RIESZ'S THEOREM FOR ORTHOGONAL MATRIX POLYNOMIALS

PEDRO LOPEZ-RODRIGUEZ

Universidad de Sevilla

ABSTRACT. We describe the image through the Stieltjes transform of the set of solutions V of a matrix moment problem. We extend Riesz's theorem to the matrix setting, proving that those matrices of measures of V for which the matrix polynomials are dense in the corresponding $\mathcal{L}^2$ space are precisely those whose Stieltjes transform is an extremal point (in the sense of convexity) of the image set.

1. Introduction.

For a positive Borel measure ν on $\mathbb{R}$ with finite moments of any order $s_n = \int_{\mathbb{R}} t^n d\nu(t)$ we denote by V the set of positive Borel measures μ on $\mathbb{R}$ satisfying $\int_{\mathbb{R}} t^n d\mu(t) = s_n$, $n \geq 0$, that is, the set of solutions to the Hamburger moment problem defined by ν. By V_n we denote the set of positive Borel measures on $\mathbb{R}$ such that $\int_{\mathbb{R}} t^k d\mu(t) = s_k$, $0 \leq k \leq n$, that is, the set of solutions to the truncated moment problem defined by ν.

We say that the measure ν is determinate if there is no other positive measure having the same moments as those of ν, that is, if $V = \{\nu\}$, otherwise we say that ν is indeterminate. This alternative is related to the index of deficiency of the operator defined on ℓ^2 by the infinite Jacobi matrix

$$J = \begin{pmatrix} b_0 & a_1 & & & \\ a_1 & b_1 & a_2 & & \\ & a_2 & b_2 & a_3 & \\ & & \ddots & \ddots & \ddots \end{pmatrix},$$

where the coefficients a_i ($\neq 0$) and b_i are the coefficients which appear in the three term recurrence relation satisfied by the orthogonal polynomials $(p_n)_n$ associated to ν,

$$t p_n(t) = a_{n+1} p_{n+1}(t) + b_n p_n(t) + a_n p_{n-1}(t), \quad n \geq 0.$$

The index of deficiency of J is 0 if the moment problem is determinate and 1 if the moment problem is indeterminate.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

In 1922 Nevanlinna proved that for a fixed non real λ the image through the Stieltjes transform of all the measures of V in the point λ

$$I(V)(\lambda) = \left\{ \int_{\mathbb{R}} \frac{d\mu(t)}{t - \lambda} : \mu \in V \right\}$$

is either a point if the moment problem is determinate or a circle if the moment problem is indeterminate, and the occurrence of these two cases does not depend on the non real λ chosen (see [**N**] or [**A**]). The measures μ for which $I(\mu)(\lambda)$ lies in the circumference of this circle $I(V)(\lambda)$ are called N-extremal (Nevanlinna-extremal).

In 1923 M. Riesz proved that in order that the set $\mathcal{P}$ of polynomials in $L^2(\mu)$ be dense, it is necessary and sufficient that the measure μ be N-extremal at every non real point λ, and for this it is sufficient that it should be N-extremal in at least one such point (see [**Ri**] or [**A**]).

The purpose of this paper is to generalize these two results for a completely indeterminate matrix moment problem.

Given $\nu = (\nu_{i,j})_{1 \leq i,j \leq N}$ a positive definite matrix of measures (for any Borel set A the numerical matrix $\nu(A)$ is positive semidefinite) with finite matrix moments

$$S_k = \int_{\mathbb{R}} t^k d\nu(t)$$

of any order $k \geq 0$, we denote by V the set of positive definite matrices of measures having the same matrix moments as those of ν, and by V_n the set of positive matrices of measures whose moments up to degree n are the same as those of ν.

We say that the positive definite matrix of measures ν is determinate if no other positive definite matrix of measures has the same moments as those of ν, that is, the positive definite matrix of measures ν is uniquely determined by the moments $\int_{\mathbb{R}} t^n d\nu(t)$, $n \geq 0$.

By $(P_n)_{n=0}^{\infty}$ we denote the sequence of orthonormal matrix polynomials with respect to ν, P_n of degree n and with non-singular leading coefficient.

These polynomials $(P_n)_n$ satisfy a three term recurrence relation of the form

$$(1.1) \qquad t P_n(t) = A_{n+1} P_{n+1}(t) + B_n P_n(t) + A_n^* P_{n-1}(t), \quad n \geq 0,$$

(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 $Q_1(t) = A_1^{-1}$.

In the matrix case the determinacy or indeterminacy of the matrix moment problem is also related to the index of deficiency of the operator J defined by the infinite N-Jacobi matrix

$$J = \begin{pmatrix} B_0 & A_1 & & & \\ A_1^* & B_1 & A_2 & & \\ & A_2^* & B_2 & A_3 & \\ & & \ddots & \ddots & \ddots \end{pmatrix}$$

on the space ℓ^2, where A_n and B_n are the coefficients which appear in the three term recurrence relation (1.1). In this case the index of deficiency can be any natural number from 0 to N, 0 in the determinate case and N in the completely indeterminate case. In the latter case the two series

$$\sum_{k=0}^{\infty} Q_k^*(\lambda) P_k(\eta) \quad \text{and} \quad \sum_{k=0}^{\infty} P_k^*(\lambda) P_k(\eta)$$

converge uniformly in the variables λ and η on every bounded set of the complex plane (see [**K**]).

In [**B**] it is proved that the rank of the limit matrix $R(\lambda) = \lim_{n \to \infty} R_n(\lambda)$ exists and is the same for every non real λ, where

$$(1.2) \qquad R_n(\lambda) = \left(\sum_{k=0}^{n} P_k^*(\overline{\lambda}) P_k(\lambda) \right)^{-1}.$$

This result is also mentioned by Krein in [**K**], who refers to [**Na**] for a proof. In this paper we assume the rank of this matrix is N and consequently the matrix $\sum_{k=0}^{\infty} P_k^*(\overline{\lambda}) P_k(\lambda)$ is invertible, for every non real λ, and equal to $R(\lambda)^{-1}$.

As in the scalar case, for a fixed non real λ, we denote by $I(V)(\lambda)$ the image through the Stieltjes transform of all the matrices of measures of V in the point λ

$$I(V)(\lambda) = \left\{ \int_{\mathbb{R}} \frac{d\mu(t)}{t - \lambda} : \mu \in V \right\}.$$

Firstly, for a fixed non real λ, we describe the set $I(V)(\lambda)$. This is the set of $N \times N$ complex matrices ω satisfying the matrix inequality

$$(1.3) \qquad [\omega + C(\lambda)] R(\overline{\lambda})^{-1} [\omega + C(\lambda)]^* \leq |\lambda - \overline{\lambda}|^{-2} R(\lambda),$$

where $C(\lambda) = \mathcal{B}(\lambda, \overline{\lambda}) \mathcal{D}(\lambda, \overline{\lambda})^{-1}$ (see the Preliminaries for the definitions of $\mathcal{B}(\lambda, \overline{\lambda})$ and $\mathcal{D}(\lambda, \overline{\lambda})$. $A \leq B$ means that $B - A$ is positive semidefinite.

The extremal points (in the sense of convexity) of the set $I(V)(\lambda)$ are the matrices ω for which equality is attained in (1.3). If μ is a matrix of measures in V for which $I(\mu)(\lambda)$ is

a extremal point of $I(V)(\lambda)$, we call this matrix of measures N-extremal, as in the scalar case.

Finally, we generalize Riesz's theorem to the matrix setting by proving that the matrices of measures of V for which the set $\mathcal{P}$ of matrix polynomials is dense in the corresponding space $\mathcal{L}^2(\mu)$ are precisely the N-extremal matrices of measures, and that the N-extremality of a matrix of measures does not depend on the non real λ chosen.

For the case $N = 1$ one recovers the very well known classical formulas exposed in [**A**, Ch. 1].

2. Preliminaries.

In what follows, if $P(\lambda)$ is a matrix polynomial, we denote by $P^*(\lambda)$ the polynomial obtained from $P(\lambda)$ by replacing each of its matrix coefficients by its hermitian conjugate, so that $P(\lambda)^* = P^*(\overline{\lambda})$. For a matrix polynomial of two variables $P(\lambda, \eta)$ the definition is the same, so that we have $P(\lambda, \eta)^* = P^*(\overline{\lambda}, \overline{\eta})$. If $F(\lambda)$ is a holomorphic matrix function we define $F^*(\lambda) = F(\overline{\lambda})^*$.

The set of positive definite matrices of measures is endowed with the vague and weak topologies. If ν a positive definite matrix of measures, the set V of matrices of measures having the same moments as ν is a compact and convex set for these topologies which coincide on V (see [**DL2**]).

For μ a positive definite matrix of measures, the space $\mathcal{L}^2(\mu)$ is defined as the set of $N \times N$ matrix functions $f : \mathbb{R} \to M_{N \times N}(\mathbb{C})$ such that $\tau(f(t)M(t)f(t)^*) \in L^1(\tau\mu)$, where $M(t)$ is the Radon-Nikodym derivative of μ with respect to its trace $(\tau\mu)$ (for a matrix $A = (a_{i,j})_{1 \leq i,j \leq N}$, we denote τA for its trace, i. e. $\tau A = \sum_{i=1}^{N} a_{i,i}$):

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

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

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

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 $\mathcal{L}^p$ spaces associated to μ, $1 \leq p < \infty$, 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 $\mathcal{L}^2(\mu)$. In any case, the polynomials $(P_k)_{k=0,\dots,n}$ are orthonormal with respect to any measure in V_{2n}.

We include here the matrix version of some classical formulas for orthonormal scalar polynomials. The proofs are easily verified using the three term recurrence relation (1.1).

(2.1)
$$\mathcal{A}_n(u,v) = (v-u) \sum_{k=0}^{n-1} Q_k^*(u)Q_k(v) = Q_{n-1}^*(u)A_n Q_n(v) - Q_n^*(u)A_n^* Q_{n-1}(v), \quad \text{for } u,v \in \mathbb{C},$$

(2.2)
$$\mathcal{B}_n(u,v) = -I + (v-u)\sum_{k=0}^{n-1} Q_k^*(u)P_k(v) = Q_{n-1}^*(u)A_nP_n(v) - Q_n^*(u)A_n^*P_{n-1}(v), \quad \text{for } u,v \in \mathbb{C},$$

(this is Green's formula),
(2.3)
$$\mathcal{C}_n(u,v) = I + (v-u)\sum_{k=0}^{n-1} P_k^*(u)Q_k(v) = P_{n-1}^*(u)A_nQ_n(v) - P_n^*(u)A_n^*Q_{n-1}(v), \quad \text{for } u,v \in \mathbb{C},$$

(2.4)
$$\mathcal{D}_n(u,v) = (v-u)\sum_{k=0}^{n-1} P_k^*(u)P_k(v) = P_{n-1}^*(u)A_nP_n(v) - P_n^*(u)A_n^*P_{n-1}(v), \quad \text{for } u,v \in \mathbb{C},$$

(this a Christoffel-Darboux formula). We will also use the Liouville-Ostrogradsky formula

(2.5)
$$Q_n(\lambda)P_{n-1}^*(\lambda) - P_n(\lambda)Q_{n-1}^*(\lambda) = A_n^{-1}, \quad \text{for } \lambda \in \mathbb{C},$$

and the relations

(2.6)
$$P_n(\lambda)Q_n^*(\lambda) = Q_n(\lambda)P_n^*(\lambda), \text{ for } \lambda \in \mathbb{C},$$

(2.7)
$$\mathcal{A}_n(u,v)\mathcal{D}_n^*(u,v) - \mathcal{B}_n(u,v)\mathcal{C}_n^*(u,v) = I, \quad \text{for } u,v \in \mathbb{C},$$

and finally

(2.8)
$$\mathcal{C}_n(u,v)\mathcal{D}_n^*(u,v) = \mathcal{D}_n(u,v)\mathcal{C}_n^*(u,v), \quad \text{for } u,v \in \mathbb{C}.$$

We will also use that

(2.9)
$$\mathcal{C}_n(\lambda,\overline{\lambda}) = -\mathcal{B}_n(\lambda,\overline{\lambda})^*, \quad \text{for } \lambda \in \mathbb{C},$$

and that

(2.10) $\quad \mathcal{D}_n(\lambda,\overline{\lambda}) = (\overline{\lambda}-\lambda)R_{n-1}(\overline{\lambda})^{-1}$ and $\mathcal{D}_n^*(\lambda,\overline{\lambda}) = (\overline{\lambda}-\lambda)R_{n-1}(\lambda)^{-1}, \quad \text{for } \lambda \in \mathbb{C}.$

By $\mathcal{A}(u,v)$, $\mathcal{B}(u,v)$, $\mathcal{C}(u,v)$ and $\mathcal{D}(u,v)$ we denote the limit matrix functions defined from $\mathcal{A}_n(u,v)$, $\mathcal{B}_n(u,v)$, $\mathcal{C}_n(u,v)$ and $\mathcal{D}_n(u,v)$ when n tends to infinity.

3. The main theorems.

For any non real λ, we define the set $B_n(\lambda)$ to be the set of $N \times N$ complex matrices ω such that

$$(3.1) \qquad [\omega + C_n(\lambda)]R_{n-1}(\overline{\lambda})^{-1}[\omega + C_n(\lambda)]^* \leq |\lambda - \overline{\lambda}|^{-2}R_{n-1}(\lambda),$$

where $C_n(\lambda) = \mathcal{B}_n(\lambda, \overline{\lambda})\mathcal{D}_n(\lambda, \overline{\lambda})^{-1}$.

The calculations in Lemma 1 (see below) show that $B_n(\lambda)$ is also the set of $N \times N$ complex matrices ω satisfying the matrix inequality

$$(3.2) \qquad \sum_{k=0}^{n-1}(Q_k^*(\lambda) + \omega P_k^*(\lambda))(Q_k(\overline{\lambda}) + P_k(\overline{\lambda})\omega^*) \leq \frac{\omega - \omega^*}{\lambda - \overline{\lambda}}.$$

We put $B_\infty(\lambda)$ for the intersection of all the sets $B_n(\lambda)$. $B_\infty(\lambda)$ is clearly the set of $N \times N$ complex matrices ω such that

$$(3.3) \qquad [\omega + C(\lambda)]R(\overline{\lambda})^{-1}[\omega + C(\lambda)]^* \leq |\lambda - \overline{\lambda}|^{-2}R(\lambda),$$

where $C(\lambda) = \mathcal{B}(\lambda, \overline{\lambda})\mathcal{D}(\lambda, \overline{\lambda})^{-1}$.

Similarly, $B_\infty(\lambda)$ is also the set of $N \times N$ complex matrices ω such that

$$(3.4) \qquad \sum_{k=0}^{\infty}(Q_k^*(\lambda) + \omega P_k^*(\lambda))(Q_k(\overline{\lambda}) + P_k(\overline{\lambda})\omega^*) \leq \frac{\omega - \omega^*}{\lambda - \overline{\lambda}}.$$

Looking at (3.1) and (3.3) it is immediate that upon a linear matrix transformation, any of the sets $B_n(\lambda)$ or $B_\infty(\lambda)$ is in a one to one correspondence with the set of $N \times N$ complex matrices T satisfying $TT^* \leq I$, which is a convex set whose extremal points are the matrices verifying $TT^* = I$, that is, the unitary matrices (this is a well-known result in operator theory which can be proved for example with the aid of the singular value decomposition of matrices). This implies that these sets $B_n(\lambda)$ and $B_\infty(\lambda)$ are convex sets whose extremal points ($\text{Ext}B_n(\lambda)$ and $\text{Ext}B_\infty(\lambda)$) are those for which equality is attained in (3.1) and (3.2) or (3.3) and (3.4) respectively.

By using formulas (2.1), (2.2), (2.3) and (2.4) in (3.2) it is straightforward to see that an equivalent condition for ω to be an extremal point of $B_n(\lambda)$ is that the matrix

$$(3.5) \qquad (\omega P_n^*(\lambda) + Q_n^*(\lambda))A_n^*(P_{n-1}(\overline{\lambda})\omega^* + Q_{n-1}(\overline{\lambda}))$$

is hermitian.

It is clear that for all $n \geq 1$ we have $B_\infty(\lambda) \subseteq B_{n+1}(\lambda) \subseteq B_n(\lambda)$. It is also clear that ω belongs to the set of interior points of $B_n(\lambda)$ or $B_\infty(\lambda)$ ($\text{Int}B_n(\lambda)$ and $\text{Int}B_\infty(\lambda)$) if a strict inequality is attained in (3.1) and (3.2) or (3.3) and (3.4) respectively.

We have the following results:

Theorem 1. *Let V denote the set of solutions to a completely indeterminate matrix moment problem defined by a natrix of measures ν and let $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Then we have*

$$B_\infty(\lambda) = I(V)(\lambda).$$

The key to prove this theorem will be the inclusions

(3.6) $$\text{Int}B_n(\lambda) \subseteq I(V_{2n-2})(\lambda) \subseteq B_n(\lambda).$$

The proofs of these inclusions present more difficulties than in the scalar case. We will prove then later in Lemmas 1 to 7. Indeed, in the scalar case the set $I(V_{2n-2})(\lambda)$ is given by

$$I(V_{2n-2})(\lambda) = B_n(\lambda) \setminus \left\{ -\frac{q_{n-1}(\lambda)}{p_{n-1}(\lambda)} \right\}.$$

The point $-q_{n-1}(\lambda)/p_{n-1}(\lambda)$ lies on the border of the circle $B_n(\lambda)$. When a moves along the real axis, the quotient

$$-\frac{q_n(\lambda) - aq_{n-1}(\lambda)}{p_n(\lambda) - ap_{n-1}(\lambda)}$$

describes all the points of the circumference of the closed disk $B_n(\lambda)$ except for the limit point $-q_{n-1}(\lambda)/p_{n-1}(\lambda)$. The well known quadrature formula (see [**A**, p. 20]) gives that every point defined by the former quotient for $a \in \mathbb{R}$ belongs to $I(V_{2n-2})(\lambda)$. It is easy to see that $-q_{n-1}(\lambda)/p_{n-1}(\lambda) \notin I(V_{2n-2})(\lambda)$, but this is of no importance because taking into account that $I(V_{2n-2})(\lambda)$ is a convex set and the simple geometry of the circles $B_n(\lambda)$ it is immediate to deduce that $\text{Int}B_n(\lambda) \subseteq I(V_{2n-2})(\lambda)$. This inclusion is not at all so immediate in the matrix case. We prove it in Lemmas 2 to 7 by means of new ideas.

Proof of Theorem 1

Suppose first that $\omega \in B_\infty(\lambda)$. $B_\infty(\lambda)$ is contained in $B_n(\lambda)$ for all n and thus we can put $\omega = \omega_n$, for all n, being ω_n in $B_n(\lambda)$. Since the interior set $\text{Int}B_n(\lambda)$ is dense in $B_n(\lambda)$, we can find η_n in $\text{Int}B_n(\lambda)$ such that $\lim_{n\to\infty} \|\omega_n - \eta_n\| = 0$. Since $\text{Int}B_n(\lambda) \subseteq I(V_{2n-2})(\lambda)$, there exists a matrix of measures σ_n in V_{2n-2} such that $\eta_n = I(\sigma_n)(\lambda)$, for all n. Since the set $\{\mu \geq \theta : \tau\mu(\mathbb{R}) \leq c\}$ is vaguely compact, where c is a positive constant (see Lemma 3.8 in [**DL2**]) and $\sigma_n(\mathbb{R}) = S_0$ for $n \geq 0$, there exists σ a vague accumulation point of a subsequence (σ_{n_p}) of (σ_n). Like in the proof of Lemma 3.10 of [**DL2**] we have $\sigma \in V$. We have $\sigma_n(\mathbb{R}) = \sigma(\mathbb{R})$ for $n \geq 1$, so by virtue of Theorem 3.1 of [**DL2**], (σ_n) converges weakly to σ. In particular

$$I(\sigma)(\lambda) = \lim_{p\to\infty} I(\sigma_{n_p})(\lambda) = \lim_{p\to\infty} \eta_{n_p} = \lim_{p\to\infty} \omega_{n_p} = \omega$$

and we have proved that $B_\infty(\lambda) \subseteq I(V)(\lambda)$.

Since $I(V_{2n-2})(\lambda) \subseteq B_n(\lambda)$, for every n, the reverse inclusion is clear.

∎

Theorem 2. *(Riesz's theorem for orthogonal matrix polynomials) Let μ be a positive definite matrix of measures corresponding to a completely indeterminate matrix moment problem. Then the following conditions are equivalent:*

(1) *There exists $\lambda_0 \in \mathbb{C} \setminus \mathbb{R}$ such that $I(\mu)(\lambda_0)$ is an extremal point (in the sense of convexity) of the set $B_\infty(\lambda_0)$.*

(2) *For any $\lambda \in \mathbb{C} \setminus \mathbb{R}$, $I(\mu)(\lambda)$ is an extremal point (in the sense of convexity) of the set $B_\infty(\lambda)$*

(3) *$\mathcal{P}$ is dense in $\mathcal{L}^2(\mu)$, equivalently $(P_n(t))_{n=0}^\infty$ is an orthonormal basis for the Hilbert space $\mathcal{L}^2(\mu)$.*

Proof of Theorem 2

$(3) \Rightarrow (2)$ The polynomials are dense in $\mathcal{L}^2(\mu)$ if for any function f in the space $\mathcal{L}^2(\mu)$ we have equality in Bessel's inequality, which is equivalent to

$$(3.7) \qquad \sum_{k=0}^{\infty}(f, P_k)(f, P_k)^* = \int_{\mathbb{R}} f(t)M(t)f^*(t)d\tau\mu(t).$$

In particular, for $f_\lambda(t) = \dfrac{I}{t - \lambda} \in \mathcal{L}^2(\mu)$ we have

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

being $I(\mu)(\lambda) = \displaystyle\int_{\mathbb{R}} \frac{d\mu(t)}{t - \lambda}$ the Stieltjes transform of μ in the point λ, and

$$\int_{\mathbb{R}} f(t)d\mu(t)f^*(t) = \int_{\mathbb{R}} \frac{d\mu(t)}{|t - \lambda|^2} = \frac{I(\mu)(\lambda) - I(\mu)(\lambda)^*}{\lambda - \overline{\lambda}} = \frac{\operatorname{Im}I(\mu)(\lambda)}{\operatorname{Im}\lambda},$$

so equality in (3.7) is

$$\sum_{k=0}^{\infty}(Q_k^*(\lambda) + \omega(\lambda)P_k^*(\lambda))(Q_k(\overline{\lambda}) + P_k(\overline{\lambda})\omega^*(\overline{\lambda})) = \frac{\operatorname{Im}I(\mu)(\lambda)}{\operatorname{Im}\lambda},$$

that is $I(\mu)(\lambda) \in \operatorname{Ext}B_\infty(\lambda)$.

$(2) \Rightarrow (1)$ is obvious.

$(1) \Rightarrow (3)$ We suppose (1) holds and we claim $f_{\lambda_0}^n = \dfrac{I}{(t - \lambda_0)^n} \in \overline{\mathcal{P}}$, for $n \geq 1$. The assertion for $n = 1$ is the assumption. We now prove that $f_{\lambda_0}^{n+1} \in \overline{\mathcal{P}}$ under the assumption $f_{\lambda_0}^n \in \overline{\mathcal{P}}$, so that the claim is established by induction. For given $\epsilon > 0$, there exists a matrix polynomial $P \in \mathcal{P}$ such that $\|f_{\lambda_0}^n - P\|_2 \leq \epsilon|\operatorname{Im}\lambda_0|$. Dividing P by $(x - \lambda_0)I$ we get $P(x) = (x - \lambda_0)Q(x) + A$, with Q another polynomial of degree $n - 1$ and A a $N \times N$ complex matrix. We have

$$\|f_{\lambda_0}^{n+1} - Af_{\lambda_0} - Q\|_2^2 =$$

$$= \tau\int_{\mathbb{R}}\left(\frac{I}{(t - \lambda_0)^{n+1}} - \frac{A}{t - \lambda_0} - Q(t)\right)M(t)\left(\frac{I}{(t - \lambda_0)^{n+1}} - \frac{A}{t - \lambda_0} - Q(t)\right)^* d\tau\mu(t)$$

$$= \tau\int_{\mathbb{R}}\frac{1}{|t - \lambda_0|^2}\left(\frac{I}{(t - \lambda_0)^n} - A - Q(t)(t - \lambda_0)\right)M(t)\left(\frac{I}{(t - \lambda_0)^n} - A - Q(t)(t - \lambda_0)\right)^* d\tau\mu(t)$$

$$\leq \frac{1}{|\mathrm{Im}\lambda_0|^2} \| \frac{I}{(t-\lambda_0)^n} - P(t) \|_2^2 \leq \epsilon^2,$$

and since $\dfrac{A}{t-\lambda_0}$ belongs to the closure of $\mathcal{P}$ we deduce that $\dfrac{I}{(t-\lambda_0)^{n+1}}$ also does.

Now, if $f \in \mathcal{L}^2(\mu)$ is orthogonal to $\overline{\mathcal{P}}$ and we consider the Stieltjes transform

$$I(f\mu)(z) = \int_{\mathbb{R}} \frac{f(t)}{t-z} d\mu(t), \quad z \in \mathbb{C} \setminus \mathbb{R}$$

using that $f_{\lambda_0}^n, f_{\overline{\lambda_0}}^n \in \overline{\mathcal{P}}$ for $n \geq 1$, we see that

$$I(f\mu)^{(n)}(z) = \theta, \text{ for } z = \lambda_0, \overline{\lambda_0}, \text{ and } n \geq 0,$$

but then we have an analytic function $I(f\mu)$ in a domain D such that $I(f\mu)^{(n)}(z_0) = \theta$ for $n \geq 0$ and a certain $z_0 \in D$. So $I(f\mu)$ is equal to θ in D. We conclude that $I(f\mu)$ is identically zero in each of the two half planes $\mathbb{C} \setminus \mathbb{R}$, and then $f = \theta$, μ a.e., hence the polynomials are dense in $\mathcal{L}^2(\mu)$.

∎

We also have the following Theorem:

Theorem 3. *If $\mu \in V_{2n-2}$ is such that $I(\mu)(\lambda) \in ExtB_n(\lambda)$, then $\overline{\mathcal{P}_{n-1}} = \mathcal{L}^2(\mu)$*

Proof of Theorem 3

The hypothesis means that equality is attained in (3.2), that is, the function $f_\lambda(t) = \dfrac{I}{t-\lambda}$ can be approximated by matrix polynomials up to degree $n-1$. Now the proof finishes exactly in the same way as the proof of Theorem 2.

∎

4. Proofs of the inclusions.

In this last section we study in detail the set $B_n(\lambda)$ and other related sets, with the purpose of proving the inclusions (3.6). We remark that these inclusions are valid without supposing the matrix moment problem to be completely indeterminate. We first prove the second inclusion of (3.6)

Lemma 1.
$$I(V_{2n-2})(\lambda) \subseteq B_n(\lambda)$$

Proof

Let's suppose μ is a measure in V_{2n-2}. We know that the first n orthonormal matrix polynomials $P_0, \dots, P_{n-1}$ form an orthonormal system in the space $\mathcal{L}^2(\mu)$. After the above calculations, from Bessel's inequality for the function $f_\lambda(t) = \dfrac{I}{t-\lambda}$, we deduce that

$$\sum_{k=0}^{n-1} (Q_k^*(\lambda) + I(\mu)(\lambda)P_k^*(\lambda))(Q_k(\overline{\lambda}) + P_k(\overline{\lambda})I(\mu)(\lambda)^*) \leq \frac{\mathrm{Im}I(\mu)(\lambda)}{\mathrm{Im}\lambda}.$$

Multiplying and using formulas (2.1), (2.2), (2.3), (2.4) and (2.9) we get

$$(4.1) \quad I(\mu)(\lambda)\left(\frac{i\mathcal{D}_n(\lambda,\overline{\lambda})}{2\mathrm{Im}\lambda}\right)I(\mu)(\lambda)^* + I(\mu)(\lambda)\left(\frac{i\mathcal{B}_n(\lambda,\overline{\lambda})}{2\mathrm{Im}\lambda}\right)^*$$
$$+ \left(\frac{i\mathcal{B}_n(\lambda,\overline{\lambda})}{2\mathrm{Im}\lambda}\right)I(\mu)(\lambda)^* + \left(\frac{i\mathcal{A}_n(\lambda,\overline{\lambda})}{2\mathrm{Im}\lambda}\right) \leq \theta.$$

Taking into account that

$$\frac{i\mathcal{D}_n(\lambda,\overline{\lambda})}{2\mathrm{Im}\lambda} = R_{n-1}(\overline{\lambda})^{-1},$$

(4.1) can be written as

$$(4.2) \quad [I(\mu)(\lambda) + C_n(\lambda)]R_{n-1}(\overline{\lambda})^{-1}[I(\mu)(\lambda) + C_n(\lambda)]^*$$
$$+ \left(\frac{i\mathcal{A}_n(\lambda,\overline{\lambda})}{2\mathrm{Im}\lambda}\right) - \left(\frac{i\mathcal{B}_n(\lambda,\overline{\lambda})}{2\mathrm{Im}\lambda}\right)\left(\frac{i\mathcal{D}_n(\lambda,\overline{\lambda})}{2\mathrm{Im}\lambda}\right)^{-1}\left(\frac{i\mathcal{B}_n(\lambda,\overline{\lambda})}{2\mathrm{Im}\lambda}\right)^* \leq \theta,$$

where we recall that $C_n(\lambda) = \mathcal{B}_n(\lambda,\overline{\lambda})\mathcal{D}_n(\lambda,\overline{\lambda})^{-1}$.

Now, using (2.7), (2.8), (2.9) and (2.10) we get

$$\left(\frac{i\mathcal{A}_n(\lambda,\overline{\lambda})}{2\mathrm{Im}\lambda}\right) - \left(\frac{i\mathcal{B}_n(\lambda,\overline{\lambda})}{2\mathrm{Im}\lambda}\right)\left(\frac{i\mathcal{D}_n(\lambda,\overline{\lambda})}{2\mathrm{Im}\lambda}\right)^{-1}\left(\frac{i\mathcal{B}_n(\lambda,\overline{\lambda})}{2\mathrm{Im}\lambda}\right)^* = |\lambda - \overline{\lambda}|^{-2}R_{n-1}(\lambda).$$

Thus, (4.2) becomes

$$(4.3) \quad [I(\mu)(\lambda) + C_n(\lambda)]R_{n-1}(\overline{\lambda})^{-1}[I(\mu)(\lambda) + C_n(\lambda)]^* \leq |\lambda - \overline{\lambda}|^{-2}R_{n-1}(\lambda).$$

and we have proved that

$$I(V_{2n-2})(\lambda) \subseteq B_n(\lambda).$$

$\blacksquare$

As we indicated above, the proof of this Lemma shows also that $B_n(\lambda)$ is described through (3.2). By a simple limit process we also have that $B_\infty(\lambda)$ is described through (3.4)

To prove the inclusion $\mathrm{Int}B_n(\lambda) \subseteq I(V_{2n-2})(\lambda)$ we introduce the set of points $\Gamma_n(\lambda)$ defined by

$$\Gamma_n(\lambda) = \{-(P_n(\lambda) - AP_{n-1}(\lambda))^{-1}(Q_n(\lambda) - AQ_{n-1}(\lambda)) : A_n A = A^* A_n^*\}.$$

We will use that
(4.4)
$$(P_n(\lambda) - AP_{n-1}(\lambda))^{-1}(Q_n(\lambda) - AQ_{n-1}(\lambda)) = (Q_n^*(\lambda) - Q_{n-1}^*(\lambda)A^*)(P_n^*(\lambda) - P_{n-1}^*(\lambda)A^*)^{-1}$$

for any complex λ. This formula can be proved like step 1 of Theorem 3.1 of [**D2**].

The three following Lemmas show the relation between $\Gamma_n(\lambda)$, $I(V_{2n-2})(\lambda)$ and $B_n(\lambda)$.

Lemma 2.

$$co(\Gamma_n(\lambda)) \subseteq I(V_{2n-2})(\lambda),$$

where $co(\Gamma_n(\lambda))$ stands for the convex hull of $\Gamma_n(\lambda)$.

Proof

By virtue of Theorem 3.2 in [**D2**], for any $N \times N$ complex matrix A for which $A_n A = A^* A_n^*$, it is possible to make the simple fraction decomposition

$$(P_n(\lambda) - AP_{n-1}(\lambda))^{-1}(Q_n(\lambda) - AQ_{n-1}(\lambda)) = \sum_{i=1}^{m} \frac{G_i}{\lambda - x_i},$$

with the numerical matrices G_i being positive semidefinite, and if we define the matrix of measures

$$\mu_A = \sum_{i=1}^{m} G_i \delta_{x_i},$$

then μ_A is a solution of the truncated matrix moment problem up to degree $2n - 2$, and

$$I(\mu)(\lambda) = \int_{\mathbb{R}} \frac{d\mu(t)}{t - \lambda} = -(P_n(\lambda) - AP_{n-1}(\lambda))^{-1}(Q_n(\lambda) - AQ_{n-1}(\lambda)).$$

We have thus proved that $\Gamma_n(\lambda) \subseteq I(V_{2n-2})(\lambda)$. Since $I(V_{2n-2})(\lambda)$ is a convex set, we have

$$co(\Gamma_n(\lambda)) \subseteq I(V_{2n-2})(\lambda).$$

∎

Lemma 3.

$$\Gamma_n(\lambda) \subseteq Ext B_n(\lambda)$$

Proof

Observe that for $I(\mu_A)(\lambda)$ in $\Gamma_n(\lambda)$ we have

$$I(\mu_A)(\lambda)(P_n^*(\lambda) - P_{n-1}^*(\lambda)A^*) = -(Q_n^*(\lambda) - Q_{n-1}^*(\lambda)A^*).$$

Multiplying by A_n and taking into account that $A_n A$ is hermitian we deduce that

$$(I(\mu_A)(\lambda)P_n^*(\lambda) + Q_n^*(\lambda))A_n^*(P_{n-1}(\overline{\lambda})I(\mu_A)(\lambda)^* + Q_{n-1}(\overline{\lambda}))$$

is a hermitian matrix. This is precisely formula (3.5) for $I(\mu_A)(\lambda)$, which means that $I(\mu_A)(\lambda)$ is an extremal point of $B_n(\lambda)$, that is we have proved that $\Gamma_n(\lambda) \subseteq Ext B_n(\lambda)$.

∎

Lemma 4.
$$I(V_{2n-2})(\lambda) \cap ExtB_n(\lambda) = \Gamma_n(\lambda).$$

Proof

In theorem 1 of [**DL3**] the authors characterize the matrices of measures μ of V_{2n-2} for which the polynomials up to degree $n-1$ are dense in $\mathcal{L}^2(\mu)$: μ is an extremal matrix of measures (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 $(P_n(\lambda) - AP_{n-1}(\lambda))$ (the zeros of a matrix polynomial $P(t)$ are the zeros of the determinant of $P(t)$), where the matrix A makes $A_n A$ hermitian:

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

(1) μ is an extremal matrix of measures of the set V_{n-1}.
(2) The matrix polynomials of degree less than or equal to $n-1$ are dense in the space $\mathcal{L}^2(\mu)$.
(3) There exists an $N \times N$ matrix A such that $A_n A = A^* A_n^*$ and for which $\mu = \sum_{i=1}^{q} G_i \delta_{x_i}$, where x_i, $i = 1, \ldots, q$, are the different zeros of the polynomial $\det(P_n(\lambda) - AP_{n-1}(\lambda))$ and G_i are the matrices which appear in the simple fraction decomposition

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

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

As a consequence of this and Theorem 3 the statement holds.

∎

The following Lemma shows which points of $ExtB_n(\lambda)$ do not lie in $\Gamma_n(\lambda)$.

Lemma 5. *If $\omega \in ExtB_n(\lambda)$, the following two conditions are equivalent:*

(1) $\omega \in \Gamma_n(\lambda)$
(2) $det(Q_{n-1}^*(\lambda) + \omega P_{n-1}^*(\lambda)) \ne 0$.

Proof

(1)⇒(2) $\omega \in \Gamma_n(\lambda)$ means that there exists a $N \times N$ matrix A with $A_n A$ hermitian such that

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

which gives

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

If $\det(Q_{n-1}^*(\lambda) + \omega P_{n-1}^*(\lambda)) = 0$, there exists a non zero vector v in $\mathbb{C}^N$ verifying $\omega P_{n-1}^*(\lambda)v = -Q_{n-1}^*(\lambda)v$. Multiplying (4.5) by $P_{n-1}^*(\lambda)v$ to the right and using that $\omega P_{n-1}^*(\lambda)v = -Q_{n-1}^*(\lambda)v$ we find

$$(Q_n(\lambda)P_{n-1}^*(\lambda) - P_n(\lambda)Q_{n-1}^*(\lambda))v = A(-P_{n-1}(\lambda)Q_{n-1}^*(\lambda) + Q_{n-1}(\lambda)P_{n-1}^*(\lambda))v.$$

By using formulas (2.5) and (2.6) we finally get $A_n^{-1} v = \theta$, which is absurd because A_n is invertible and $v \neq \theta$.

$(1) \Rightarrow (2)$ If $\omega \in \mathrm{Ext} B_n(\lambda)$, then the matrix

$$(\omega P_n^*(\lambda) + Q_n^*(\lambda)) A_n^* (P_{n-1}(\overline{\lambda}) \omega^* + Q_{n-1}(\overline{\lambda}))$$

is hermitian (3.5). Since $(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))$ is invertible, this is equivalent to

$$(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))^{-1} (\omega P_n^*(\lambda) + Q_n^*(\lambda)) A_n^*$$

being hermitian. We put this matrix equal to $A_n A = A^* A_n^*$ and find

$$A^* = (\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))^{-1} (\omega P_n^*(\lambda) + Q_n^*(\lambda)).$$

With this matrix A we have

$$\omega = -(P_n(\lambda) - A P_{n-1}(\lambda))^{-1} (Q_n(\lambda) - A Q_{n-1}(\lambda))$$

and thus $\omega \in \Gamma_n(\lambda)$.

■

In the following two Lemmas we show the topological relation between $\Gamma_n(\lambda)$ and $B_n(\lambda)$.

Lemma 6. *The set $\Gamma_n(\lambda)$ is dense in $\mathrm{Ext} B_n(\lambda)$.*

Proof

For $\omega \in \mathrm{Ext} B_n(\lambda)$ we are going to define $A(p)$ for $p \geq 1$ such that $A_n A(p)$ is hermitian and

$$\omega_{A(p)}(\lambda) = -(P_n(\lambda) - A(p) P_{n-1}(\lambda))^{-1} (Q_n(\lambda) - A(p) Q_{n-1}(\lambda))$$

verifies $\lim_{p \to \infty} \omega_{A(p)}(\lambda) = \omega$.

We suppose ω is an extremal point of $B_n(\lambda)$ such that $\det(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda)) = 0$ (otherwise $\omega \in \Gamma_n(\lambda)$ as we have proved in Lemma 5). We consider the kernel and image sets of the linear operator defined on $\mathbb{C}^N$ by the $N \times N$ matrix $(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))$:

$$\mathrm{Im}(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda)) = \{v(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda)), \quad v \in \mathbb{C}^N\},$$

$$\mathrm{Ker}(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda)) = \{v \in \mathbb{C}^N : \quad v(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda)) = \theta\}.$$

We choose an orthonormal basis $\{u_1, \ldots, u_m\}$ of $\mathrm{Im}(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))$ and vectors $\{v_1, \ldots, v_m\}$ in $\mathbb{C}^N$ such that $u_i = v_i(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))$, for $1 \leq i \leq m$. We also choose $\{v_{m+1}, \ldots, v_N\}$ a basis of $\mathrm{Ker}(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))$ so that the set $\{v_1, \ldots, v_N\}$ is a basis of $\mathbb{C}^N$.

Since ω is an extremal point of $B_n(\lambda)$, by (3.5) the matrix

$$T = (\omega P_n^*(\lambda) + Q_n^*(\lambda)) A_n^* (P_{n-1}(\overline{\lambda}) \omega^* + Q_{n-1}(\overline{\lambda}))$$

is hermitian. We define a linear operator H on $\text{Im}(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))$ by putting

(4.7)
$$u_i H = v_i(\omega P_n^*(\lambda) + Q_n^*(\lambda))A_n^*, \quad 1 \le i \le m$$

and extending it by linearity. With this definition, for any vector u in $\text{Im}(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))$ we have that uHu^* is a real number. Indeed, we can express $u = \sum_{i=1}^m \alpha_i u_i$, being α_i complex numbers, and we have

$$uHu^* = \left(\sum_{i=1}^m \alpha_i v_i(\omega P_n^*(\lambda) + Q_n^*(\lambda))A_n^*\right) \left(\sum_{j=1}^m \overline{\alpha_j}(P_{n-1}(\overline{\lambda})\omega^* + Q_{n-1}(\overline{\lambda}))v_j^*\right)$$

$$= \sum_{i=1}^m |\alpha_i|^2 v_i T v_i^* + \sum_{i<j}(\alpha_i\overline{\alpha_j} v_i T v_j^* + \alpha_j\overline{\alpha_i} v_j T v_i^*)$$

which is a real number because T is a hermitian matrix.

We now consider the linear subspace of $\mathbb{C}^N$ orthogonal to $\text{Im}(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))$:

$$\text{Im}(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))^\perp = \{v \in \mathbb{C}^N : \quad u_i v^* = 0, \quad 1 \le i \le m\},$$

and a basis of it formed by the orthonormal vectors $\{u_{m+1}, \ldots u_N\}$. $\mathbb{C}^N$ is the direct sum of $\text{Im}(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))$ and $\text{Im}(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))^\perp$.

The matrix H_u representing the operator $H : \text{Im}(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda)) \to \mathbb{C}^N$ with respect to the basis $\{u_1, \ldots, u_m\}$ in $\text{Im}(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))$ and $\{u_1, \ldots, u_N\}$ in $\mathbb{C}^N$ is

$$H_u = \begin{pmatrix} u_1 H \\ \vdots \\ u_m H \end{pmatrix} C^*,$$

where C is the $N \times N$ matrix whose rows are the vectors $u_1^*, \ldots, u_N^*$ in this order. If we consider the matrix H_u formed by the $m \times m$ matrix $H_{u,1}$ and the $m \times (N-m)$ matrix $H_{u,2}$ so that $H_u = (H_{u,1}, H_{u,2})$, the condition $uHu^* \in R$ for any $u \in \text{Im}(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))$ means that the matrix $H_{u,1}$ is hermitian.

Now, for any real number $p \ge 1$ we extend H to H_p on $\mathbb{C}^N$ by defining H_p on $\text{Im}(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))^\perp$ so that the matrix M_p representing H_p with respect to the basis $\{u_1, \ldots, u_N\}$ on both sides is

$$M_p = \begin{pmatrix} H_{u,1} & H_{u,2} \\ H_{u,2}^* & pI_{(N-m)\times(N-m)} \end{pmatrix},$$

$(H_p = C^* M_p C)$. H_p is clearly hermitian. We now define $A(p) = A_n^{-1} H_p$. The matrix $A_n A(p) = H_p$ is hermitian and consequently

$$\omega_{A(p)}(\lambda) = -(P_n(\lambda) - A(p)P_{n-1}(\lambda))^{-1}(Q_n(\lambda) - A(p)Q_{n-1}(\lambda))$$

belongs to $\Gamma_n(\lambda)$. We finish by proving that

$$\lim_{p\to\infty} \omega_{A(p)} = \omega.$$

Firstly, since for $1 \le i \le m$ we have defined

$$v_i(\omega P_n^*(\lambda) + Q_n^*(\lambda))A_n^* = v_i(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))H$$

and H_p is an extension of H to $\mathbb{C}^N$, for $1 \le i \le m$ we have

$$v_i(\omega P_n^*(\lambda) + Q_n^*(\lambda))A_n^* = v_i(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))H_p.$$

Operating on this formula we find

$$v_i\omega = -v_i(Q_n^*(\lambda) - Q_{n-1}^*(\lambda)H_pA_n^{*\,-1})(P_n^*(\lambda) - P_{n-1}^*(\lambda)H_pA_n^{*\,-1})^{-1}, \quad 1 \le i \le m$$

and by using (4.4) we obtain

$$v_i\omega = -v_i(P_n(\lambda) - A_n^{-1}H_pP_{n-1}(\lambda))^{-1}(Q_n(\lambda) - A_n^{-1}H_pQ_{n-1}(\lambda)), \quad 1 \le i \le m,$$

that is, for $1 \le i \le m$ we have $v_i\omega = v_i\omega_{A(p)}$ and so

(4.8)
$$\lim_{p\to\infty} v_i\omega_{A(p)} = v_i\omega, \quad \text{for } 1 \le i \le m.$$

We finish by proving (4.8) for $m + 1 \le i \le N$. For $m + 1 \le i \le N$ we have $v_i(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda)) = \theta$, which gives $v_i\omega = -v_iQ_{n-1}^*(\lambda)P_{n-1}^*(\lambda)^{-1}$.

By using (2.5) and (2.6) we get

$$(P_n(\lambda) - A(p)P_{n-1}(\lambda))^{-1}(Q_n(\lambda) - A(p)Q_{n-1}(\lambda)) - Q_{n-1}^*(\lambda)P_{n-1}^*(\lambda)^{-1}$$
$$= (P_n(\lambda) - A(p)P_{n-1}(\lambda))^{-1}\cdot$$
$$\cdot \left[(Q_n(\lambda) - A(p)Q_{n-1}(\lambda))P_{n-1}^*(\lambda) - (P_n(\lambda) - A(p)P_{n-1}(\lambda))Q_{n-1}^*(\lambda)\right]P_{n-1}^*(\lambda)^{-1}$$
$$= (P_n(\lambda) - A_n^{-1}H_pP_{n-1}(\lambda))^{-1}A_n^{-1}P_{n-1}^*(\lambda)^{-1}$$
$$= P_{n-1}(\lambda)^{-1}[A_nP_n(\lambda)P_{n-1}(\lambda)^{-1} - H_p]^{-1}P_{n-1}^*(\lambda)^{-1}.$$

Thus it remains to be proved that for $m + 1 \le i \le N$ we have

(4.9)
$$\lim_{p\to\infty} v_iP_{n-1}(\lambda)^{-1}[A_nP_n(\lambda)P_{n-1}(\lambda)^{-1} - H_p]^{-1}P_{n-1}^*(\lambda)^{-1} = \theta.$$

For $m + 1 \le i \le N$ we have $v_i\omega = -v_iQ_{n-1}^*(\lambda)P_{n-1}^*(\lambda)^{-1}$, and using (4.4) we get $v_i\omega = -v_iP_{n-1}(\lambda)^{-1}Q_{n-1}(\lambda)$. Using (2.5), for $m + 1 \le i \le N$ we get

$$\begin{aligned}
v_iP_{n-1}(\lambda)^{-1} &= v_iP_{n-1}(\lambda)^{-1}A_n^{*\,-1}A_n^* \\
&= v_iP_{n-1}(\lambda)^{-1}[-Q_{n-1}(\lambda)P_n^*(\lambda) + P_{n-1}(\lambda)Q_n^*(\lambda)]A_n^* \\
&= v_i[-P_{n-1}(\lambda)^{-1}Q_{n-1}(\lambda)P_n^*(\lambda) + Q_n^*(\lambda)]A_n^* \\
&= v_i[\omega P_n^*(\lambda) + Q_n^*(\lambda)]A_n^*.
\end{aligned}$$

Since v_i belongs to $\mathrm{Ker}(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))$, and using that T is hermitian, we have that for any $m + 1 \leq i \leq N$ and for any $1 \leq j \leq m$,

$$v_i(\omega P_n^*(\lambda) + Q_n^*(\lambda))A_n^* u_j^* = v_i(\omega P_n^*(\lambda) + Q_n^*(\lambda))A_n^*(P_{n-1}(\overline{\lambda})\omega^* + Q_{n-1}(\overline{\lambda})v_j^*$$
$$= v_i(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))A_n(P_n(\overline{\lambda})\omega^* + Q_n(\overline{\lambda}))v_j^* = \theta.$$

This means that $v_i(\omega P_n^*(\lambda) + Q_n^*(\lambda))A_n^*$ belongs to $\mathrm{Im}(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))^\perp$, and consequently (4.9) holds if we prove that for any vector u in $\mathrm{Im}(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))^\perp$ we have

$$\lim_{p \to \infty} u(A_n P_n(\lambda)P_{n-1}(\lambda)^{-1} - H_p)^{-1}P_{n-1}^*(\lambda)^{-1} = \theta.$$

Since $P_{n-1}^*(\lambda)$ is an invertible matrix and $\{u_{m+1}, \ldots, u_N\}$ is a basis of $\mathrm{Im}(\omega P_{n-1}^*(\lambda) + Q_{n-1}^*(\lambda))^\perp$ it is enough to prove that for $m + 1 \leq i \leq N$ we have

$$\lim_{p \to \infty} u_i[A_n P_n(\lambda)P_{n-1}(\lambda)^{-1} - H_p]^{-1} = \theta.$$

Observe that

$$u_i[A_n P_n(\lambda)P_{n-1}(\lambda)^{-1} - H_p]^{-1}$$
$$= u_i[A_n P_n(\lambda)P_{n-1}(\lambda)^{-1} - C^* M_p C]^{-1}$$
$$= u_i C^*[C A_n P_n(\lambda)P_{n-1}(\lambda)^{-1}C^* - M_p]^{-1}C$$
$$= e_i[C A_n P_n(\lambda)P_{n-1}(\lambda)^{-1}C^* - M_p]^{-1}C,$$

where $e_i = (0, \ldots, 1, \ldots, 0)$, being the 1 in the position i.

The matrix $C A_n P_n(\lambda)P_{n-1}(\lambda)^{-1}C^* - M_p$ is of the form

$$\begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,m} & \alpha_{1,m+1} & \cdots & \alpha_{1,N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,m} & \alpha_{m,m+1} & \cdots & \alpha_{m,N} \\ \alpha_{m+1,1} & \cdots & \alpha_{m+1,m} & \alpha_{m+1,m+1} - p & \cdots & \alpha_{m+1,N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{N,1} & \cdots & \alpha_{N,m} & \alpha_{N,m+1} & \cdots & \alpha_{N,N} - p \end{pmatrix}.$$

The determinant of this matrix is a polynomial in the variable p of degree $N - m$, whereas the principal minors $A_{i,j}$ for i or j bigger than m are polynomials in the variable p of degree $N - m - 1$. For this reason, any entry $E_{i,j}$ of $[C A_n P_n(\lambda)P_{n-1}(\lambda)^{-1}C^* - M_p]^{-1}$ with i or j bigger than m tends to 0 when p tends to infinity, and consequently

$$\lim_{p \to \infty} e_i[C A_n P_n(\lambda)P_{n-1}(\lambda)^{-1}C^* - M_p]^{-1} = \theta, \text{ for } m + 1 \leq i \leq N$$

which proves the result.

■

We finish by proving that the convex hull of the set $\Gamma_n(\lambda)$ contains the set $\mathrm{Int}B_n(\lambda)$ of interior points of $B_n(\lambda)$.

Lemma 7.

$$IntB_n(\lambda) \subseteq co(\Gamma_n(\lambda))$$

Proof

Since $\Gamma_n(\lambda)$ is dense in $\mathrm{Ext}B_n(\lambda)$, that is we have $\overline{\Gamma_n(\lambda)} = \mathrm{Ext}B_n(\lambda)$, we deduce that

$$\overline{\mathrm{co}\,\Gamma_n(\lambda)} = \mathrm{co}\,\overline{\Gamma_n(\lambda)} = \mathrm{co}(\mathrm{Ext}B_n(\lambda)) = B_n(\lambda)$$

the last equality by virtue of Krein-Millman's theorem. Now, applying well-known arguments of convexity we have that

$$\overline{\mathrm{Int}(\mathrm{co}(\Gamma_n(\lambda)))} = \overline{\mathrm{co}\,\Gamma_n(\lambda)} = B_n(\lambda).$$

As we have previously mentioned, there exists an invertible linear operator L defined on the $N \times N$ complex matrices transforming $B_n(\lambda)$ bijectively onto the set $B = \{T : TT^* \leq I\}$. It is immediate that the image set $L(\Gamma_n(\lambda))$ is dense in $\mathrm{Ext}B = \{T : TT^* = I\}$, and that $\overline{\mathrm{Int}(\mathrm{co}(L(\Gamma_n(\lambda))))} = B$. To prove that $\mathrm{Int}B_n(\lambda) \subseteq \mathrm{co}(\Gamma_n(\lambda))$ it is enough to prove that $\mathrm{Int}B \subseteq \mathrm{co}(L(\Gamma_n(\lambda)))$.

For this, if there exists $x \in \mathrm{Int}B \backslash \mathrm{Int}(\mathrm{co}(L(\Gamma_n(\lambda))))$, we can separate x and $\mathrm{Int}(\mathrm{co}(L(\Gamma_n(\lambda))))$ with a linear operator Λ such that $\Lambda(x) = 1$ and $\Lambda(z) \leq 1$ for any z in $\mathrm{Int}(\mathrm{co}(L(\Gamma_n(\lambda))))$. Since $\overline{\mathrm{Int}(\mathrm{co}(L(\Gamma_n(\lambda))))} = B$ we have that $\Lambda(z) \leq 1$ for any z in B and thus $\|\Lambda\| \leq 1$, but this is in contradiction with $\Lambda(x) = 1$ because x in an interior point of B.

∎

Acknowledgements

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.

[B]. Ju. M. Berezanskii, *Expansions in eigenfunctions of selfadjoint operators*, Translations of Mathematical Monographs, American Mathematical Society, Providence, Rhode Island, 1968.

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

[D2]. A. J. Durán, *Markov's Theorem for orthogonal matrix polynomials*, Can. J. Math. **48 (6)** (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]. A. J. Durán and P. López-Rodríguez, *The L^p space of a positive definite matrix of measures and density of matrix polynomials in L^1*, J. Approx. Theory **90** (1997), 299-318.

[DL3]. A. J. Durán and P. López-Rodríguez, *Density questions for the truncated matrix moment problem*, Canadian J. of Math. **49 (4)** (1997), 708-721.

[K]. M. Krein, *Infinite J-matrices and a matrix moment problem*, Dokl.Akad. Nauk SSSR **69, nr. 2** (1949), 125-128 (translation from russian by Walter Van Assche in a personal note).

[N]. Nevanlinna, R., *Asymptotische Entwickelungen beschränkter Funktionen und das Stieltjessche Momentenproblem*, Ann. Acad. Sci. Fenn. A 18 **5** (1922).

[Na]. H. Nagel, *Über die quadrierbaren Hermiteschen Matrizen entstehenden Operatoren*, Math. Ann. **112 109** (1936), 247-285.

[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.

[Ri]. M. Riesz, *Sur le probleme des moments et le théoreme de Parseval correspondant*, Acta Litt. ac Sci. (Szeged) **1** (1922), 209-225.

PEDRO LÓPEZ RODRÍGUEZ, DEPARTAMENTO DE ANÁLISIS MATEMÁTICO, UNIVERSIDAD DE SEVILLA, APDO. 1160. 41080-SEVILLA, SPAIN. E-MAIL: PLOPEZ@CICA.ES