MARGARITA MATHEMATICA EN MEMORIA DE
JOSÉ JAVIER (CHICHO) GUADALUPE HERNÁNDEZ
(Luis Español y Juan L. Varona, editores),
Servicio de Publicaciones, Universidad de La Rioja,
Logroño, Spain, 2001.

THE MATRIX MOMENT PROBLEM

ANTONIO J. DURÁN AND PEDRO LÓPEZ-RODRÍGUEZ

A la memoria de nuestro amigo Chicho

Abstract. We expose the recent extensions to the matrix case of classical results in the theory of the moment problem: the theorem of Riesz, the parametrization of Nevanlinna and properties of the N-extremal matrices of measures.

1. The classical theory

The purpose of this survey is to show the recent extensions to the matrix case of classical results in the theory of the moment problem. The interest that Chicho always showed for these questions makes it interesting for a publication devoted to his memory.

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.

Given a sequence of numbers $s_0, s_1, s_2, \dots$, Hamburger's theorem from 1920 states that a necessary and sufficient condition for the existence of a positive measure with infinite support having moments $s_0, s_1, s_2, \dots$ is that the sequence $s_0, s_1, s_2, \dots$ is positive definite, or equivalently that all the Hankel matrices

$$H_n = (s_{i+j})_{0\leq i,j\leq n} = \begin{pmatrix} s_0 & s_1 & s_2 & \dots & s_n \\ s_1 & s_2 & s_3 & \dots & s_{n+1} \\ s_2 & s_3 & s_4 & \dots & s_{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_n & s_{n+1} & s_{n+2} & \dots & s_{2n} \end{pmatrix}$$

are positive definite, which is equivalent to $\det H_n > 0$, for $n \geq 0$. In 1894 Stieltjes had already established the corresponding result for measures supported in $[0, \infty)$.

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

2000 *Mathematics Subject Classification.* 42C05, 44A60.

Key words and phrases. Orthogonal matrix polynomials, matrix moment problem, Pick matrix functions, N-extremal matrices of measures.

This work has been partially supported by DGI, project BFM2000-0206-04-C02.

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,

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

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

The polynomials of the second kind $(q_n)_n$ are given by

$$q_n(x) = \int_{\mathbb{R}} \frac{p_n(x) - p_n(t)}{x - t}\, d\nu(t).$$

They also satisfy the three term recurrence relation, taking initial conditions $q_0 = 0$ and $q_1 = \dfrac{1}{a_1}$. These polynomials play an important role in the theory.

The main results in the truncated moment problem are exposed in the following theorem. In the sequel $\mathbb{P}$ denotes de space of polynomials and $\mathbb{P}_n$ denotes de space of polynomials up to degree n. In the matrix case we will use the same notation.

Theorem 1.1 (The truncated case)**.** *Given a positive measure ν in $\mathbb{R}$ with finite moments of any order, for μ in V_{2n-2}, the following statements are equivalent:*

(1) *The measure μ is an extremal point (in the sense of convexity) of the set V_{n-1}.*
(2) *The polynomials up to degree $n-1$ are dense in the space $L^2(\mu)$.*
(3) *There exists a real number a such that the Stieltjes transform of μ is given by*

$$\omega(\lambda) = \int_{\mathbb{R}} \frac{d\mu(x)}{x - \lambda} = -\frac{q_n(\lambda) - aq_{n-1}(\lambda)}{p_n(\lambda) - ap_{n-1}(\lambda)}. \tag{1.1}$$

(4) *For any given non real λ, the value of the Stieltjes transform $I(\mu)$ in the point λ is extremal in the set*

$$I(V_{2n-2})(\lambda) = \left\{ \int_{\mathbb{R}} \frac{d\mu(t)}{t - \lambda} : \mu \in V_{2n-2} \right\}.$$

The set $I(V_{2n-2})(\lambda)$ is the circle defined by the image of the real line through the Möbius linear transformation

$$\omega_n(\lambda, a) = -\frac{q_n(\lambda) - aq_{n-1}(\lambda)}{p_n(\lambda) - ap_{n-1}(\lambda)} \tag{1.2}$$

defined by (1.1), except for the single point $-q_{n-1}(\lambda)/p_{n-1}(\lambda)$, which is just on its border.

(5) *At some point x_0 of its support (and then at every), the measure μ supports the highest possible weight at the point x_0 for a measure in V_{2n-2}, which is given by*

$$\mu(\{x_0\}) = \left(\sum_{k=0}^{n-1} p_k^2(x_0)\right)^{-1}.$$

As we have exposed, when a moves through the real line, the Möbius linear transformation (1.2) describes a circumference except for one point. The circles defined by these circumferences are the so called Hellinger-Nevanlinna circles, we denote them by $B_n(\lambda)$. A single calculation gives that the centers $\alpha_n(\lambda)$ of these circles are given by

$$\alpha_n(\lambda) = \frac{q_{n-1}(\lambda)p_n(\overline{\lambda}) - q_n(\lambda)p_{n-1}(\overline{\lambda})}{p_{n-1}(\lambda)p_n(\overline{\lambda}) - p_n(\lambda)p_{n-1}(\overline{\lambda})}$$

and that their radius $r_n(\lambda)$ are given by

$$r_n(\lambda) = \left(|\lambda - \overline{\lambda}| \sum_{k=0}^{n-1} |p_k(\lambda)|^2\right)^{-1}.$$

Observe that $B_{n+1}(\lambda) \subseteq B_n(\lambda)$ and that their borders have the common point

$$\omega_n(\lambda, 0) = \omega_{n+1}(\lambda, \infty) = -q_n(\lambda)/p_n(\lambda).$$

Algebraic calculations with the polynomials p_n and q_n give another expression for the set $B_n(\lambda)$. It is the set of complex ω satisfying

$$\sum_{k=0}^{n-1} |\omega p_k(\lambda) + q_k(\lambda)|^2 \leq \frac{\omega - \overline{\omega}}{\lambda - \overline{\lambda}}.$$

We put $B_\infty(\lambda)$ for the intersection of all these nested circles $B_n(\lambda)$. The set $B_\infty(\lambda)$ is a closed disc that can degenerate into a single point, and its radius $r_\infty(\lambda)$ is the limit of the radius of $B_n(\lambda)$:

$$r_\infty(\lambda) = \left(|\lambda - \overline{\lambda}| \sum_{k=0}^{\infty} |p_k(\lambda)|^2\right)^{-1}.$$

It is clear that $B_\infty(\lambda)$ is a circle when the sequence $(p_k(\lambda))_k$ belongs to ℓ^2. This occurrence does not depend on the chosen non real λ:

Theorem 1.2 (R. Nevanlinna, 1922). *If $B_\infty(\lambda_0)$ is a non degenerate circle for some non real λ_0, then $B_\infty(\lambda)$ is always a non degenerate circle (see* [N]*). Furthermore, in such case the series $\sum |p_n(\lambda)|^2$ and $\sum |q_n(\lambda)|^2$ converge uniformly on any compact subset of $\mathbb{C}$.*

An essential result in the theory is the following

Theorem 1.3 (R. Nevanlinna, 1922). *Given a positive definite sequence* $(s_n)_n$ *or equivalently a positive measure* ν *and a number* $\lambda \in \mathbb{C} \setminus \mathbb{R}$, *we have*

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

The measures μ for which $I(\mu)(\lambda)$ lies in the circumference of this circle $I(V)(\lambda)$ are called N-extremal (Nevanlinna-extremal). They play an important role in the theory and have interesting properties that we will expose later.

The Theorem of Riesz gives valuable information about these N-extremal measures:

Theorem 1.4 (M. Riesz). *Let* μ *be a positive measure corresponding to an indeterminate 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) $\mathbb{P}$ *is dense in* $L^2(\mu)$, *equivalently* $(p_n(t))_{n=0}^\infty$ *is an orthonormal basis for the Hilbert space* $L^2(\mu)$.

For a proof see [A] or [R].

In the indeterminate case the series $\sum |p_n(\lambda)|^2$, $\sum |q_n(\lambda)|^2$ converge uniformly on compact subsets of $\mathbb{C}$, which makes it possible to define four important entire functions on $\mathbb{C}$ by

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

These functions depend only on the moment sequence $(s_n)_{n\geq 0}$ of ν, 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 $\mathcal{P}$ of Pick functions, which are holomorphic functions in the upper half-plane $\mathbb{H}$ with non-negative imaginary part. Pick functions are also called Herglotz or Nevanlinna functions.

Theorem 1.5 (R. Nevanlinna, 1922). *There exists a homeomorphism* $\varphi \to \nu_\varphi$ *of* $\mathcal{P} \cup \{\infty\}$ *onto* V *given by*

$$\int_{\mathbb{R}} \frac{d\nu_\varphi(t)}{t-\lambda} = -\frac{a(\lambda)\varphi(\lambda) - c(\lambda)}{b(\lambda)\varphi(\lambda) - d(\lambda)}, \quad \text{for } \lambda \in \mathbb{C} \setminus \mathbb{R}. \tag{1.3}$$

This means that the Stieltjes transform of any solution $\nu \in V$ is given by (1.3) for a unique Pick function φ or by the point ∞ (see [A] or [N]).

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

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

which are holomorphic functions in $\mathbb{C} \setminus \mathbb{R}$. This is just as good, since $\mu \to I(\mu)$ is a one-to-one mapping from the set $M(\mathbb{R})$ of finite complex measures on $\mathbb{R}$ to the set $\mathcal{H}(\mathbb{C} \setminus \mathbb{R})$ of holomorphic functions in $\mathbb{C} \setminus \mathbb{R}$. 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 $M(\mathbb{R})$ of positive measures on the real line as dual space of $\mathcal{C}_0(\mathbb{R})$ (continuous functions on $\mathbb{R}$ vanishing at infinity).

As we have just exposed, N-extremal measures in V are those for which the set $\mathbb{P}$ of polynomials is dense in its corresponding space $L^2(\mu)$. In the parametrization (1.3) of Nevanlinna, they are the ones whose corresponding Pick functions $\varphi(\lambda)$ are real constants or ∞. These constant real functions are extremal in $\mathcal{P}$ in an obvious sense.

If we define $\mathcal{V}$ to be the set of holomorphic functions $v(\lambda)$ in the upper half-plane $\mathbb{H}$ 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 $\mathcal{P} \cup \{\infty\}$ onto $\mathcal{V}$ bijectively, if we accept that the limit function $\varphi(\lambda) = \infty$ is transformed into $v(\lambda) = -1$. Its inverse is given by

$$\varphi(\lambda) = i[1 - v(\lambda)][1 + v(\lambda)]^{-1}. \tag{1.4}$$

If we make this change in (1.3) we obtain the expression

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

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.

The N-extremal measures have some interesting properties that we state in the following theorem.

Theorem 1.6.

(1) *An N-extremal measure is discrete with mass in countably many points, which are the zeros of a certain entire function of minimal exponential type (see* [A, Th. 2.4.3]*).*

(2) *For every real number t there is one and only one N-extremal measure μ_t having a mass point at t (see* [A, Th. 3.4.1]*).*

(3) *If μ is the N-extremal measure having a mass point at t then for any positive real number a the measure $\mu + a\delta_t$ is N-extremal, and the measure $\mu - \mu(\{t\})\delta_t$ is determinate (see* [A, Th. 3.4]*).*

(4) *The N-extremal measure μ having a mass point at t reaches the maximum mass which can be concentrated in t for any solution of the indeterminate moment problem, i.e.:*

$$\mu_t(\{t\}) = \sup\{\nu(\{t\}) : \nu \in V_\mu\}.$$

Moreover, this maximum is uniquely attained by μ_t and

$$\mu_t(\{t\}) = \frac{1}{\sum_{n=0}^{\infty} |p_n(t)|^2}$$

(see [A, Th. 3.4.1]*).*

2. The matrix extension

A matrix of measures μ of size N in the real line is a matrix of size $N \times N$ whose entries are complex Borel measures:

$$\mu = \begin{pmatrix} \mu_{11} & \dots & \mu_{1N} \\ \vdots & \ddots & \vdots \\ \mu_{N1} & \dots & \mu_{NN} \end{pmatrix}.$$

The matrix of measures μ is said to be positive definite if for any Borel set A in the real line the numerical matrix $\mu(A)$ is positive semidefinite. This implies that μ_{ii} are positive measures and that $\mu_{ij} = \overline{\mu_{ji}}$.

A matrix polynomial $P(t)$ of size N and degree n is a square matrix of size $N \times N$ whose entries are polynomials:

$$P(t) = \begin{pmatrix} p_{11}(t) & \cdots & p_{1N}(t) \\ \vdots & \ddots & \vdots \\ p_{N1}(t) & \cdots & p_{NN}(t) \end{pmatrix},$$

or equivalently, a polynomial of the form

$$P(t) = A_n t^n + A_{n-1} t^{n-1} + \cdots + A_0,$$

where $A_0, \dots, A_n$ are numerical matrices of size $N \times N$.

A number a is a zero of the polynomial $P(t)$ if it is a zero of the scalar polynomial $\det P(t)$, that is, if the matrix $P(a)$ is singular, or equivalently, if 0 is an eigenvalue of $P(a)$. The multiplicity of a as a zero of $P(t)$ is the multiplicity of a as a zero of $\det P(t)$.

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

$$tP_n(t) = A_{n+1}P_{n+1}(t) + B_n P_n(t) + A_n^* P_{n-1}(t), \quad n \geq 0, \tag{2.1}$$

(A_n and B_n being $N \times N$ matrices such that $\det(A_n) \neq 0$ and $B_n^* = B_n$), with initial condition $P_{-1}(t) = \theta$. (Here and in the rest of this paper, we write θ for the null matrix, the dimension of which can be determined from the context. For instance, here θ is the $N \times N$ null matrix.)

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

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

which also satisfy the recurrence relation (2.1), 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 indices of deficiency of the operator J defined by the $(2N+1)$-banded 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 (2.1). 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. When the matrix of measures is indeterminate but not completely indeterminate we call it just indeterminate.

For μ a positive definite matrix of measures, the space $L^2(\mu)$ is defined as the set of $N \times N$ matrix functions $f : \mathbb{R} \to M_{N\times N}(\mathbb{C})$ 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 $L^2(\mu)$ 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 L^p spaces associated to μ, $1 \leq p < \infty$, see [DL2].)

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

The sequence of moments S_n associated to the matrix of measures μ is given by

$$S_n = \int_{\mathbb{R}} t^n \, d\mu.$$

As in the scalar case, for a given positive definite matrix of measures ν we denote by V the set of positive definite matrices of measures μ on $\mathbb{R}$ having the same moments as μ, that is, the set of solutions to the Hamburger moment problem defined by μ. By V_n we denote the set of positive definite matrices of measures on $\mathbb{R}$ such that $\int_{\mathbb{R}} t^k\, d\mu(t) = S_k$, $0 \le k \le n$, that is, the set of solutions to the truncated moment problem defined by μ.

We first give two examples showing that the determinacy or indeterminacy of matrices of measures is essentially different to the scalar case. Both examples have the form

$$W_\mu = \begin{pmatrix} \mu & \mu(\{x\})\delta_x - \mu(\{y\})\delta_y \\ \mu(\{x\})\delta_x - \mu(\{y\})\delta_y & \mu \end{pmatrix}, \tag{2.2}$$

where $x, y \in \operatorname{supp}\mu$ and μ is an indeterminate measure, but in one case the matrix of measures W_μ is determinate and in the other it is completely indeterminate.

(1) If μ is an N-extremal measure then W_μ given by (2.2) is determinate. This is indeed equivalent to show that

$$\frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} W_\mu \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

is determinate, but we have

$$\frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} W_\mu \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} \mu - \mu(\{x\})\delta_x + \mu(\{y\})\delta_y & 0 \\ 0 & \mu + \mu(\{x\})\delta_x - \mu(\{y\})\delta_y \end{pmatrix}$$

and from Theorem 1.6 (3) the measures

$$(\mu + \mu(\{y\})\delta_y) - \mu(\{x\})\delta_x \quad \text{and} \quad (\mu + \mu(\{x\})\delta_x) - \mu(\{y\})\delta_y$$

are both determinate and then it is not difficult to see that the matrix of measures

$$\begin{pmatrix} \mu - \mu(\{x\})\delta_x + \mu(\{y\})\delta_y & 0 \\ 0 & \mu + \mu(\{x\})\delta_x - \mu(\{y\})\delta_y \end{pmatrix}$$

is also determinate.

(2) If we put $\mu = \chi_{[0,1]}(t)dt + \nu$, with ν being N-extremal, then W_μ given by (2.2) is completely indeterminate. Again, this is equivalent to see that

$$\frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} W_\mu \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

is completely indeterminate, but

$$\frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} W_\mu \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \chi_{[0,1]}(t)dt + \nu - \nu(\{x\})\delta_x + \nu(\{y\})\delta_y & 0 \\ 0 & \chi_{[0,1]}(t)dt + \nu + \nu(\{x\})\delta_x - \nu(\{y\})\delta_y \end{pmatrix},$$

and it is not difficult to see that

$$\chi_{[0,1]}(t)dt + \nu - \nu(\{x\})\delta_x + \nu(\{y\})\delta_y \quad \text{and} \quad \chi_{[0,1]}(t)dt + \nu + \nu(\{x\})\delta_x - \nu(\{y\})\delta_y$$

are both indeterminate.

We begin with the generalization of some of the points in Theorem 1.1:

Theorem 2.1. *For a positive definite matrix of measures μ in V_{2n-2} the following statements are equivalent:*

(1) *μ is an extremal measure of the set V_{n-1} (in the sense of convexity).*
(2) *P_{n-1} is dense in the space $L^2(\mu)$.*
(3) *There exists an $N \times N$ numerical 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, $i = 1, \dots, q$ 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, $i = 1, \dots, q$.

See [DL3] for the proof.

Property (5) of Theorem 1.1 is not true in the matrix case. Indeed, the matrices in V_{2n-2} extremal in V_{n-1} may not support the maximum mass in V_{2n-2}. This happens exactly when the mass is a non-singular matrix, in which case the polynomial $P_n(\lambda) - AP_{n-1}(\lambda)$ has a zero of maximum multiplicity:

Theorem 2.2. *If μ is a matrix of measures extremal in V_{n-1}, put*

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

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

(1) *μ reaches in x_{i_0} the maximum mass permitted in x_{i_0} for a matrix of measures in V_{2n-2}, more concretely,*

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

(2) *G_{i_0} is non-singular.*
(3) *$P_n(\lambda) - AP_{n-1}(\lambda)$ has a zero of maximum multiplicity (N) in x_{i_0}.*
(4) *$P_{n-1}(x_{i_0})$ is non-singular and $A = P_n(x_{i_0}) P_{n-1}^{-1}(x_{i_0})$.*

See [DL3] for the proof.

Property (4) of Theorem 1.1 is still an open problem in the matrix case, we will give more details later.

The generalization of the Theorem of Riesz to the matrix case also presents important difficulties. We have proved this theorem in the matrix case in the completely indeterminate case, or equivalently, when the indices of deficiency of the operator J are both N. In this 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]).

This convergence permits to define the following four analytic matrix functions,

$$A(\lambda) = \lambda \sum_{k=0}^{\infty} Q_k^*(0) Q_k(\lambda), \qquad B(\lambda) = -I + \lambda \sum_{k=0}^{\infty} Q_k^*(0) P_k(\lambda),$$
$$C(\lambda) = I + \lambda \sum_{k=0}^{\infty} P_k^*(0) Q_k(\lambda), \quad D(\lambda) = \lambda \sum_{k=0}^{\infty} P_k^*(0) P_k(\lambda),$$

which together with their partial sums

$$A_n(\lambda) = \lambda \sum_{k=0}^{n-1} Q_k^*(0) Q_k(\lambda), \qquad B_n(\lambda) = -I + \lambda \sum_{k=0}^{n-1} Q_k^*(0) P_k(\lambda),$$
$$C_n(\lambda) = I + \lambda \sum_{k=0}^{n-1} P_k^*(0) Q_k(\lambda), \quad D_n(\lambda) = \lambda \sum_{k=0}^{n-1} P_k^*(0) P_k(\lambda),$$

play an important role in the theory.

We also put

$$R_n(\lambda) = \left(\sum_{k=0}^{n} P_k^*(\overline{\lambda}) P_k(\lambda) \right)^{-1}. \tag{2.3}$$

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

$$[\omega + \alpha_n(\lambda)] R_{n-1}(\overline{\lambda})^{-1} [\omega + \alpha_n(\lambda)]^* \leq |\lambda - \overline{\lambda}|^{-2} R_{n-1}(\lambda), \tag{2.4}$$

where

$$\alpha_n(\lambda) = \left(\frac{I}{2i \operatorname{Im} \lambda} + \sum_{k=0}^{n-1} Q_k^*(\lambda) P_k(\overline{\lambda}) \right) \left(\sum_{k=0}^{n-1} P_k^*(\lambda) P_k(\overline{\lambda}) \right)^{-1}.$$

The sets $B_n(\lambda)$ are the matrix equivalents to the Hellinger-Nevanlinna circles. Some algebraic calculations with the polynomials P_n and Q_n show that $B_n(\lambda)$ is also the set of $N \times N$ complex matrices ω satisfying the matrix inequality

$$\sum_{k=0}^{n-1} (Q_k^*(\lambda) + \omega P_k^*(\lambda))(Q_k(\overline{\lambda}) + P_k(\overline{\lambda})\omega^*) \leq \frac{\omega - \omega^*}{\lambda - \overline{\lambda}}. \tag{2.5}$$

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

$$[\omega + C(\lambda)] R(\overline{\lambda})^{-1} [\omega + C(\lambda)]^* \leq |\lambda - \overline{\lambda}|^{-2} R(\lambda), \tag{2.6}$$

where

$$C(\lambda) = \left(\frac{I}{2i \operatorname{Im} \lambda} + \sum_{k=0}^{\infty} Q_k^*(\lambda) P_k(\overline{\lambda}) \right) \left(\sum_{k=0}^{\infty} P_k^*(\lambda) P_k(\overline{\lambda}) \right)^{-1}.$$

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

$$\sum_{k=0}^{\infty}(Q_k^*(\lambda)+\omega P_k^*(\lambda))(Q_k(\overline{\lambda})+P_k(\overline{\lambda})\omega^*) \leq \frac{\omega-\omega^*}{\lambda-\overline{\lambda}}. \tag{2.7}$$

Looking at (2.4) and (2.6) 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 ($\operatorname{Ext} B_n(\lambda)$ and $\operatorname{Ext} B_\infty(\lambda)$) are those for which equality is attained in (2.4) and (2.5) or (2.6) and (2.7) respectively.

Other calculations with algebraic formulas involving P_n and Q_n show that an equivalent condition for ω to be an extremal point of $B_n(\lambda)$ is that the matrix

$$(\omega P_n^*(\lambda)+Q_n^*(\lambda))A_n^*(P_{n-1}(\overline{\lambda})\omega^*+Q_{n-1}(\overline{\lambda})) \tag{2.8}$$

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 $B_\infty(\lambda)$ ($\operatorname{Int} B_n(\lambda)$ and $\operatorname{Int} B_\infty(\lambda)$) if a strict inequality is attained in (2.4) and (2.5) or (2.6) and (2.7) respectively.

The Theorem of Nevanlinna has the same formulation in the matrix case:

Theorem 2.3. *Let V denote the set of solutions to a completely indeterminate matrix moment problem defined by a matrix of measures ν and let $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Then we have*

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

See [L1] for the proof.

The statement is exactly the same as in the scalar case, however the proof given in the scalar case fails completely in the matrix case, because in the scalar case a complete description of the Stieltjes transforms of the solutions of the truncated problem is known in terms of the circles of Hellinger-Nevanlinna. However in the matrix case this description is far more complicated. The key to prove this theorem in the matrix case is the inclusions

$$\operatorname{Int} B_n(\lambda) \subseteq I(V_{2n-2})(\lambda) \subseteq B_n(\lambda). \tag{2.9}$$

The proofs of these inclusions present much more difficulties than in the scalar case. 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 $\operatorname{Int} B_n(\lambda) \subseteq I(V_{2n-2})(\lambda)$. This inclusion is not at all so immediate in the matrix case, and it is done in several steps:

Step 1.

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

where

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

and $\operatorname{co}(\Gamma_n(\lambda))$ *stands for the convex hull of* $\Gamma_n(\lambda)$.

Step 2.

$$\Gamma_n(\lambda) \subseteq \operatorname{Ext} B_n(\lambda),$$

where $\operatorname{Ext} B_n(\lambda)$ *stands for the set of extremal points of* $B_n(\lambda)$.

Step 3.

$$I(V_{2n-2})(\lambda) \cap \operatorname{Ext} B_n(\lambda) = \Gamma_n(\lambda).$$

Step 4. *If* $\omega \in \operatorname{Ext} B_n(\lambda)$, *the following two conditions are equivalent:*

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

Step 5. *The set* $\Gamma_n(\lambda)$ *is dense in* $\operatorname{Ext} B_n(\lambda)$.

Step 6. *The following inclusion holds:*

$$\operatorname{Int} B_n(\lambda) \subseteq \operatorname{co}(\Gamma_n(\lambda)).$$

We also have the generalization to the matrix case of the Theorem of Riesz:

Theorem 2.4 (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* $L^2(\mu)$, *equivalently* $(P_n(t))_{n=0}^\infty$ *is an orthonormal basis for the Hilbert space* $L^2(\mu)$.

See [L1] for the proof.

For the Nevanlinna parametrization, in the matrix case the parameter space is the space $\mathcal{V}$ of holomorphic matrix functions $V(\lambda)$ in the upper half-plane $\mathbb{H}$ such that $V(\lambda)^*V(\lambda) \leq I$. The theorem reads as follows:

Theorem 2.5. *There exists a homeomorphism between the set V and the set $\mathcal{V}$ given by*

$$\int_{\mathbb{R}} \frac{d\nu(t)}{t-\lambda} = -\left\{C^*(\lambda)[I+V(\lambda)] - iA^*(\lambda)[I-V(\lambda)]\right\} \times \left\{D^*(\lambda)[I+V(\lambda)] - iB^*(\lambda)[I-V(\lambda)]\right\}^{-1}. \tag{2.10}$$

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

See [L2] for the proof.

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 $\mathbb{H}$ such that for any z in $\mathbb{H}$ 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 $\mathbb{H}$, then we can define

$$\Phi(\lambda) = i[I-V(\lambda)][I+V(\lambda)]^{-1}, \tag{2.11}$$

which is a Pick matrix function:

$$\begin{aligned} \frac{\Phi(\lambda)-\Phi(\lambda)^*}{2i} &= \frac{1}{2i}\left\{i[I-V(\lambda)][I+V(\lambda)]^{-1} + i[I+V(\lambda)^*]^{-1}[I-V(\lambda)^*]\right\} \\ &= \frac{1}{2}[I+V(\lambda)^*]^{-1}\left\{[I+V(\lambda)^*][I-V(\lambda)] + [I-V(\lambda)^*][I+V(\lambda)]\right\}[I+V(\lambda)]^{-1} \\ &= [I+V(\lambda)^*]^{-1}\left\{I-V(\lambda)^*V(\lambda)\right\}[I+V(\lambda)]^{-1} \geq 0 \end{aligned} \tag{2.12}$$

because $V(\lambda)$ belongs to $\mathcal{V}$. In this case (2.10) becomes

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

which is the matrix version of (1.3).

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

$$\int_{\mathbb{R}} \frac{d\nu(t)}{t-\lambda} = -\left\{C^*(\lambda)[I+U] - iA^*(\lambda)[I-U]\right\}\left\{D^*(\lambda)[I+U] - iB^*(\lambda)[I-U]\right\}^{-1} \tag{2.13}$$

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

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

is hermitian, and (2.13) reduces to

$$\int_{\mathbb{R}} \frac{d\nu(t)}{t-\lambda} = -\left\{C^*(\lambda) - A^*(\lambda)H\right\}\left\{D^*(\lambda) - B^*(\lambda)H\right\}^{-1}, \tag{2.14}$$

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

The N-extremal matrices of measures also satisfy a number of interesting properties although the matrix structure creates important divergences; for instance: for every real number t, and for any natural number m, $0 \leq m < N$, there are infinitely many N-extremal measures having a mass point at t of rank m, but only one having a mass point at t of rank N.

To present the results in full, we need some definitions and previous results. For t_0 a real number, the matrix $D^*(t_0) + iB^*(t_0)$ is non-singular and the matrix

$$U_{t_0} = -(D^*(t_0) + iB^*(t_0))^{-1}(D^*(t_0) - iB^*(t_0)) \tag{2.15}$$

is unitary. For any unitary matrix U we write

$$A_U = \{u \in \mathbb{C}^N : Uu^* = U_{t_0}u^*\}.$$

We have the following theorem which characterizes the N-extremal matrices of measures having a mass point at t_0:

Theorem 2.6. *The Nevanlinna parametrization (2.13) establishes a bijective mapping between the sets*

$$\left\{U : U \text{ is a unitary matrix, } \dim\left(A_U\right) = m\right\}$$

and

$$\left\{\nu : \nu \text{ is an } N\text{-extremal matrix of measures with } \operatorname{rank}\left(\nu(\{t_0\})\right) = m\right\}.$$

Moreover:

(1) *If ν is N-extremal, the matrix $\nu(\{t_0\})$ is the inverse of the positive definite matrix $\sum_{k=0}^{\infty} P_k^*(t_0)P_k(t_0)$ on the orthogonal subspace of* $\operatorname{Ker}\left(\nu(\{t_0\})\right)$, *that is, if $u, v \in \operatorname{Ker}^{\perp}\left(\nu(\{t_0\})\right)$ then*

$$\begin{aligned} &u\left(\nu(\{t_0\})\right)\left(\sum_{k=0}^{\infty} P_k^*(t_0)P_k(t_0)\right)v^* \\ &= u\left(\sum_{k=0}^{\infty} P_k^*(t_0)P_k(t_0)\right)\left(\nu(\{t_0\})\right)v^* \\ &= uv^*. \end{aligned} \tag{2.16}$$

(2) *If ν is N-extremal, the matrix $\nu(\{t_0\})$ attains, on the orthogonal subspace of* $\operatorname{Ker}\left(\nu(\{t_0\})\right)$, *the maximum mass which can be concentrated at t_0 for*

any solution of the indeterminate matrix moment problem, that is, if $u \in \operatorname{Ker}^{\perp}\Big(\nu(\{t_0\})\Big)$ *and* μ *has the same matrix moments as those of* ν *then:*

$$u\nu(\{t_0\})u^* \geq u\mu(\{t_0\})u^*.$$

As a consequence, there is only one N*-extremal matrix of measures with a non-singular mass at the point* t_0*: the* N*-extremal matrix of measures associated to the unitary matrix* U_{t_0} *defined by (2.15). In this case the mass at* t_0 *is*

$$\left(\sum_{k=0}^{\infty} P_k^*(t_0)P_k(t_0)\right)^{-1}.$$

See [DL4] for the proof.

We stress the important differences between our Theorem 2.6 and the property (2) of N-extremal measures pointed out above. The property (3) also has a more complicated interpretation in the matrix case which depends on the rank of the mass that the N-extremal solution supports on t:

Theorem 2.7. *If* ν *is an* N*-extremal matrix of measures then the deficiency index of the matrix of measures* $\nu - \nu(\{t\})\delta_t$ *is less than or equal to* $N - \operatorname{rank}\Big(\nu(\{t\})\Big)$.

See [DL4] for the proof.

Finally, the property (4) in Theorem 1.6 has an analogous in the matrix case:

Corollary 2.8. *The* N*-extremal matrix of measures associated to the unitary matrix* U_{t_0} *is the only solution of the indeterminate matrix moment problem having maximum mass at the point* t_0*.*

See [DL4] for the proof.

References

[A] N. I. Akhiezer, *The classical moment problem and some related questions in analysis*, Oliver and Boyd, Edinburgh, 1965. (Translated from the russian by N. Kemmer.)

[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 **17**, American Mathematical Society, Providence, RI, 1968.

[D1] A. J. Durán, A generalization of Favard's theorem for polynomials satisfying a recurrence relation, *J. Approx. Theory* **74** (1993), 83–109.

[D2] A. J. Durán, On orthogonal polynomials with respect to a positive definite matrix of measures, *Canad. J. Math.* **47** (1995), 88–112.

[D3] A. J. Durán, Markov's theorem for orthogonal matrix polynomials, *Canad. J. Math.* **48** (1996), 1180–1195.

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

[DL2] 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, *Canad. J. Math.* **49** (1997), 708–721.

[DL4] A. J. Durán and P. López-Rodríguez, N-extremal matrices of measures for an indeterminate matrix moment problem, *J. Funct. Anal.* **174** (2000), 301–321.

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

[K] M. Krein, Infinite J-matrices and a matrix moment problem (Russian), *Dokl. Akad. Nauk SSSR (N.S.)* **69** (1949), 125–128. (Translation by Walter Van Assche in a personal note.)

[L1] P. López-Rodríguez, Riesz's theorem for orthogonal matrix polynomials, *Constr. Approx.* **15** (1999), 135–151.

[L2] P. López-Rodríguez, The Nevanlinna parametrization for a matrix moment problem, *Math. Scand.*, to appear.

[Na] H. Nagel, Über die quadrierbaren Hermiteschen Matrizen entstehenden Operatoren, *Math. Ann.* **5** (1936), 247–285.

[N] R. Nevanlinna, Asymptotische Entwickelungen beschränkter Funktionen und das Stieltjessche Momentenproblem, *Ann. Acad. Sci. Fenn. Ser. A* **18** (1922), 88–112.

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

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

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

E-mail address: `duran@cica.es, plopez@cica.es`