MATH. SCAND. 89 (2001), 245–267

THE NEVANLINNA PARAMETRIZATION FOR A MATRIX MOMENT PROBLEM

PEDRO LOPEZ-RODRIGUEZ*

Abstract

We obtain the Nevanlinna parametrization for an indeterminate matrix moment problem, giving a homeomorphism between the set V of solutions to the matrix moment problem and the set $\mathscr{V}$ of analytic matrix functions in the upper half plane such that $V(\lambda)^* V(\lambda) \leq I$. We characterize the N-extremal matrices of measures (those for which the space of matrix polynomials is dense in their L^2-space) as those whose corresponding matrix function $V(\lambda)$ is a constant unitary matrix.

1. Introduction

Let $(s_n)_{n\geq 0}$ be an indeterminate Hamburger moment sequence and let V be the set of positive Borel measures μ on R satisfying $\int_{\mathsf{R}} t^n \, d\mu(t) = s_n$, $n \geq 0$. It is clear that V is an infinite convex set, and well-known that V is compact in the weak topology. Let $(p_n)_n$ be the corresponding orthonormal polynomials satisfying

$$\int_{\mathsf{R}} p_n(t) p_m(t) \, d\mu(t) = \delta_{n,m}, \qquad \text{for any} \quad \mu \in V.$$

$(p_n)_n$ is uniquely determined if we assume that p_n is of degree n with positive leading coefficient. The polynomials $(q_n)_n$ of the second kind are given by

$$q_n(t) = \int_{\mathsf{R}} \frac{p_n(t) - p_n(x)}{t - x} \, d\mu(x), \qquad \text{for any} \quad \mu \in V.$$

It is well-known that the series $\sum |p_n(\lambda)|^2$, $\sum |q_n(\lambda)|^2$ converge uniformly on compact subsets of C, which makes it possible to define four important entire functions on C by

$$a(\lambda) = \lambda \sum_{k=0}^{\infty} q_k(0) q_k(\lambda), \qquad b(\lambda) = -1 + \lambda \sum_{k=0}^{\infty} q_k(0) p_k(\lambda),$$

* This work has been partially supported by DGES ref. PB96-1321-C02.
Received December 17, 1997; in revised form December 4, 1998.

$$c(\lambda) = 1 + \lambda \sum_{k=0}^{\infty} p_k(0) q_k(\lambda), \qquad d(\lambda) = \lambda \sum_{k=0}^{\infty} p_k(0) p_k(\lambda).$$

These functions depend only on the moment sequence $(s_n)_{n\geq 0}$ (or equivalently on V).

The set V of all solutions μ to the indeterminate moment problem was parametrized by Nevanlinna in 1922 using these functions. The parameter space is the one-point compactification of the set $\mathscr{P}$ of Pick functions, which are holomorphic functions in the upper half-plane H with non-negative imaginary part. Pick functions are also called Herglotz or Nevanlinna functions. The Nevanlinna parametrization is the homeomorphism $\varphi \to \nu_\varphi$ of $\mathscr{P} \cup \{\infty\}$ onto V given by

$$\text{(1.1)} \qquad \int_{\mathsf{R}} \frac{d\nu_\varphi(t)}{t-\lambda} = -\frac{a(\lambda)\varphi(\lambda) - c(\lambda)}{b(\lambda)\varphi(\lambda) - d(\lambda)}, \qquad \text{for} \quad \lambda \in \mathsf{C} \setminus \mathsf{R},$$

which means that the Stieltjes transform of any solution $\nu \in V$ is given by (1.1) for a unique Pick function φ or by the point ∞ (see [1] or [14]).

Strictly speaking it is not the set V which is parametrized but the set of its Stieltjes transforms

$$I(\mu)(\lambda) = \int_{\mathsf{R}} \frac{d\mu(t)}{t-\lambda}, \qquad \lambda \in \mathsf{C} \setminus \mathsf{R},$$

which are holomorphic functions in $\mathsf{C} \setminus \mathsf{R}$. This is just as good, since $\mu \to I(\mu)$ is a one-to-one mapping from the set $M(\mathsf{R})$ of finite complex measures on R to the set $\mathscr{H}(\mathsf{C} \setminus \mathsf{R})$ of holomorphic functions in $\mathsf{C} \setminus \mathsf{R}$. The inverse mapping is given by the Perron-Stieltjes inversion formula

$$\mu = \lim_{\epsilon \to 0^+} \frac{1}{2\pi i} \left\{ I(\mu)(x+i\epsilon) - I(\mu)(x-i\epsilon) \right\},$$

where the convergence is in the weak topology on the space $M(\mathsf{R})$ as dual space of $\mathscr{C}_0(\mathsf{R})$ (continuous functions on R vanishing at infinity).

The measures in V for which the set P of polynomials is dense in its corresponding space $L^2(\mu)$ are those whose corresponding Pick functions are real constants or ∞. These constant real functions are extremal in $\mathscr{P}$ in an obvious sense. In honour of Nevanlinna these measures are called N-extremal (Nevanlinna-extremal). Nevanlinna also proved that for a fixed $\lambda \in \mathsf{C} \setminus \mathsf{R}$, the set

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

is a circle in the complex plane, and in 1923 M. Riesz gave a beautiful geometric characterization of the N-extremal measures, proving that a measure ν is N-extremal if and only if $I(\nu)(\lambda)$ lies on the boundary of the circle $I(V)(\lambda)$, and that this property does not depend on the chosen λ (see [1] or [16]). In the parametrization (1.1), these N-extremal measures are obtained when $\varphi(\lambda)$ is taken to be a constant real number.

If we define $\mathscr{V}$ to be the set of holomorphic functions $v(\lambda)$ in the upper half-plane H such that $|v(\lambda)|^2 = v(\lambda)\overline{v(\lambda)} \leq 1$, then the mapping

$$v(\lambda) = -[\varphi(\lambda) + i]^{-1}[\varphi(\lambda) - i]$$

transforms the set $\mathscr{P} \cup \{\infty\}$ onto $\mathscr{V}$ bijectively, if we accept that the limit function $\varphi(\lambda) = \infty$ is transformed into $v(\lambda) = -1$. Its inverse is given by

$$\varphi(\lambda) = i[1 - v(\lambda)][1 + v(\lambda)]^{-1}. \tag{1.2}$$

If we make this change in (1.1) we obtain the expression

$$\int_{\mathsf{R}} \frac{d\nu(t)}{t - \lambda} = -\frac{a(\lambda)i[1 - v(\lambda)] - c(\lambda)[1 + v(\lambda)]}{b(\lambda)i[1 - v(\lambda)] - d(\lambda)[1 + v(\lambda)]}, \qquad \text{for} \quad \lambda \in \mathsf{H},$$

and the N-extremal measures are obtained when $v(\lambda)$ is a constant complex number a with $|a| = 1$. This expression is more suitable to be generalized to the matrix case. The reason is that, whereas in the scalar case there is only one limit Pick function $\varphi(\lambda) = \infty$, in the matrix case a Pick matrix function can be "big" in many different ways, as we will see later.

The purpose of this paper is to generalize the parametrization of Nevanlinna to the matrix case, and to characterize those matrices of measures for which the matrix polynomials are dense in the corresponding L^2-space, that is, the N-extremal matrices of measures.

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 $S_k = \int_{\mathsf{R}} t^k \, d\nu(t)$ 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 definite matrices of measures having the same moments as those of ν up to degree n.

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

$$tP_n(t) = A_{n+1}P_{n+1}(t) + B_nP_n(t) + A_n^*P_{n-1}(t), \qquad n \geq 0, \tag{1.3}$$

(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 [3] or [7]).

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

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

which also satisfy the recurrence relation (1.3), with initial conditions $Q_0(t) = \theta$ and $Q_1(t) = A_1^{-1}$.

As in the scalar case the determinacy or indeterminacy of the matrix moment problem is also related to the indices 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 (1.3). In this case the indices of deficiency can be any natural number from 0 to N, being both equal to 0 in the determinate case and both equal to N in the completely indeterminate case.

In this paper we assume the matrix moment problem has the highest possible degree of indetermination, that is, these indices are both equal to N. In section 3 we will prove that in this case the two series

$$(1.4) \qquad \sum_{k=0}^{\infty} Q_k^*(\lambda) P_k(\eta) \qquad \text{and} \qquad \sum_{k=0}^{\infty} P_k^*(\lambda) P_k(\eta)$$

converge uniformly in the variables λ and η on every bounded set of the complex plane.

This permits to define four holomorphic matrix functions $A(\lambda)$, $B(\lambda)$, $C(\lambda)$ and $D(\lambda)$ which depend only on the sequence $(S_n)_{n\geq 0}$ and play a fundamental role in the parametrization of the solutions to the matrix moment problem (see (2.16)).

In the matrix case the parameter space is the space $\mathscr{V}$ of holomorphic matrix functions $V(\lambda)$ in the upper half-plane H such that $V(\lambda)^* V(\lambda) \leq I$. We will prove the following theorem:

THEOREM 1.1. *There exists a homeomorphism between the set V and the set $\mathscr{V}$ given by*

$$\int_{\mathsf{R}} \frac{d\nu(t)}{t-\lambda} = -\left\{C^*(\lambda)[I+V(\lambda)] - iA^*(\lambda)[I-V(\lambda)]\right\} \left\{D^*(\lambda)[I+V(\lambda)] - iB^*(\lambda)[I-V(\lambda)]\right\}^{-1}. \tag{1.5}$$

The N-extremal matrices of measures in V correspond to the constant unitary matrices in $\mathscr{V}$.

In most cases this expression can be given in terms of a Pick matrix function. A Pick matrix function is a holomorphic matrix function $\Phi(\lambda)$ in the upper half-plane H such that for any z in H the matrix

$$\operatorname{Im}\Phi(\lambda) = \frac{\Phi(\lambda)-\Phi(\lambda)^*}{2i}$$

is positive semidefinite.

If we suppose the matrix function $[I+V(\lambda)]$ to be invertible for every λ in H, then we can define

$$\Phi(\lambda) = i[I-V(\lambda)][I+V(\lambda)]^{-1}, \tag{1.6}$$

which is a Pick matrix function:

$$\begin{aligned}\frac{\Phi(\lambda)-\Phi(\lambda)^*}{2i} &= \frac{1}{2i}\left\{i[I-V(\lambda)][I+V(\lambda)]^{-1}\right.\\ &\qquad \left.+\, i[I+V(\lambda)^*]^{-1}[I-V(\lambda)^*]\right\}\\ &= \frac{1}{2}[I+V(\lambda)^*]^{-1}\left\{[I+V(\lambda)^*][I-V(\lambda)]\right.\\ &\qquad \left.+\,[I-V(\lambda)^*][I+V(\lambda)]\right\}[I+V(\lambda)]^{-1}\\ &= [I+V(\lambda)^*]^{-1}\left\{I-V(\lambda)^*V(\lambda)\right\}[I+V(\lambda)]^{-1}\\ &\geq 0\end{aligned} \tag{1.7}$$

because $V(\lambda)$ belongs to $\mathscr{V}$. In this case (1.5) becomes

$$\int_{\mathsf{R}} \frac{d\nu(t)}{t-\lambda} = -\left\{C^*(\lambda) - A^*(\lambda)\Phi(\lambda)\right\}\left\{D^*(\lambda) - B^*(\lambda)\Phi(\lambda)\right\}^{-1},$$

which is the matrix version of (1.1). For any given Pick matrix function $\Phi(\lambda)$, the inverse mapping of (1.6) is

$$V(\lambda) = -[\Phi(\lambda)+iI]^{-1}[\Phi(\lambda)-iI].$$

Observe that $[\Phi(\lambda) + iI]$ is always invertible, for if there exists λ_0 in H such that $\det[\Phi(\lambda_0) + iI] = 0$, then there exists a non-zero vector v in C^N such that $\Phi(\lambda_0)v^* = -iv^*$ (vectors v in C^N are considered as row vectors, and v^* is the column of complex conjugate entries of v) . Using this we get

$$v\left(\frac{\Phi(\lambda_0) - \Phi(\lambda_0)^*}{2i}\right)v^* = -vv^* < 0$$

which is absurd because $\Phi(\lambda)$ is a Pick matrix function.

For $V(\lambda)$ defined in this way, we have

$$I + V(\lambda) = I - [\Phi(\lambda) + iI]^{-1}[\Phi(\lambda) - iI] = [\Phi(\lambda) + iI]^{-1}2iI,$$

which is always invertible. Consequently (1.7) holds for every λ in H and it immediately gives that $V(\lambda)^*V(\lambda) \leq I$, that is, $V(\lambda)$ belongs to $\mathscr{V}$.

If ν is N-extremal, then its Stieltjes transform is

$$\int_{\mathsf{R}} \frac{d\nu(t)}{t-\lambda} = -\left\{C^*(\lambda)[I+U] - iA^*(\lambda)[I-U]\right\} \left\{D^*(\lambda)[I+U] - iB^*(\lambda)[I-U]\right\}^{-1} \tag{1.8}$$

for a certain unitary matrix U. If $I + U$ is invertible, then

$$H = i[I-U][I+U]^{-1}$$

is hermitian, and (1.8) reduces to

$$\int_{\mathsf{R}} \frac{d\nu(t)}{t-\lambda} = -\left\{C^*(\lambda) - A^*(\lambda)H\right\}\left\{D^*(\lambda) - B^*(\lambda)H\right\}^{-1}, \tag{1.9}$$

but observe that not every N-extremal matrix of measures can be represented in this way for a hermitian matrix.

During the forties the matrix moment problem appears mentioned in several papers by a few soviet authors, who obtained some results using operator theory. Indeed, Theorem 1.1 appears without proof in the 1949 paper [10] by Krein, who refers to previous papers on operator theory by himself and M. Krasnoselskii ([11] and [12]) and to a paper by Nagel ([15]) from 1936. As far as we know, no proof of Theorem 1.1 has ever appeared published. We can only think that Krein saw this result as a consequence of his investigations in operator theory. The proof we present in this paper is obtained without using any techniques of operator theory, and together with our previous paper ([13]), it gives a direct extension and a geometric interpretation to the matrix case of the 1922 and 1923 results of R. Nevanlinna and M. Riesz.

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 $P(\lambda)^* = P^*(\overline{\lambda})$. If $F(\lambda)$ is a holomorphic function on a domain Ω containing 0 we denote by $F^*(\lambda)$ the matrix function obtained from $F(\lambda)$ by replacing each of the matrix coefficients in its power series expansion at zero by its hermitian conjugate, and similarly we have $F(\lambda)^* = F^*(\overline{\lambda})$. We say that λ_0 is a zero of the analytic matrix function $F(\lambda)$ if $\det F(\lambda_0) = 0$.

The set of positive definite matrices of measures is endowed with the vague and weak topologies. The vague topology is the coarsest topology for which the mappings $\mu \to \int_{\mathsf{R}} f d\mu$ are continuous, where $f \in \mathscr{C}_c(\mathsf{R})$ is arbitrary. By $\mathscr{C}_c(\mathsf{R})$ we denote the set of continuous functions with compact support defined on R. The weak topology is the coarsest topology for which the mappings $\mu \to \int_{\mathsf{R}} f d\mu$ are continuous, where $f \in \mathscr{C}_b(\mathsf{R})$ is arbitrary. By $\mathscr{C}_b(\mathsf{R})$ we denote the set of continuous and bounded functions defined on R.

Since $\mathscr{C}_c(\mathsf{R})$ is strictly included in $\mathscr{C}_b(\mathsf{R})$ it is clear that the vague topology is coarser, that is, it has fewer open sets than the weak topology. It is not hard to see that both topologies are Hausdorff. V is a compact and convex set for these topologies which coincide on V (see [8]).

For μ a positive definite matrix of measures, the space $L^2(\mu)$ is defined as the set of $N \times N$ matrix functions $f : \mathsf{R} \to M_{N\times N}(\mathsf{C})$ 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 $A = (a_{i,j})_{1\le i,j\le N}$, 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

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

and is a Hilbert space. The duality works as for the scalar case (see [17] or [8] for more details). For the definition of the L^p spaces associated to μ, $1 \le p < \infty$, see also [8].

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

$$\begin{aligned}\mathscr{A}_n(u,v) &= (v-u)\sum_{k=0}^{n-1} Q_k^*(u)Q_k(v) \\ &= Q_{n-1}^*(u)A_nQ_n(v) - Q_n^*(u)A_n^*Q_{n-1}(v), \quad \text{for} \quad u,v\in \mathbb{C},\end{aligned} \tag{2.1}$$

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

(this is Green's formula),

$$\begin{aligned}\mathscr{C}_n(u,v) &= I + (v-u)\sum_{k=0}^{n-1} P_k^*(u)Q_k(v) \\ &= P_{n-1}^*(u)A_nQ_n(v) - P_n^*(u)A_n^*Q_{n-1}(v), \quad \text{for} \quad u,v\in \mathbb{C},\end{aligned} \tag{2.3}$$

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

(this is Christoffel-Darboux formula).

By using the Liouville-Ostrogradsky formula

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

and

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

one can obtain the three following formulae strightforwardly:

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

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

and

$$\mathscr{A}_n(u,v)\mathscr{B}_n^*(u,v) = \mathscr{B}_n(u,v)\mathscr{A}_n^*(u,v), \qquad \text{for} \quad u,v\in\mathbb{C}. \tag{2.9}$$

We define the four matrix functions

$$\begin{aligned} A_n(\lambda) &= \mathscr{A}_n(0,\lambda), & B_n(\lambda) &= \mathscr{B}_n(0,\lambda), \\ C_n(\lambda) &= \mathscr{C}_n(0,\lambda), & D_n(\lambda) &= \mathscr{D}_n(0,\lambda). \end{aligned} \tag{2.10}$$

For $u = 0$ and $v = \lambda$ in (2.7), (2.8) and (2.9) we get the identities

$$A_n(\lambda)D_n^*(\lambda) - B_n(\lambda)C_n^*(\lambda) = I, \qquad \text{for} \quad \lambda \in \mathbb{C}, \tag{2.11}$$

$$C_n(\lambda)D_n^*(\lambda) = D_n(\lambda)C_n^*(\lambda), \qquad \text{for} \quad \lambda \in \mathbb{C}, \tag{2.12}$$

and

$$A_n(\lambda)B_n^*(\lambda) = B_n(\lambda)A_n^*(\lambda), \qquad \text{for} \quad \lambda \in \mathbb{C}. \tag{2.13}$$

By using (2.5) and (2.6) we obtain the four following formulae:

$$\begin{aligned} P_n(\lambda) &= Q_n(0)D_n(\lambda) - P_n(0)B_n(\lambda), & &\text{for} \quad \lambda \in \mathbb{C}, \\ P_{n-1}(\lambda) &= Q_{n-1}(0)D_n(\lambda) - P_{n-1}(0)B_n(\lambda), & &\text{for} \quad \lambda \in \mathbb{C}, \\ Q_n(\lambda) &= Q_n(0)C_n(\lambda) - P_n(0)A_n(\lambda), & &\text{for} \quad \lambda \in \mathbb{C}, \\ Q_{n-1}(\lambda) &= Q_{n-1}(0)C_n(\lambda) - P_{n-1}(0)A_n(\lambda), & &\text{for} \quad \lambda \in \mathbb{C}. \end{aligned} \tag{2.14}$$

We also have

$$B_n^*(\lambda)D_n(\overline{\lambda}) - D_n^*(\lambda)B_n(\overline{\lambda}) = 2i\,\operatorname{Im}\lambda \sum_{k=0}^{n-1} P_k^*(\lambda)P_k(\overline{\lambda}), \tag{2.15}$$

for $\lambda \in \mathbb{C}$.

As we explained in the Introduction, we can define four holomorphic matrix functions $A(\lambda)$, $B(\lambda)$, $C(\lambda)$ and $D(\lambda)$ by taking limit in (2.10) for n tending to ∞:

$$\begin{aligned} A(\lambda) &= \lambda \sum_{k=0}^{\infty} Q_k^*(0)Q_k(\lambda), & B(\lambda) &= -I + \lambda \sum_{k=0}^{\infty} Q_k^*(0)P_k(\lambda), \\ C(\lambda) &= I + \lambda \sum_{k=0}^{\infty} P_k^*(0)Q_k(\lambda), & D(\lambda) &= \lambda \sum_{k=0}^{\infty} P_k^*(0)P_k(\lambda). \end{aligned} \tag{2.16}$$

By taking limit in (2.11), (2.12), (2.13) and (2.15) we get the following formulae:

$$A(\lambda)D^*(\lambda) - B(\lambda)C^*(\lambda) = I, \qquad \text{for} \quad \lambda \in \mathbb{C}, \tag{2.17}$$

$$C(\lambda)D^*(\lambda) = D(\lambda)C^*(\lambda), \qquad \text{for} \quad \lambda \in \mathsf{C}, \tag{2.18}$$

$$A(\lambda)B^*(\lambda) = B(\lambda)A^*(\lambda), \qquad \text{for} \quad \lambda \in \mathsf{C}, \tag{2.19}$$

and

$$B^*(\lambda)D(\overline{\lambda}) - D^*(\lambda)B(\overline{\lambda}) = 2i \operatorname{Im} \lambda \sum_{k=0}^{\infty} P_k^*(\lambda)P_k(\overline{\lambda}), \tag{2.20}$$

for $\lambda \in \mathsf{C}$.

If we suppose $Q_{n-1}(0)$ is invertible we can express

$$B_n(\lambda) = Q_{n-1}^*(0)A_n\{P_n(\lambda) - A_n^{-1}Q_{n-1}^*(0)^{-1}Q_n^*(0)A_n^*P_{n-1}(\lambda)\}.$$

Putting $u = v = 0$ in (2.1) we get that the matrix $Q_{n-1}^*(0)^{-1}Q_n^*(0)A_n^*$ is hermitian and then remark 2.3 of [7] gives that the zeros of $B_n(\lambda)$ are all real.

If $Q_{n-1}(0)$ is not invertible, then $Q_{n-1}(t)$ is invertible, for $|t|$ sufficiently small, for the number of zeros of $Q_{n-1}(t)$ is at most $(n-1)N$, and then we can express

$$\begin{aligned} Q_{n-1}^*(t)A_nP_n(\lambda) - Q_n^*(t)A_n^*P_{n-1}(\lambda) \\ = Q_{n-1}^*(t)A_n\{P_n(\lambda) - A_n^{-1}Q_{n-1}^*(t)^{-1}Q_n^*(t)A_n^*P_{n-1}(\lambda)\}. \end{aligned}$$

Again, putting $u = v = t$ in (2.1) we see that the matrix $Q_{n-1}^*(t)^{-1}Q_n^*(t)A_n^*$ is hermitian and consequently the zeros of $Q_{n-1}^*(t)A_nP_n(\lambda) - Q_n^*(t)A_n^*P_{n-1}(\lambda)$ are real. Since the zeros of $B_n(\lambda)$ are obtained for $|t|$ tending to 0, these must be real as well. As a consequence the zeros of the limit matrix function $B(\lambda) = \lim_{n\to\infty} B_n(\lambda)$ are also real.

In a similar way, taking into account that $Q_n(t)$ is also a sequence of orthogonal matrix polynomials we deduce that the zeros of $A_n(\lambda)$, $C_n(\lambda)$, $D_n(\lambda)$, $A(\lambda)$, $C(\lambda)$ and $D(\lambda)$ are all real.

In [13] the sets $\mathsf{B}_n(\lambda)$ and $\mathsf{B}_\infty(\lambda)$ were defined and used. For $\lambda \in \mathsf{C} \setminus \mathsf{R}$, $\mathsf{B}_n(\lambda)$ is the set of $N \times N$ complex matrices ω such that the matrix inequality

$$[\omega + G_n(\lambda)]R_{n-1}(\overline{\lambda})^{-1}[\omega + G_n(\lambda)]^* \le |\lambda - \overline{\lambda}|^{-2}R_{n-1}(\lambda) \tag{2.21}$$

holds, where $G_n(\lambda) = \mathscr{B}_n(\lambda, \overline{\lambda})\mathscr{D}_n(\lambda, \overline{\lambda})^{-1}$ and

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

Equation (2.21) is equivalent to

$$\sum_{k=0}^{n-1}(Q_k^*(\lambda)+\omega P_k^*(\lambda))(Q_k(\overline{\lambda})+P_k(\overline{\lambda})\omega^*)\leq\frac{\omega-\omega^*}{\lambda-\overline{\lambda}}. \tag{2.22}$$

The set $\mathsf{B}_\infty(\lambda)$ is the intersection of all the sets $\mathsf{B}_n(\lambda)$, for $n\in\mathsf{N}$. Its equation is

$$[\omega+G(\lambda)]R(\overline{\lambda})^{-1}[\omega+G(\lambda)]^*\leq|\lambda-\overline{\lambda}|^{-2}R(\lambda), \tag{2.23}$$

where $G(\lambda)=\mathscr{B}(\lambda,\overline{\lambda})\mathscr{D}(\lambda,\overline{\lambda})^{-1}$ and $R(\lambda)$ is the limit matrix

$$R(\lambda)=\lim_{n\to\infty}R_n(\lambda),$$

or equivalently

$$\sum_{k=0}^{\infty}(Q_k^*(\lambda)+\omega P_k^*(\lambda))(Q_k(\overline{\lambda})+P_k(\overline{\lambda})\omega^*)\leq\frac{\omega-\omega^*}{\lambda-\overline{\lambda}}. \tag{2.24}$$

By using formulae (2.1), (2.2), (2.3) and (2.4) in (2.22) it is straightforward to see that $\mathsf{B}_n(\lambda)$ is also the set of $N\times N$ complex matrices ω such that

$$\frac{\operatorname{Im}\{[Q_{n-1}^*(\lambda)+\omega P_{n-1}^*(\lambda)]A_n[P_n(\overline{\lambda})\omega^*+Q_n(\overline{\lambda})]\}}{\operatorname{Im}\lambda}\geq\theta. \tag{2.25}$$

The condition for ω to be an extremal point of $\mathsf{B}_n(\lambda)$ (in the sense of convexity) is that the matrix

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

is hermitian.

By substituting formulae (2.14) in (2.25) we get that $\mathsf{B}_n(\lambda)$ is also described by the matrix inequality

$$\frac{\operatorname{Im}\{[C_n^*(\lambda)+\omega D_n^*(\lambda)][A_n(\overline{\lambda})+B_n(\overline{\lambda})\omega^*]\}}{\operatorname{Im}\lambda}\geq\theta, \tag{2.26}$$

and by taking limit for n tending to ∞ we get that $\mathsf{B}_\infty(\lambda)$ is described by the inequality

$$\frac{\operatorname{Im}\{[C^*(\lambda)+\omega D^*(\lambda)][A(\overline{\lambda})+B(\overline{\lambda})\omega^*]\}}{\operatorname{Im}\lambda}\geq\theta. \tag{2.27}$$

In [13] it is proved that for a fixed non-real λ the image through the Stieltjes transform of all the matrices of measures of V in the point λ

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

is the set $\mathsf{B}_\infty(\lambda)$. This is the generalization to the matrix case of the same theorem proved by Nevanlinna in 1922 for the scalar case.

The extremal points (in the sense of convexity) of the set $I(V)(\lambda) = \mathsf{B}_\infty(\lambda)$ are the matrices ω for which equality is attained in (2.27), that is the $N \times N$ complex matrices ω such that the matrix

$$[C^*(\lambda) + \omega D^*(\lambda)][A(\overline{\lambda}) + B(\overline{\lambda})\omega^*]$$

is hermitian. If μ is a matrix of measures in V for which $I(\mu)(\lambda)$ is an extremal point of $I(V)(\lambda)$, we call this matrix of measures N-extremal, as in the scalar case.

Finally, [13] generalizes Riesz's theorem by proving that the matrices of measures of V for which the set P of matrix polynomials is dense in the corresponding space $L^2(\mu)$ 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. The questions of density for the truncated matrix moment problem were solved in [9].

3. The indices of deficiency

Given a complex λ in the upper half-plane, the index of deficiency δ_+ is the dimension of the kernel of the operator $J^* - \lambda I$. The index δ_- is defined in the same way for λ in the lower half-plane. It is a well-known result in operator theory (see [2]) that the indices δ_+ and δ_- do not depend on the λ chosen in the upper or lower half-plane respectively.

If we solve the equation $(J^* - \lambda I)x^* = \theta$ (x^* denotes an infinite column vector), we obtain that the solutions x^* must be necessarily of the form

$$x^* = \begin{pmatrix} P_0(\lambda)v^* \\ P_1(\lambda)v^* \\ \vdots \end{pmatrix},$$

for a certain vector v in C^N. The condition for x^* to belong to ℓ^2 is

$$\sum_{k=0}^{\infty} \|P_k(\lambda)v^*\|_2^2 < \infty,$$

and consequently the index of deficiency δ_+ is

$$\begin{aligned}\delta_+ &= \dim\left\{v \in \mathbb{C}^N : \sum_{k=0}^{\infty} \| P_k(\lambda)v^* \|_2^2 < \infty\right\} \\ &= \dim\left\{v \in \mathbb{C}^N : vR_n(\lambda)^{-1}v^* \text{ is bounded}\right\}.\end{aligned}$$

We have the following results:

THEOREM 3.1. *The index of deficiency of the operator $J^* - \lambda I$ is equal to the rank of the limit matrix $R(\lambda)$.*

This theorem appears in [18]. For the convenience of the reader we include a proof of it here. A more general result can be found in [6, Th. 2.6].

PROOF. We prove that a vector v in $\mathbb{C}^N$ belongs to the image of $R(\lambda)$ if and only if $vR_n(\lambda)^{-1}v^*$ is a bounded sequence. Observe that since $R(\lambda)$ is a hermitian matrix, the subspaces $\mathrm{Im}(R(\lambda))$ and $\mathrm{Ker}(R(\lambda))$ of $\mathbb{C}^N$ are orthogonal complements of each other.

($\Longrightarrow$) We prove first that for any $n \geq 1$ we have

$$R(\lambda)R_n(\lambda)^{-1}R(\lambda) \leq R(\lambda). \tag{3.1}$$

To see this observe that for $n \geq 1$ we have $R_n(\lambda) \geq R(\lambda)$ and thus for any $\epsilon > 0$ we have $R_n(\lambda) + \epsilon I \geq R(\lambda) + \epsilon I$. Since $R(\lambda) + \epsilon I$ is invertible, we obtain

$$(R_n(\lambda) + \epsilon I)^{-1} \leq (R(\lambda) + \epsilon I)^{-1},$$

and thus

$$(R(\lambda) + \epsilon I)(R_n(\lambda) + \epsilon I)^{-1}(R(\lambda) + \epsilon I) \leq (R(\lambda) + \epsilon I).$$

Now (3.1) is obtained by letting $\epsilon \to 0$.

We prove now the first implication. Given $v \in \mathrm{Im}(R(\lambda))$, there exists w in $\mathbb{C}^N$ such that $v = wR(\lambda)$, and we have

$$vR_n(\lambda)^{-1}v^* = wR(\lambda)R_n(\lambda)^{-1}R(\lambda)w^* \leq wR(\lambda)w^*,$$

which is bounded.

($\Longleftarrow$) We choose a unitary matrix $U_n(\lambda)$ such that

$$U_n(\lambda)R_n(\lambda)^{-1}U_n(\lambda)^* = \mathrm{diag}(\lambda_1^{(n)}, \dots, \lambda_N^{(n)}),$$

with $\lambda_1^{(n)} \leq \lambda_2^{(n)} \leq \cdots \leq \lambda_N^{(n)}$. A subsequence $U_{n_p}(\lambda)$ of $U_n(\lambda)$ converges to a unitary matrix $U(\lambda)$, and without loss of generality we can assume $U_{n_p}(\lambda)$ is the same $U_n(\lambda)$. We suppose now v is a vector in $\mathbb{C}^N$ such that $vR_n(\lambda)^{-1}v^*$ is bounded, that is,

$$vU_n(\lambda)^* \operatorname{diag}(\lambda_1^{(n)}, \ldots, \lambda_N^{(n)})U_n(\lambda)v^* \leq K,$$

for a certain constant K.

We know that $\lambda_1^{(n)}, \ldots, \lambda_N^{(n)}$ are all increasing sequences, and thus they are convergent, respectively to certain λ_j, for $n \to \infty$, and

$$vU(\lambda)^* \operatorname{diag}(\lambda_1, \ldots, \lambda_N)U(\lambda)v^* \leq K.$$

We call $\tilde{v} = vU(\lambda)^* = (\tilde{v}_1, \ldots, \tilde{v}_N)$, and we get

$$\sum_{j=1}^{N} \lambda_j |\tilde{v}_j|^2 \leq K.$$

If we call $I = \{j = 1, \ldots, N : \quad \lambda_j = \infty\}$, we have that $\tilde{v}_j = 0$ for $j \in I$. Now,

$$U_n(\lambda)R_n(\lambda)U_n(\lambda)^* = \operatorname{diag}\left(\frac{1}{\lambda_1^{(n)}}, \ldots, \frac{1}{\lambda_N^{(n)}}\right),$$

and hence

$$U(\lambda)R(\lambda)U(\lambda)^* = \operatorname{diag}\left(\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_N}\right),$$

where $1/\lambda_j = 0$ for $j \in I$. To finish we see now that $v \in \operatorname{Im}(R(\lambda))$, by proving that $vw^* = 0$, for any $w \in \operatorname{Ker}(R(\lambda))$. For $w \in \operatorname{Ker}(R(\lambda))$ we have

$$0 = wR(\lambda)w^* = \sum_{j=1}^{N} \frac{1}{\lambda_j} |\tilde{w}_j|^2,$$

with $\tilde{w} = wU(\lambda)^*$, hence $\tilde{w}_j = 0$ for $j \notin I$, and consequently

$$vw^* = vU(\lambda)^* U(\lambda)w^* = \tilde{v}\tilde{w}^* = \sum_{j=1}^{N} \tilde{v}_j \overline{\tilde{w}_j} = 0$$

because $\tilde{v}_j = 0$ for $j \in I$ and $\tilde{w}_j = 0$ for $j \in I$.

THEOREM 3.2. *If for a fixed complex λ_0 the series*

$$\sum_{k=0}^{\infty} P_k^*(\overline{\lambda_0}) P_k(\lambda_0)$$

is convergent, then

(1) *The series*

$$\sum_{k=0}^{\infty} Q_k^*(\overline{\lambda_0})Q_k(\lambda_0)$$

is also convergent, and

(2) *The series*

$$\sum_{k=0}^{\infty} P_k^*(\overline{\lambda})P_k(\lambda) \qquad \textit{and} \qquad \sum_{k=0}^{\infty} Q_k^*(\overline{\lambda})Q_k(\lambda)$$

are both uniformly convergent in any compact subset of the complex plane.

PROOF. (1) In the hypothesis, the set $\mathsf{B}_\infty(\overline{\lambda_0})$ described by 2.23 is isomorphic to $H^*H \leq I$. Formula 2.24 gives that for any matrix ω_0 in $\mathsf{B}_\infty(\overline{\lambda_0})$, the series

$$\sum_{k=0}^{\infty}(Q_k^*(\overline{\lambda_0}) + \omega_0 P_k^*(\overline{\lambda_0}))(Q_k(\lambda_0) + P_k(\lambda_0)\omega_0^*)$$

is convergent. To deduce the result it is enough to choose any ω_0 in $\mathsf{B}_\infty(\overline{\lambda_0})$ and to take into account that the sequences of $N \times N$ matrices M_n for which the series $\sum_{n=1}^{\infty} M_n^* M_n$ is convergent form a vector space.

(2) The proof known for the scalar case (see [1, p. 16]) works for the matrix case with minimal adjustments, so we omit it.

By virtue of Theorem 3.2 the claim about the series (1.4) stated in the Introduction follows, which permits to define the four important holomorphic matrix functions (2.16).

4. Pick matrix functions

We denote by $\mathscr{P}$ the set of Pick matrix functions, that is, the set of holomorphic matrix functions $\Phi(\lambda)$ defined on $\mathsf{H} = \{\operatorname{Im}\lambda > 0\}$ such that for any λ in H the matrix

$$\operatorname{Im}\Phi(\lambda) = \frac{\Phi(\lambda) - \Phi(\lambda)^*}{2i}$$

is positive semidefinite. Any Pick matrix function can be extended to the half plane $\mathsf{H}^* = \{\operatorname{Im} z < 0\}$ by putting $\Phi(z) = \Phi(\overline{z})^*$. Of course the functions obtained in this way are not in general analytic continuation of each other. Thus we can assume the space $\mathscr{P}$ to consist of all the functions $\Phi(\lambda)$ holomorphic in $\mathsf{C} \setminus \mathsf{R}$ such that

$$\Phi(\lambda) = \Phi(\bar{\lambda})^* \quad \text{and} \quad \frac{\Phi(\lambda) - \Phi(\lambda)^*}{i \operatorname{Im} \lambda} \geq \theta, \qquad \text{for} \quad \operatorname{Im} \lambda \neq 0.$$

For example, if $\nu(t)$ is a positive matrix of measures with $\tau\nu(\mathsf{R}) < \infty$ (τ denotes the trace of the matrix), then its Stieltjes transform $\omega(\lambda) = \int_{\mathsf{R}} \frac{d\nu(t)}{t - \lambda}$ is a Pick matrix function, for

$$\frac{\omega(\lambda) - \omega(\lambda)^*}{2i} = \int_{\mathsf{R}} \frac{\operatorname{Im} \lambda}{|t - \lambda|^2} \, d\nu(t).$$

We now give some results about Pick matrix functions. We only give the proof of the fourth theorem. For the first three theorems the proofs work as in the scalar case (see [1]). The presentation here follows the treatment of the classical theory as given in [4].

THEOREM 4.1. *The formula*

$$\Phi(\lambda) = \alpha\lambda + \beta + \int_{\mathsf{R}} \frac{t\lambda + 1}{t - \lambda} \, d\nu(t), \qquad \lambda \in \mathsf{H} \tag{4.1}$$

establishes a one-to-one correspondence between Pick matrix functions $\Phi(\lambda)$ and triples (α, β, ν), where α is a positive semidefinite numerical matrix, β is a hermitian matrix and ν is a positive definite matrix of measures such that $\tau\nu(\mathsf{R}) < \infty$.

THEOREM 4.2. *The matrix functions $I(\sigma)$, where σ is a positive definite matrix of measures with $\tau\sigma(\mathsf{R}) < \infty$ are characterized as the Pick matrix functions Φ for which*

$$\Phi(iy) = O\left(\frac{1}{y}\right) \qquad \text{for} \quad y \to \infty. \tag{4.2}$$

THEOREM 4.3. *Let $(S_n)_n$ be a matrix moment sequence. For every representing matrix of measures σ, the Pick matrix function $\Phi(\lambda) = I(\sigma)(\lambda)$ has the asymptotic expansion*

$$\Phi(\lambda) \sim \sum_{n=0}^{\infty} -\frac{S_n}{\lambda^{n+1}}, \qquad \text{for} \quad |\lambda| \to \infty \tag{4.3}$$

in any V_δ, $0 < \delta \leq \frac{\pi}{2}$, where the set V_δ is

$$V_\delta = \{z \in \mathsf{C} \setminus \{0\} \text{ such that } \delta \leq \arg z \leq \pi - \delta\}, \qquad 0 < \delta \leq \frac{\pi}{2}.$$

Reciprocally, if $(S_n)_{n\geq 0}$ *is a sequence of matrices and* $\Phi(\lambda)$ *is a Pick matrix function such that (4.3) holds in* $V_{\frac{\pi}{2}}$, *then* $(S_n)_n$ *is a matrix moment sequence and* $\Phi(\lambda) = I(\sigma)(\lambda)$ *for some representing measure* σ.

We define the set $\mathscr{N}$ of holomorphic functions to be

$$\mathscr{N} = \{F \in \mathscr{H}(\mathsf{H}) : \forall\lambda \in \mathsf{H},\ F(\lambda) \in \mathsf{B}_\infty(\lambda)\},$$

endowed with the usual topology of uniform convergence on compact subsets of H.

THEOREM 4.4. *The mapping* $I : V \to \mathscr{N}$ *given by*

$$I(\sigma)(\lambda) = \int_{\mathsf{R}} \frac{d\sigma(t)}{t-\lambda}$$

is a homeomorphism. As a consequence $\mathscr{N}$ *is compact.*

PROOF. We already know that if $\sigma \in V$ then $I(\sigma)(\lambda) \in \mathsf{B}_\infty(\lambda)$ (see Preliminaries). The mapping is one-to-one by the Perron-Stieltjes inversion formula. To see it is onto it suffices to show that any $F \in \mathscr{N}$ has the asymptotic expansion

$$F(iy) \sim \sum_{n=0}^{\infty} -\frac{S_n}{(iy)^{n+1}}, \qquad \text{for} \quad y \to \infty.$$

Since every $I(\sigma)$ has this asymptotic expansion for $\sigma \in V$, it is enough to prove that

$$\lim_{y\to\infty} y^n[I(\sigma)(iy) - F(iy)] = \theta, \qquad \text{for} \quad n \in \mathsf{N}.$$

We have

$$y^n[I(\sigma)(iy) - F(iy)] = y^n[I(\sigma)(iy) + G(iy)] - y^n[G(iy) + F(iy)]. \tag{4.4}$$

We consider the Frobenius norm $\|A\| = \tau(AA^*)^{\frac{1}{2}}$. We have to prove that the norms of the two matrices summing in the right hand side of (4.4) tend to 0 when $y \to \infty$. For the first one, observe that $I(\sigma)(iy) \in \mathsf{B}_\infty(iy)$ (2.23) implies that

$$\left\|[\omega + G(iy)]R(-iy)^{-\frac{1}{2}}\right\| \leq \frac{1}{2|y|}\left\|R(iy)^{\frac{1}{2}}\right\|.$$

Using this we have

$$
\begin{aligned}
&\left\| y^n[I(\sigma)(iy) + G(iy)]\right\| \\
&\quad = |y|^n \left\| y^n[I(\sigma)(iy) + G(iy)]R(-iy)^{-\frac{1}{2}} R(-iy)^{\frac{1}{2}} \right\| \\
&\quad \le \frac{1}{2}|y|^{n-1} \left\| R(iy)^{\frac{1}{2}} \right\| \left\| R(-iy)^{\frac{1}{2}} \right\| \\
&\quad = \frac{1}{2}|y|^{n-1} \left\{ \tau \left(\sum_{k=0}^{\infty} P_k^*(-iy)P_k(iy) \right)^{-1} \right\}^{\frac{1}{2}} \left\{ \tau \left(\sum_{k=0}^{\infty} P_k^*(iy)P_k(-iy) \right)^{-1} \right\}^{\frac{1}{2}} \\
&\quad \le \frac{1}{2}|y|^{n-1} \left\{ \tau(P_n^*(-iy)P_n(iy))^{-1} \right\}^{\frac{1}{2}} \left\{ \tau(P_n^*(iy)P_n(-iy))^{-1} \right\}^{\frac{1}{2}}.
\end{aligned}
$$

Since $P_n(\lambda)$ has a non-singular leading coefficient, $\{\tau(P_n^*(-iy)P_n(iy))^{-1}\}^{\frac{1}{2}}$ and $\{\tau(P_n^*(iy)P_n(-iy))^{-1}\}^{\frac{1}{2}}$ are both equivalent to y^{-n} for $y \to \infty$, and consequently the above expresion tends to 0 for $y \to \infty$.

For the second matrix the proof works the same since by hypothesis $F(iy) \in B_\infty(iy)$.

We see now that the mapping is continuous. Let's suppose $\sigma_n \to \sigma$ weakly for σ_n, σ in V. This gives that $I(\sigma_n)(\lambda) \to I(\sigma)(\lambda)$ for all $\lambda \in \mathsf{H}$. We have to show that $I(\sigma_n)(\lambda)$ converges to $I(\sigma)(\lambda)$ uniformly on any compact subset K of H. If $\delta = \inf\{\operatorname{Im} z : z \in K\}$, then $\delta > 0$ and $|t - \lambda| \ge \delta$, for $t \in \mathsf{R}$, $z \in K$. For a given $\epsilon > 0$, there exists finitely many points $z_1, \ldots, z_p$ in K such that

$$
K \subseteq \bigcup_{i=1}^{p} D(z_i, \epsilon),
$$

and we can choose N in N such that

$$
\|I(\sigma)(z_i) - I(\sigma_n)(z_i)\| < \epsilon, \qquad \text{for} \quad n \ge N,\ i = 1, \ldots p.
$$

For $z \in K$ we choose i so that $z \in D(z_i, \epsilon)$ and get for $n \ge N$

$$
\begin{aligned}
\|I(\sigma)(z) - I(\sigma_n)(z)| &\le \left\| I(\sigma - \sigma_n)(z) - I(\sigma - \sigma_n)(z_i) \right\| + \epsilon \\
&\le \left\| \int_{\mathsf{R}} \left| \frac{1}{t - z} - \frac{1}{t - z_i} \right| d(\sigma + \sigma_n)(t) \right\| + \epsilon \\
&= \left\| \int_{\mathsf{R}} \frac{|z - z_i|}{|t - z||t - z_i|} d(\sigma + \sigma_n)(t) \right\| + \epsilon \\
&\le \epsilon \left(\frac{2}{\delta^2} \|S_0\| + 1 \right)
\end{aligned}
$$

which shows that

$$\lim_{n\to\infty} \sup_{z\in K} \left\| I(\sigma)(z) - I(\sigma_n)(z) \right\| = 0.$$

A continuous bijection of a compact space onto a Hausdorff space is a homeomorphism and the image $\mathscr{N}$ is compact. This completes the proof.

5. Proof of Theorem 1.1

By virtue of Theorem 4.4, to finish proving Theorem 1.1, we only have to prove that there is a bijection between $\mathscr{V}$ and $\mathscr{N}$ given by

$$\text{(5.1)}\quad \omega(\lambda) = -\left\{C^*(\lambda)[I + V(\lambda)] - iA^*(\lambda)[I - V(\lambda)]\right\} \left\{D^*(\lambda)[I + V(\lambda)] - iB^*(\lambda)[I - V(\lambda)]\right\}^{-1},$$

where $\omega(\lambda)$ belongs to $\mathscr{N}$ and $V(\lambda)$ belongs to $\mathscr{V}$. This bijection clearly respects uniform convergence on compact subsets of H.

To begin with, let's see that for a given $V(\lambda)$ in $\mathscr{V}$, (5.1) defines $\omega(\lambda)$ holomorphic in H. For this it is enough to prove that the zeros of the matrix function

$$\left\{D^*(\lambda)[I + V(\lambda)] - iB^*(\lambda)[I - V(\lambda)]\right\}$$

are all real. Suppose on the contrary that there exists λ_0 in H and a non-zero vector v in C^N such that

$$\text{(5.2)}\qquad \left\{D^*(\lambda_0)[I + V(\lambda_0)] - iB^*(\lambda_0)[I - V(\lambda_0)]\right\} v^* = \theta.$$

Since $B^*(\lambda_0)$ is invertible, this is equivalent to

$$B^*(\lambda_0)^{-1}D^*(\lambda_0)[I + V(\lambda_0)]v^* = i[I - V(\lambda_0)]v^*.$$

Using this and (2.20) we obtain

$$\begin{aligned}
v[I + V(\lambda_0)^*] & \left\{D(\overline{\lambda_0})B(\overline{\lambda_0})^{-1} - B^*(\lambda_0)^{-1}D^*(\lambda_0)\right\} [I + V(\lambda_0)]v^* \\
&= v[I + V(\lambda_0)^*]B^*(\lambda_0)^{-1} \left\{B^*(\lambda_0)D(\overline{\lambda_0}) - D^*(\lambda_0)B(\overline{\lambda_0})\right\} B(\overline{\lambda_0})^{-1}[I + V(\lambda_0)]v^* \\
&= 2i(\operatorname{Im}\lambda_0)v[I + V(\lambda_0)^*]B^*(\lambda_0)^{-1} \left(\sum_{k=0}^{\infty} P_k^*(\lambda_0)P_k(\overline{\lambda_0})\right)B(\overline{\lambda_0})^{-1} - [I + V(\lambda_0)]v^*
\end{aligned}$$

which is i multiplied by a non negative number. This must be equal to

$$-iv\left\{[I-V(\lambda_0)^*][I+V(\lambda_0)]+[I+V(\lambda_0)^*][I-V(\lambda_0)]\right\}v^* \\ = -2iv[I-V(\lambda_0)^*V(\lambda_0)]v^*,$$

which is i multiplied by a non-positive number. For these two expressions to coincide it must be $[I+V(\lambda_0)]v^*=\theta$, which together with (5.2) gives

$$\theta = D^*(\lambda_0)[I+V(\lambda_0)]v^* = 2iB^*(\lambda_0)v^*,$$

which is absurd because $B^*(\lambda_0)$ is an invertible matrix.

The inverse mapping of (5.1) is given by

$$V(\lambda) = -\left\{[C^*(\lambda)+\omega(\lambda)D^*(\lambda)]+i[A^*(\lambda)+\omega(\lambda)B^*(\lambda)]\right\}^{-1} \\ \cdot\left\{[C^*(\lambda)+\omega(\lambda)D^*(\lambda)]-i[A^*(\lambda)+\omega(\lambda)B^*(\lambda)]\right\}. \tag{5.3}$$

Let's see that if $\omega(\lambda)$ belongs to $\mathscr{N}$, then (5.3) defines $V(\lambda)$ holomorphic in H, for which it is enough to prove that the zeros of the matrix

$$\left\{[C^*(\lambda)+\omega(\lambda)D^*(\lambda)]+i[A^*(\lambda)+\omega(\lambda)B^*(\lambda)]\right\}$$

are all real. Suppose on the contrary that there exists λ_0 in H and a non-zero vector v in C^N such that

$$v\left\{[C^*(\lambda_0)+\omega(\lambda_0)D^*(\lambda_0)]+i[A^*(\lambda_0)+\omega(\lambda_0)B^*(\lambda_0)]\right\}=\theta,$$

which gives

$$v[C^*(\lambda_0)+\omega(\lambda_0)D^*(\lambda_0)] = -iv[A^*(\lambda_0)+\omega(\lambda_0)B^*(\lambda_0)]. \tag{5.4}$$

From this we deduce that

$$v\,\frac{\operatorname{Im}\left\{[C^*(\lambda_0)+\omega(\lambda_0)D^*(\lambda_0)][A(\overline{\lambda_0})+B(\overline{\lambda_0})\omega^*(\overline{\lambda_0})]\right\}}{\operatorname{Im}\lambda_0}\,v^* \\ = \frac{1}{2i\operatorname{Im}\lambda_0}\left\{v[C^*(\lambda_0)+\omega(\lambda_0)D^*(\lambda_0)][A(\overline{\lambda_0})+B(\overline{\lambda_0})\omega^*(\overline{\lambda_0})]v^*\right. \\ \left. - v[A^*(\lambda_0)+\omega(\lambda_0)B^*(\lambda_0)][C(\overline{\lambda_0})+D(\overline{\lambda_0})\omega^*(\overline{\lambda_0})]v^*\right\} \\ = -\frac{1}{\operatorname{Im}\lambda_0}v[A^*(\lambda_0)+\omega(\lambda_0)B^*(\lambda_0)][A(\overline{\lambda_0})+B(\overline{\lambda_0})\omega^*(\overline{\lambda_0})]v^*. \tag{5.5}$$

But $\omega(\lambda_0)$ belongs to $\mathsf{B}_\infty(\lambda_0)$ because $\omega(\lambda)$ is a function in $\mathscr{N}$, and this means that

$$\frac{\operatorname{Im}\left\{[C^*(\lambda_0)+\omega(\lambda_0)D^*(\lambda_0)][A(\overline{\lambda_0})+B(\overline{\lambda_0})\omega^*(\overline{\lambda_0})]\right\}}{\operatorname{Im}\lambda_0}\geq\theta.$$

Together with (5.5) this gives the only possibility

$$v[A^*(\lambda_0) + \omega(\lambda_0)B^*(\lambda_0)] = \theta,$$

and (5.4) yields

$$v[C^*(\lambda_0) + \omega(\lambda_0)D^*(\lambda_0)] = \theta.$$

From these two equations we get

$$-v\omega(\lambda_0) = vC^*(\lambda_0)D^*(\lambda_0)^{-1} = vA^*(\lambda_0)B^*(\lambda_0)^{-1},$$

which using (2.18) and (2.17) leads to

$$\begin{aligned}
\theta &= v[A^*(\lambda_0)B^*(\lambda_0)^{-1} - C^*(\lambda_0)D^*(\lambda_0)^{-1}] \\
&= v[A^*(\lambda_0)B^*(\lambda_0)^{-1} - D(\lambda_0)^{-1}C(\lambda_0)] \\
&= vD(\lambda_0)^{-1}[D(\lambda_0)A^*(\lambda_0) - C(\lambda_0)B^*(\lambda_0)]B^*(\lambda_0)^{-1} \\
&= vD(\lambda_0)^{-1}B^*(\lambda_0)^{-1}
\end{aligned}$$

which is absurd because $B(\lambda_0)$ and $D(\lambda_0)$ are both invertible matrices and $v \neq \theta$.

Finally, let's see that $V(\lambda)$ belongs to $\mathscr{V}$ if and only if $\omega(\lambda)$ belongs to $\mathscr{N}$. For this, observe that for $V(\lambda)$ given by (5.3), we have

$$V(\lambda)^*V(\lambda) \leq I$$

$$\Longleftrightarrow$$

$$\begin{aligned}
&\big\{[C(\bar\lambda) + D(\bar\lambda)\omega^*(\bar\lambda)] + i[A(\bar\lambda) + B(\bar\lambda)\omega^*(\bar\lambda)]\big\} \\
&\cdot \big\{[C(\bar\lambda) + D(\bar\lambda)\omega^*(\bar\lambda)] - i[A(\bar\lambda) + B(\bar\lambda)\omega^*(\bar\lambda)]\big\}^{-1} \\
&\cdot \big\{[C^*(\lambda) + \omega(\lambda)D^*(\lambda)] + i[A^*(\lambda) + \omega(\lambda)B^*(\lambda)]\big\}^{-1} \\
&\cdot \big\{[C^*(\lambda) + \omega(\lambda)D^*(\lambda)] - i[A^*(\lambda) + \omega(\lambda)B^*(\lambda)]\big\} \leq I
\end{aligned}$$

$$\Longleftrightarrow$$

$$\begin{aligned}
&\big\{[C(\bar\lambda) + D(\bar\lambda)\omega^*(\bar\lambda)] - i[A(\bar\lambda) + B(\bar\lambda)\omega^*(\bar\lambda)]\big\}^{-1} \\
&\cdot \big\{[C^*(\lambda) + \omega(\lambda)D^*(\lambda)] + i[A^*(\lambda) + \omega(\lambda)B^*(\lambda)]\big\}^{-1} \\
\leq &\big\{[C(\bar\lambda) + D(\bar\lambda)\omega^*(\bar\lambda)] + i[A(\bar\lambda) + B(\bar\lambda)\omega^*(\bar\lambda)]\big\}^{-1} \\
&\cdot \big\{[C^*(\lambda) + \omega(\lambda)D^*(\lambda)] - i[A^*(\lambda) + \omega(\lambda)B^*(\lambda)]\big\}^{-1}
\end{aligned}$$

$$\Longleftrightarrow$$

$$\begin{aligned}&\left\{[C^*(\lambda)+\omega(\lambda)D^*(\lambda)]+i[A^*(\lambda)+\omega(\lambda)B^*(\lambda)]\right\}\\&\cdot\left\{[C(\bar\lambda)+D(\bar\lambda)\omega^*(\bar\lambda)]-i[A(\bar\lambda)+B(\bar\lambda)\omega^*(\bar\lambda)]\right\}\\&\geq\left\{[C^*(\lambda)+\omega(\lambda)D^*(\lambda)]-i[A^*(\lambda)+\omega(\lambda)B^*(\lambda)]\right\}\\&\cdot\left\{[C(\bar\lambda)+D(\bar\lambda)\omega^*(\bar\lambda)]+i[A(\bar\lambda)+B(\bar\lambda)\omega^*(\bar\lambda)]\right\}\\&\Longleftrightarrow\\&2i\left\{[A^*(\lambda)+\omega(\lambda)B^*(\lambda)][C(\bar\lambda)+D(\bar\lambda)\omega^*(\bar\lambda)]\right.\\&\quad\left.-[C^*(\lambda)+\omega(\lambda)D^*(\lambda)][A(\bar\lambda)+B(\bar\lambda)\omega^*(\bar\lambda)]\right\}\geq\theta\\&\Longleftrightarrow\\&\operatorname{Im}\left\{[C^*(\lambda)+\omega(\lambda)D^*(\lambda)][A(\bar\lambda)+B(\bar\lambda)\omega^*(\bar\lambda)]\right\}\geq\theta,\end{aligned}$$

which is precisely the condition for $\omega(\lambda)$ to belong to $\mathsf{B}_\infty(\lambda)$.

The above calculations show that ν is a N-extremal matrix of measures if and only if its corresponding $V(\lambda)$ verifies $V(\lambda)^*V(\lambda)=I$. But a holomorphic matrix function $V(\lambda)$ on H such that $V(\lambda)^*V(\lambda)=I$ reduces to a constant unitary matrix. To see this, observe that for any unitary vector v in C^N, the vector function

$$F(\lambda)=vV(\lambda)^*=(f_1(\lambda),\dots,f_N(\lambda))$$

satisfies $F(\lambda)F(\lambda)^*=vV(\lambda)^*V(\lambda)v^*=1$, so if we take any point λ_0 in H and define

$$G(\lambda)=\overline{f_1(\lambda_0)}f_1(\lambda)+\ldots+\overline{f_N(\lambda_0)}f_N(\lambda),$$

$G(\lambda)$ is holomorphic in H, and Cauchy-Schwarz inequality yields that for any λ in H, $|G(\lambda)|\leq 1$. We have $G(\lambda_0)=1$, so the maximum modulus theorem gives that $G(\lambda)=1$, for every λ in H. Now, Cauchy-Schwarz inequality gives that $f_i(\lambda)=f_i(\lambda_0)$, for every $1\leq i\leq N$ and for every λ in H. Thus we have that for any vector v on C^N, $vV(\lambda)^*$ is constant, which immediately gives that $V(\lambda)$ is constant and equal to a unitary matrix U.

REFERENCES

1. Akhiezer, N. I., *The classical moment problem and some related questions in analysis*, english translation, Oliver and Boyd, Edinburgh, 1965.
2. Akhiezer, N. I. and Glazman, I. M., *Theory of linear operators in Hilbert Space, english translation, Vol. II*, Frederick Ungar Publishing Co., New York, 1966.
3. Aptekarev, A. I. and Nikishin, E. M., *The scattering problem for a discrete Sturm-Liouville operator*, Math. USSR-Sb. 49 (1984), 325–355.
4. Berg, C., *Moment problems and orthogonal polynomials*, Lecture Notes, Department of Mathematics, University of Copenhagen, 1994.

5. Bauer, H., *Probability theory and elements of measure theory*, Academic Press, New York, London, 1978.
6. Berezanskii, Ju. M., *Expansions in eigenfunctions of selfadjoint operators*, Translations of Mathematical Monographs, American Mathematical Society, Providence, Rhode Island, 1968.
7. Duran, A. J. and Lopez-Rodriguez, P., *Orthogonal matrix polynomials: zeros and Blumenthal's Theorem*, J. Approx. Theory 84 (1996), 96–118.
8. Duran, A. J. and Lopez-Rodriguez, P., *The L^p space of a positive definite matrix of measures and density of matrix polynomials in L^1*, J. Approx. Theory 90 (1997).
9. Duran, A. J. and Lupez-Rodriguez, P., *Density questions for the truncated matrix moment problem*, Canad. J. Math. 49 (1997), 708–721.
10. Krein, M., *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).
11. Krein, M., *Fundamental aspects of the representation theory of Hermitian operators with deficiency index* (m, m), Ukrain. Mat. Zh. 1 (1949), 3–66; (English translation in Amer. Math. Soc. Transl. (2) 97 (1970), 75–143).
12. Krein, M. and Krasnoselskii, M., *Fundamental theorems of the extension of Hermitian operators and some applications to the theory of orthogonal polynomials and the moment problem*, Uspehi Mat. Nauk 2 (1947), nr. 3 (19), 60–106
13. Lopez-Rodriguez, P., *Riesz's theorem for orthogonal matrix polynomials*, Constr. Approx. 15 (1) (1999), 135–151.
14. Nevanlinna, R., *Asymptotische Entwicklungen beschränkter Funktionen und das Stieltjessche Momentenproblem*, Ann. Acad. Sci. Fenn. A 18 5 (1922).
15. Nagel, H., *Über die quadrierbaren Hermiteschen Matrizen entstehenden Operatoren*, Math. Ann. 112 (1936), 247–285.
16. Riesz, M., *Sur le problème des moments et le théorème de Parseval correspondant*, Acta Litt. ac Sci. (Szeged) 1 (1922), 209–225.
17. Rosenberg, M., *The square-integrability of matrix-valued functions with respect to a non-negative hermitian measure*, Duke Math. J. 31 (1964), 291–298.
18. Zhani, D., *Problème des moments matriciels sur la droite: construction d'une famille de solutions et questions d'unicité*, Publications du Departement de Mathématiques, 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: plopez@cica.es