \in \mathbb{C} \setminus \mathbb{R}$. By using now $A_nA = A^*A^*_n$ we obtain

$$\theta = [Q^*_{n-1}(\lambda)A^* - Q^*_n(\lambda)] + \omega(\lambda)[P^*_{n-1}(\lambda)A^* - P^*_n(\lambda)], \quad \lambda \in \mathbb{C} \setminus \mathbb{R}$$

which gives

$$\omega(\lambda) = -\big(Q^*_n(\lambda) - Q^*_{n-1}(\lambda)A^*\big)\big(P^*_n(\lambda) - P^*_{n-1}(\lambda)A^*\big)^{-1}, \quad \lambda \in \mathbb{C} \setminus \mathbb{R}$$

which by virtue of Lemma 2.1 is

$$\omega(\lambda) = -\big(P_n(\lambda) - AP_{n-1}(\lambda)\big)^{-1}\big(Q_n(\lambda) - AQ_{n-1}(\lambda)\big), \quad \lambda \in \mathbb{C} \setminus \mathbb{R}.$$

Taking into account $\omega(\lambda) = \sum_{i=1}^{q} \frac{G_i}{x_i-\lambda}$ and the simple fraction decomposition of $\big(P_n(\lambda) - AP_{n-1}(\lambda)\big)^{-1}\big(Q_n(\lambda) - AQ_{n-1}(\lambda)\big)$, the theorem is proved. ∎

REMARK. Observe that in case the polynomials of degree less than or equal to $n-1$ are dense in $L^2(\mu)$, no closure operation is necessary, because in this case the space $L^2(\mu)$ is of finite dimension.

3. **The maximum mass Theorem.** It is convenient to compare Theorem 1 with its scalar version. The representation of the measures in V_{2n-2} extremal in V_{n-1} in terms of the polynomials $p_n - ap_{n-1}$ is almost trivial in the scalar case, because it is possible to find directly $a = \frac{p_n(x_i)}{p_{n-1}(x_i)}$. As we have seen, the matrix structure introduces some complications which make the result harder to obtain, requiring new ideas.

These differences of structure between the matrix and scalar case create important divergences as we are going to see next. Indeed, as we pointed out in the introduction, in the scalar case, the points of the support of the extremal measures support the maximum mass in V_{2n-2}. This property is not true in the matrix case. The matrices in V_{2n-2} extremal in V_{n-1} may not support the maximum mass in V_{2n-2}.

THEOREM 3.1. *If $\mu \in V_{2n-2}$ is a matrix of measures extremal in V_{n-1}, put*

$$\mu = \sum_{i=1}^{q} G_i \delta_{x_i}$$

where x_i are the zeros of $P_n(\lambda) - AP_{n-1}(\lambda)$, for certain A such that $A_n A = A^* A_n^*$, then the following conditions are equivalent:

(1) μ reaches in x_{i_0} the maximum mass which can be concentrated in V_{2n-2}, more concretely,

$$G_{i_0} = \Big(\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0})\Big)^{-1}.$$

(2) G_{i_0} is non-singular.

(3) $P_n(\lambda) - AP_{n-1}(\lambda)$ has a zero of maximum multiplicity (N) in x_{i_0} .

(4) $P_{n-1}(x_{i_0})$ is non-singular and $A = P_n(x_{i_0}) P_{n-1}^{-1}(x_{i_0})$.

PROOF. (1) $\Rightarrow$ (2) The maximum mass which can be concentrated for $\mu \in V_{2n-2}$ in x_{i_0} is

$$\Big(\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0})\Big)^{-1},$$

because

$$\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0}) = \int_{\mathbb{R}} \Big(\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(t)\Big)\, d\mu(t) \Big(\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(t)\Big)^*$$

and hence

$$\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0}) \geq \Big(\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0})\Big) B \Big(\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0})\Big)^*$$

where we denote by B the mass of the matrix of measures in x_{i_0}. We then have

$$\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0}) \Big[\Big(\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0})\Big)^{-1} - B\Big] \sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0}) \geq 0$$

from which we deduce that

$$\Big(\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0})\Big)^{-1} - B \geq 0.$$

This result is also contained in [Z].

So, if μ supports the maximum mass which can be concentrated in x_{i_0}, this is given by

$$G_{i_0} = \Big(\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0})\Big)^{-1},$$

which is a non-singular matrix since it is the inverse of an invertible matrix:

$$\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0}) \geq P_0^*(x_{i_0}) P_0(x_{i_0}) = I.$$

(2) $\Leftrightarrow$ (3) Since rank G_{i_0}, as mentioned earlier, is equal to the multiplicity of the zero x_{i_0} of $P_n(\lambda) - AP_{n-1}(\lambda)$, we see that G_{i_0} is non-singular if and only if x_{i_0} is of multiplicity N.

(3) $\Leftrightarrow$ (4) If $P_n(\lambda) - AP_{n-1}(\lambda)$ has a zero of multiplicity N in x_{i_0}, we deduce from the Remark 2.3 of [DL1] that $P_n(x_{i_0}) - AP_{n-1}(x_{i_0}) = \theta$, and hence that $P_n(x_{i_0}) = AP_{n-1}(x_{i_0})$. If $P_{n-1}(x_{i_0})$ was singular we would have that x_{i_0} would be a zero of $P_{n-1}(\lambda)$ and $P_n(\lambda)$, and moreover $P_n(x_{i_0})$ and $P_{n-1}(x_{i_0})$ would have a common eigenvector associated to 0, which is in contradiction with Theorem 1.1(2) of [DL1]. In consequence, $P_{n-1}(x_{i_0})$ is non-singular and thus $A = P_n(x_{i_0})P_{n-1}^{-1}(x_{i_0})$. On the other hand, if $A = P_n(x_{i_0})P_{n-1}^{-1}(x_{i_0})$, it is clear that $P_n(\lambda) - AP_{n-1}(\lambda)$ has a zero of multiplicity N in x_{i_0}.

(2) $\Rightarrow$ (1) Taking limit as $\lambda \longrightarrow x_i$ in (2.8) we get

$$\begin{aligned}\theta = {}&\left(Q_{n-1}^*(x_i)A_nP_n(x_i) - Q_n^*(x_i)A_n^*P_{n-1}(x_i)\right)G_i\\ &+ \sum_{j\neq i} G_j \frac{1}{x_j - x_i}\left(P_{n-1}^*(x_i)A_nP_n(x_i) - P_n^*(x_i)A_n^*P_{n-1}(x_i)\right)G_i\\ &+ \lim_{\lambda\longrightarrow x_i} G_i \frac{1}{x_i - \lambda}\left(P_{n-1}^*(\lambda)A_nP_n(x_i) - P_n^*(\lambda)A_n^*P_{n-1}(x_i)\right)G_i.\end{aligned}$$

From (2.3) in Lemma 2.1,

$$\begin{aligned}Q_{n-1}^*(x_i)A_nP_n(x_i) - Q_n^*(x_i)A_n^*P_{n-1}(x_i) &= \left(P_n^*(x_i)A_n^*Q_{n-1}(x_i) - P_{n-1}^*(x_i)A_nQ_n(x_i)\right)^*\\ &= -I^* = -I\end{aligned}$$

and from (2.1) in Lemma 2.1,

$$P_{n-1}^*(x_i)A_nP_n(x_i) - P_n^*(x_i)A_n^*P_{n-1}(x_i) = \theta$$

so that

$$-G_i = G_i\left(\lim_{\lambda\longrightarrow x_i} \frac{P_n^*(\lambda)A_n^*P_{n-1}(x_i) - P_{n-1}^*(\lambda)A_nP_n(x_i)}{x_i - \lambda}\right)G_i.$$

Therefore

$$G_i^{-1} = -\lim_{\lambda\longrightarrow x_i} \frac{P_n^*(\lambda)A_n^*P_{n-1}(x_i) - P_{n-1}^*(\lambda)A_nP_n(x_i)}{x_i - \lambda},$$

since G_i was invertible by assumption.

Finally this gives

$$G_i^{-1} = -\lim_{\lambda\longrightarrow x_i} \frac{\lambda - x_i}{x_i - \lambda} \sum_{k=0}^{n-1} P_k^*(\lambda)P_k(x_i) = \sum_{k=0}^{n-1} P_k^*(x_i)P_k(x_i). \qquad \blacksquare$$

https://doi.org/10.4153/CJM-1997-034-5 Published online by Cambridge University Press

REFERENCES

[A] N. I. Akhiezer, *The classical moment problem and some related questions in analysis,* English Translation, Oliver and Boyd, Edinburgh, 1965.

[AN] A. I. Aptekarev and E. M. Nikishin, *The scattering problem for a discrete Sturm-Liouville operator,* Math. USSR-Sb. **49**(1984), 325–355.

[D1] A. J. Durán, *On orthogonal polynomials with respect to a positive definite matrix of measures,* Canad. J. Math. **47**(1995), 88–112.

[D2] ______, *Markov's Theorem for orthogonal matrix polynomials,* Canad. J. Math. **48**(1996), 1180–1195.

[DL1] A. J. Durán and P. López-Rodríguez, *Orthogonal matrix polynomials: zeros and Blumenthal's Theorem,* J. Approx. Theory **84**(1996), 96–118.

[DL2] ______, *The L^p space of a positive definite matrix of measures and density of matrix polynomials in L^1*, J. Approx. Theory, to appear.

[K] A. N. Kochubei, *The extreme points of the set of solutions of the truncated matrix moment problem,* Dopovīdī Akad. Nauk Ukraïn. RSR Ser. A. (1972), 418–421.

[R] M. Rosenberg, *The square-integrability of matrix-valued functions with respect to a non-negative hermitian measure,* Duke Math. J. **31**(1964), 291–298.

[Z] D. Zhani, *Problème des moments matriciels sur la droite: construction d'une famille de solutions et questions d'unicité,* Publ. Dept. Math., Université Claude-Bernard-Lyon I Nouvelle serie **21**(1984), 1–84.

Departamento de Análisis Matemático
Universidad de Sevilla
Apdo. 1160
41080 Sevilla
Spain
e-mail: duran@cica.es

Departamento de Análisis Matemático
Universidad de Sevilla
Apdo. 1160
41080 Sevilla
Spain
e-mail: plopez@cica.es