JOURNAL OF APPROXIMATION THEORY **90**, 97–116 (1997)
ARTICLE NO. AT963064

Algebraic Aspects of Matrix Orthogonality for Vector Polynomials

Vladimir N. Sorokin

Mechanics–Mathematics Department, Moscow State University, Moscow, Russia
E-mail: vnsormm@nw.math.msu.su

and

Jeannette Van Iseghem

UFR de Mathématiques, Université de Lille, 59655 Villeneuve d'Ascq Cedex, France
E-mail: jvaniseg@ano.univ-lille1.fr

Communicated by Walter Van Assche

Received October 20, 1995; accepted in revised form May 31, 1996

An algebraic theory of orthogonality for vector polynomials with respect to a matrix of linear forms is presented including recurrence relations, extension of the Shohat–Favard theorem, of the Christoffel–Darboux formula, and its converse. The connection with orthogonal matrix polynomials is described. © 1997 Academic Press

1. INTRODUCTION

Classical orthogonal polynomials and matrix polynomials orthogonal with respect to some Hermitian positive matrix of measures satisfy several algebraic properties, for example, recurrence relations. Matrix orthogonality has already been studied by several authors. In [6] a part of our study, considering polynomials with matrix coefficients, is made in the case of square Hermitian matrices and a more precise reference to their work is given at the end. The work in [5] is also concerned with this problem but from another point of view; the difference is obvious on the matrix called here A ($xH = AH$ formula (15)) which is a Hessenberg matrix in their framework and which is a banded matrix here. This problem was partly investigated in [11]. This study started in [1], with other aims and notation, linked with previous work [7, 8].

This paper deals with algebraic aspects of matrix orthogonality. In particular we want to show how the structure of orthogonal polynomials with respect to a $p \times q$ matrix of arbitrary linear forms is a canonical one. The

0021-9045/97 $25.00

scalar case is recovered if the matrix is of size 1 ($p=q=1$) in the sense defined in [2], sometimes called formal orthogonality. The vector case, i.e., $q=1$, as defined in [12, 13], is also recovered, and as in this last case, only approximation with "regular degree" is considered (section 2).

Algebraic properties such as recurrence relations, a generalization of the Shohat–Favard theorem, and a Christoffel–Darboux formula, are studied in the case of orthogonality with respect to an abstract matrix of measures known by their moments. As in the scalar case the moments can also be considered as given by the coefficients of power series, here this means the coefficients of a $p \times q$ matrix Θ of ordinary scalar power series with real or complex coefficients. The link with the Hermite-Padé point of view will be given, i.e., the simultaneous approximation of a matrix of functions.

The orthogonality is defined as right-orthogonality of a family of vector polynomials $H^{(n)} = (H_1^{(n)}, ..., H_q^{(n)})^t$ through $\Theta(H^{(n)} x^\nu) = 0$, $\nu = \ldots$. By simple transposition this gives rise to left-orthogonality with respect to Θ^*. If $(H^{(n)})_{n \geqslant 0}$ are the right-orthogonal polynomials and $(E^{(m)})_{m \geqslant 0}$ the left-orthogonal polynomials with respect to Θ, the Shohat–Favard theorem will be presented in Section 7 in terms of the orthogonal family $(H^{(n)})_{n \geqslant 0}$, and in terms of the biorthogonal family $(H^{(n)}, E^{(m)})_{n,m}$.

A kind of Christoffel–Darboux formula will be given in Section 8. As in the scalar case ([3]), a converse result is proved, i.e., this identity characterizes the matrix orthogonality through the recurrence relations.

One question is to define orthogonal polynomials with respect to a matrix of linear forms as matrix polynomials or vector polynomials. The choice which is done here is to define vector polynomials, but some remarks are given at the end concerning this point which show that under some restrictions on the order (see Section 2) of the vector polynomials, these two points of view are equivalent. Moreover vector orthogonal polynomials and square matrix polynomials are then two extreme cases of only one theory. The basic idea would be that a matrix of polynomials is a set of vector polynomials which are orthogonal. Through the recurrence relation satisfied by the orthogonal polynomials of each kind, this idea is made clear in Section 9.

2. THE MATRIX HERMITE–PADÉ PROBLEM

Let p and q be any fixed natural numbers, $\mathbb{F}(z)$ is a $p \times q$ matrix of functions

$$f_{k,j}(z) = \sum_{\nu=0}^{\infty} \frac{f_{k,j}^{\nu}}{z^{\nu+1}}, \qquad k = 1, ..., p, \quad j = 1, ..., q,$$

each function being a formal power series with complex coefficients. Fix two multi-indices $\bar{m}=(m_1, ..., m_q)$ and $\bar{n}=(n_1, ..., n_p)$ of order $|\bar{m}|=m_1+\cdots+m_q$ and $|\bar{n}|=n_1+\cdots+n_p$ which satisfy $|\bar{m}|=|\bar{n}|+1$.

The following problem is now considered as the Hermite–Padé problem at infinity. We look for scalar polynomials $H_1, ..., H_q$, not simultaneously zero, of degree not greater than respectively $m_1-1, ..., m_q-1$ and such that for some polynomials K_i, $i=1, ..., p$ the following relations are satisfied

$$\begin{cases} R_1=H_1 f_{1,1}+\cdots+H_q f_{1,q}-K_1=O(1/z^{n_1+1}) \\ \quad\vdots \\ R_p=H_1 f_{p,1}+\cdots+H_q f_{p,q}-K_p=O(1/z^{n_p+1}). \end{cases} \tag{1}$$

There always exists non trivial solutions because the problem reduces to solving a linear system of $|\bar{n}|$ equations for $|\bar{m}|=|\bar{n}|+1$ unknowns (the coefficients of the polynomials H_k, $k=1, ..., q$). The polynomials $K_1, ..., K_p$ are defined automatically as the polynomial part of the series $H_1 f_{i,1}+\cdots+H_q f_{i,q}$, $i=1, ..., p$; if some negative degree occurs for a polynomial, then the polynomial is zero. In general, the solution is not unique, even up to the multiplication by constants. The problem can, equivalently, be considered in a neighbourhood of zero. In that case a recursive algorithm for solving (1) can be found in [8] where no regularity assumptions are required.

If $p=1$, the problem is the Hermite–Padé approximation of the first kind, if $q=1$ it is the Hermite–Padé problem of the second kind, also called vector Padé approximation [13].

The problem (1) can be rewritten in matrix form with $H=(H_1, ..., H_q)^t$ (for $i=1, ..., q$, $\deg H_i\leqslant m_i-1$) and $K=(K_1, ..., K_p)^t$

$$R(z)=\mathbb{F}(z)\,H(z)-K(z)=O(1/z^{\bar{n}+1}), \qquad \deg H\leqslant \bar{m}, \quad |\bar{m}|=|\bar{n}|+1 \tag{2}$$

where $\mathbb{F}=(f_{k,j})$, $k=1, ..., p$, $j=1, ..., q$, R, K, H are column matrices of power series or polynomials, of size p for R and K, q for H, and the $O(1/z^{\bar{n}+1})$ being understood as in (1).

3. REGULAR MULTI–INDICES AND WEAKLY PERFECT SYSTEMS

In the following we will restrict ourselves to regular multi-indices $\bar{n}$ ([10]).

Definition 1. A multi-index $\bar{k}=(k_1, ..., k_d)\in\mathbb{Z}^d$ is regular if

$$k_1\geqslant k_2\geqslant\cdots\geqslant k_d\geqslant k_1-1$$

Regular multi-indices are uniquely defined by their order, if n is the order and $n=\nu d+k$, $k<d$, then $\bar{n}=(n_i)_{i=1,\ldots,d}$, $n_1=\cdots=n_k=\nu+1$, $n_{k+1}=\cdots=n_d=\nu$.

We now consider the following "canonical" basis for the vector space of vector polynomials of size q

$$h_0=\begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix}, \ldots, h_{q-1}=\begin{pmatrix}0\\0\\0\\\vdots\\1\end{pmatrix}, h_q=\begin{pmatrix}x\\0\\0\\\vdots\\0\end{pmatrix}, \ldots, h_{2q}=\begin{pmatrix}x^2\\\cdots\\0\\\vdots\\0\end{pmatrix}, \ldots.$$

For $n=\nu q+k, k<q, n\geqslant 0$, h_n has 0 for its components except for the component $(k+1)$ which is x^ν.

Any linear combination of $h_0, \ldots, h_n$ is said to be of order less than or equal to n. A vector polynomial, $c_0h_0+\cdots+c_nh_n$, satisfying $c_n\neq 0$ will be called of maximum order; this is the analog of having maximum degree in the scalar case.

So for all $n\geqslant 0$, n defines a regular multi-index $(n_1, \ldots, n_p)$, and $H^{(n)}$ will denote a vector polynomial of order n which satisfies

$$\begin{cases} H^{(n)}=c_0h_0+\cdots+c_nh_n, \qquad c_n\neq 0\\ R^{(n)}=\mathbb{F}H^{(n)}-K^{(n)}=O(1/z^{\bar{n}+1})=O(1/z^{n_1+1}), \ldots, O(1/z^{n_p+1}))^t \\ \text{The order of approximation is maximum.}\end{cases} \tag{3}$$

The assumption that the approximation is maximum means that $R^{(n)}=O(1/z^{\bar{n}+1})$, and if $n'=n+1$, then $R^{(n)}\neq O(1/z^{\bar{n}'+1})$. This was already the sense given in the case of vector approximation ([12]). The assumption that the order of approximation is maximum limits the study to the classical non-degenerate or "normal" case.

To a vector polynomial $H=(H_1, \ldots, H_q)$ one can associate, as in [1] or in [6], the scalar polynomial

$$p(x)=\sum_{i=1}^{q} x^{i-1}H_i(x^q).$$

The degree of p and the order of H play the same role. Conversely if p is known, $x^{i-1}H_i(x^q)$ is formed by the terms of p with powers equal to $i-1$ (mod q). Recurrence relations for the H's and the p's are the same.

Let us define the matrix moments $f^\nu=(f^\nu_{k,j})_{k=1,\ldots,p,j=1,\ldots,q}$, then

$$\mathbb{F}(z)=\sum_{\nu=0}^{\infty}\frac{f^\nu}{z^{\nu+1}}.$$

The infinite generalized Hankel matrix can be considered in block form, or in standard scalar form (which defines the complex constants $h_{i,j}$, $i \geqslant 0$, $j \geqslant 0$)

$$\mathscr{H} = \begin{pmatrix} f^0 & f^1 & f^2 & \cdots \\ f^1 & f^2 & f_3 & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix} = \begin{pmatrix} h_{0,0} & h_{0,1} & h_{0,2} & \cdots \\ h_{1,0} & h_{1,1} & h_{1,2} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}.$$

The principal minors of size $n \times n$ of the scalar writing on the right hand side are called the generalized Hankel determinants and are denoted by $\mathscr{H}_n$; by convention $\mathscr{H}_0 = 1$. Then, for a regular index $(n_1, ..., n_p)$, looking for a solution $H^{(n)}$ of maximum order, expanded in the basis (h_k), leads to the following linear system with respect to the constant c_i, $i = 0, ..., n$, $c_n \neq 0$

$$\begin{cases} h_{0,0} c_0 + \cdots + h_{0,n-1} c_{n-1} + h_{0,n} c_n = 0 \\ \quad \vdots \qquad\qquad\qquad\qquad\qquad \vdots \\ h_{n-1,0} c_0 + \cdots + h_{n-1,n-1} c_{n-1} + h_{n-1,n} c_n = 0 \end{cases} \tag{4}$$

Definition 2. *An index $n \geqslant 0$ is said to be normal if $\mathscr{H}_n \neq 0$.*

Lemma 1. *If n is normal, then* (4) *has a unique solution, up to the multiplication by a constant a_n, where the leading coefficient is non zero. We get as solution of* (4), *a vector polynomial*

$$H^{(n)} = \frac{a_n}{\mathscr{H}_n} \begin{vmatrix} h_{0,0} & h_{0,1} & \cdots & h_{0,n} \\ h_{1,0} & h_{1,1} & \cdots & h_{1,n} \\ \vdots & \vdots & \vdots & \\ h_{n-1,0} & h_{n-1,1} & \cdots & h_{n-1,n} \\ h_0 & h_1 & \cdots & h_n \end{vmatrix}, \tag{5}$$

where the last row of the determinant is composed of vectors.

The proof is a direct consequence of Cramer's formulae.

Definition 3. The matrix $\mathbb{F}$ is called weakly perfect if all n are normal.

In the following, we assume weak perfectness of $\mathbb{F}$ or, in linear algebra terminology, strong regularity of the matrix $\mathscr{H}$.

4. BIORTHOGONALITY WITH RESPECT TO A BILINEAR FORM

For each k and l, the linear functional $\Theta_{k,l}$ is defined on the space of polynomials $\mathbb{C}[x]$ by

$$\Theta_{k,l}(x^\nu) = f^{\nu}_{k,l}, \qquad \nu \geqslant 0.$$

If $H^{(n)}$ is the vector polynomial $(H_1^{(n)}, ..., H_q^{(n)})^t$ then relations (1) or (5) imply

$$\begin{cases} \Theta_{1,1}(H_1^{(n)}(x)x^\nu) + \cdots + \Theta_{1,q}(H_q^{(n)}(x)x^\nu) = 0, & \nu = 0, ..., n_1 - 1 \\ \cdots & \\ \Theta_{p,1}(H_1^{(n)}(x)x^\nu) + \cdots + \Theta_{p,q}(H_q^{(n)}(x)x^\nu) = 0, & \nu = 0, ..., p_1 - 1 \end{cases} \tag{6}$$

These relations can be put in matrix form. If $\mathbb{C}^q[x]$ (resp. $\mathbb{C}^p[x]$) is the vector space of vector polynomials of size q (resp. of size p), then we define Θ as the bilinear form from $\mathbb{C}^p[x] \times \mathbb{C}^q[x]$ to $\mathbb{C}$ by: $H = (H_1, ..., H_q)^t$, $E = (E_1, ..., E_p)^t$

$$\langle E, H \rangle_\Theta = E^* \Theta H = \sum_{k=1}^{p} \sum_{j=1}^{q} \Theta_{k,j}(H_j(x)\, \bar{E}_k(x)), \tag{7}$$

so that Θ can also be considered as the matrix of linear forms $(\Theta_{k,j})$, $k = 1, ..., j = 1, ..., q$.

Similar to the $(h_n)_{n \geqslant 0}$, we define the canonical basis $(e_k)_{k \geqslant 0}$ of $\mathbb{C}[x]^p$

$$e_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \cdots \\ 0 \end{pmatrix}, ..., e_{p-1} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, e_p = \begin{pmatrix} x \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, ..., e_{2p} = \begin{pmatrix} x^2 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix}, ...$$

As a remark, if $\tilde{e}_n = \sum_{i=n-p+1}^{n} e_i$, then (6) is the expanded writing of

$$\langle \tilde{e}_m, H^{(n)} \rangle_\Theta = 0, \qquad m = 0, ..., n-1.$$

which is equivalent to

$$\langle e_m, H^{(n)} \rangle_\Theta = 0, \qquad m = 0, ..., n-1. \tag{8}$$

As another remark, $\langle e_m, h_n \rangle_\Theta = h_{m,n}$, and so the formula (5) can be directly deduced from the orthogonality relations (6) or (8). Moreover the fact that the order of approximation is maximum is given by

$$\langle e_n, H^{(n)} \rangle_\Theta = \frac{a_n}{\mathscr{H}_n} \mathscr{H}_{n+1} \neq 0. \tag{9}$$

Θ can also be considered as a linear form from $\mathbb{C}[x]$ to $\mathscr{M}_{p,q}(\Theta(x^\nu) = (\Theta_{k,l}(x^\nu))_{k,l})$, and in that sense it follows formally as in the scalar case ([2]) (in the notation the index x means that Θ acts on the variable x)

$$\mathbb{F}(z) = \Theta_x \left(\frac{1}{z-x} \right).$$

From the beginning the conjugate transposed problem of (2) can also be considered for $\mathbb{F}^*$.

If $\mathbb{F}$ is weakly perfect, then $\mathbb{F}^*$ is also weakly perfect because its Hankel matrix is $\mathscr{H}^*$. Then denote the solution of the transposed problem $S^{(n)} = \mathbb{F}^* E^{(n)} - M^{(n)}$, $n \geqslant 0$, and let α_n be normalizing constants for $E^{(n)}$, then

$$\langle h_m, E^{(n)} \rangle_{\Theta^*} = \langle E^{(n)}, h_m \rangle_\Theta^* = \begin{cases} 0 & \text{if} \quad m < n, \\ \alpha_n \overline{\mathscr{H}}_{n+1} / \overline{\mathscr{H}}_n & \text{if} \quad n = m. \end{cases} \tag{10}$$

From (8), (9), (10), we obtain sequences of biorthogonal polynomials ([4])

Lemma 2. *The sequences $(H^{(n)})_{n \geqslant 0}$ and $(E^{(n)})_{n \geqslant 0}$ are biorthogonal sequences with respect to the bilinear form Θ*

$$\forall n, m \leqslant 0 \qquad \langle E^{(m)}, H^{(n)} \rangle_\Theta = \begin{cases} 0 & \textit{if} \quad n \neq m, \\ a_n \bar{\alpha}_n \dfrac{\mathscr{H}_{n+1}}{\mathscr{H}_n} & \textit{if} \quad n = m. \end{cases} \tag{11}$$

5. RECURRENCE RELATIONS

The previous non zero constants a_n and α_n are the leading coefficient of respectively, $H^{(n)}$ and $E^{(n)}$ (and may be changed in order to normalize the vector polynomials in one way or another)

$$H^{(n)} = a_n h_n + \cdots, \qquad E^{(n)} = \alpha_n e_n + \cdots.$$

THEOREM 1. *There exists a unique set of complex coefficients* $a_n^{(m)}$, $m = -p, ..., q$, $n \geqslant 0$, $n+m \geqslant 0$ *such that the sequence of vector polynomials* $(H^{(k)})_k$ *is the unique solution of the recurrence relation*

$$\begin{aligned} a_n^{(q)} H^{(n+q)} + \cdots + a_n^{(1)} H^{(n+1)} + a_n^{(0)} H^{(n)} \\ + a_n^{(-1)} H^{(n-1)} + \cdots + a_n^{(-p)} H^{(n-p)} = x H^{(n)} \end{aligned} \tag{12}$$

with the initial conditions

$$H^{(-p)} = \cdots = H^{(-1)} = 0$$

$$H^{(j)} = \frac{a_j}{\mathscr{H}_j} \begin{vmatrix} h_{0,0} & \cdots & h_{0,j} \\ & \cdots & \\ h_{0,0} & \cdots & h_{0,j} \\ h_0 & \cdots & h_j \end{vmatrix}, \qquad j = 0, ..., q-1$$

and the coefficients are given by the solution of the system (14). *In particular*

$$\begin{aligned} a_n^{(q)} &= \frac{a_n}{a_{n+q}}, && n \geqslant 0 \\ a_{n+p}^{(-p)} &= \frac{a_{n+p}}{a_n} \frac{\mathscr{H}_{n+p+1}}{\mathscr{H}_{n+p}} \frac{\mathscr{H}_n}{\mathscr{H}_{n+1}}, && n \geqslant 0 \end{aligned} \tag{13}$$

Proof. The required recurrence relation is written in the following form

$$\begin{aligned} a_n^{(q)} H^{(n+q)} = {} & -a_n^{(q-1)} H^{(n+q-1)} - \cdots - a_n^{(1)} H^{(n+1)} + (x - a_n^{(0)})\, H^{(n)} \\ & - a_n^{(-1)} H^{(n-1)} - \cdots - a_n^{(-p)} H^{(n-p)}. \end{aligned}$$

Because $x h_k = h_{k+q}$, the vector polynomial on the right hand side is of order $n+q$, so identifying the leading coefficient gives the expression for $a_n^{(q)}$, $n \geqslant 0$. Now $H^{(n+q)}$ satisfies the orthogonality relations

$$\langle e_k, H^{(n+q)} \rangle_\Theta = 0, \qquad k = 0, ..., n+q-1$$

and the right hand side of the relation is orthogonal to $(e_k, k = 0, ..., n-p-1)$, because $x e_k = e_{k+p}$. Let us develop this orthogonality

$$\begin{cases} a_n^{(-p)} \langle e_{n-p}, H^{(n-p)} \rangle_\Theta = \langle e_{n-p}, x H^{(n)} \rangle_\Theta \\ a_n^{(-p)} \langle e_{n-p+1}, H^{n-p)} \rangle_\Theta + a_n^{(-p+1)} \langle e_{n-p+1}, H^{(n-p+1)} \rangle_\Theta \\ \qquad = \langle e_{n-p+1}, x H^{(n)} \rangle_\Theta \\ \vdots \\ a_n^{(-p)} \langle e_{b+q+1}, H^{(n-p)} \rangle_\Theta + \cdots + a_n^{(q-1)} \langle e_{n+q-1}, H^{(n+q-1)} \rangle_\Theta \\ \qquad = \langle e_{n+q-1}, x H^{(n)} \rangle_\Theta. \end{cases} \tag{14}$$

We get a system of $p+q$ linear equations with $p+q$ unknowns. The matrix of the system is triangular and the diagonal terms are $\langle e_k, H_k \rangle_\Theta$, $k=n-p, ..., b+q-1$, which are nonzero, and thus there exists a unique solution. In particular $a_n^{(-p)}$ can be given from the first equation

$$a_n^{(-p)} = \frac{\langle e_{n-p}, xH^{(n)} \rangle_\Theta}{\langle e_{n-p}, H^{(n-p)} \rangle_\Theta} = \frac{\langle e_n, H^{(n)} \rangle_\Theta}{\langle e_{n-p}, H^{(n-p)} \rangle_\Theta}, \qquad \langle e_n, H^{(n)} \rangle_\Theta = a_n \frac{\mathscr{H}_{n+1}}{\mathscr{H}_n}$$

from which the formulas in (13) are obtained.

The recurrence relation (12) can be written in matrix form as

$$AH = xH, \tag{15}$$

where H is the infinite column vector $(H^{(0)}, H^{(1)}, ...)^t$ (each term being a vector, H could be written as a scalar matrix $(\infty \times q)$) and A a scalar infinite band matrix with $p+q+1$ diagonals

$$A = \begin{pmatrix} a_0^{(0)} & \cdots & \cdots & a_0^{(q)} & 0 & 0 \\ \vdots & a_1^{(0)} & \cdots & \cdots & a_1^{(q)} & 0 \\ \vdots & & \ddots & & & \ddots \\ a_p^{(-p)} & \cdots & \cdots & & & \\ 0 & a_{p+1}^{(-p)} & \cdots & \cdots & & \\ 0 & 0 & \ddots & & & \end{pmatrix}.$$

Equation (15) means that A is the matrix of the operator multiplying the variable x in the set $\mathbb{C}^q[x]$ in the basis $(H^{(n)})_{n \geqslant 0}$.

If the solution of the dual problem $(E^{(n)})_{n \geqslant 0}$ is considered, then in this basis, the system (14) is a diagonal one. If moreover the $E^{(n)}$ are normalized, such that

$$\langle E^{(n)}, H^{(n)} \rangle_\Theta = 1, \qquad n \geqslant 0,$$

then the sequences $(H^{(n)})_{n \geqslant 0}$ and $(E^{(n)})_{n \geqslant 0}$ are biorthonormal, and the $a_n^{(k)}$ are defined by

$$a_n^{(k)} = \langle E^{(n+k)}, xH^{(n)} \rangle_\Theta.$$

The matrix of the same operator in $\mathbb{C}^p[x]$ is, in the basis $(E^{(n)})_{n \geqslant 0}$, given by A^*. This means that the infinite column $E = (E^{(n)})_{n \geqslant 0}$ satisfies the recurrence relations

$$A^*E = xE.$$

In explicit form, the $(E^{(n)})_{n \geqslant 0}$ satisfy (α_n being the leading coefficient of $E^{(n)}$ in basis (e_k))

$$\bar{a}_{n+p}^{(-p)} E^{(n+p)} + \cdots + \bar{a}_{n+1}^{(-1)} E^{(n+1)} + \bar{a}_n^{(0)} E^{(n)} + \bar{a}_{n-1}^{(1)} E^{(n-1)} + \cdots + \bar{a}_{n-q}^{(q)} E^{n-q} = x E^{(n)} \tag{16}$$

with the initial conditions

$$E^{(-q)} = \cdots = E^{(-1)} = 0$$

$$E^{(k)} = \frac{\alpha_k}{\mathscr{H}_k} \begin{vmatrix} h_{0,0} & \cdots & h_{0,k-1} & e_0 \\ & \vdots & & \\ h_{k,0} & \cdots & h_{k,k-1} & e_k \end{vmatrix}, \qquad k = 0, ..., p-1$$

Of course $M^{(n)}$ and $S^{(n)}$ ($S^{(n)} = \mathbb{F}^* E^{(n)} - M^{(n)}$) satisfy the same recurrence relation with different initial conditions.

LEMMA 3. *If the leading coefficients of $H^{(n)}$ and $E^{(n)}$ are respectively a_n, and α_n defined by*

$$a_n = \left(\frac{\mathscr{H}_n}{\mathscr{H}_{n+1}} \right)^{p/p+q}, \qquad \bar{\alpha}_n = \left(\frac{\mathscr{H}_n}{\mathscr{H}_{n+1}} \right)^{q/p+q},$$

then

$$(a_n^{(q)} \cdots a_{n+p-1}^{(q)})^q = (a_{n+p}^{(-p)} \cdots a_{n+p+q-1}^{(-p)})^p$$

(*the q and p exterior to the parentheses are powers, those interior to the parentheses are indices, as before*).

Proof. We have

$$\langle E^{(n)}, H^{(n)} \rangle_\Theta = \left\langle \alpha_n e_n, \frac{a_n}{\mathscr{H}_n} \begin{vmatrix} h_{0,0} & \cdots & h_{0,n} \\ & \vdots & \\ h_{n-1,0} & \cdots & h_{n-1,n} \\ h_0 & \cdots & h_n \end{vmatrix} \right\rangle_\Theta = a_n \bar{\alpha}_n \frac{\mathscr{H}_{n+1}}{\mathscr{H}_n} = 1. \tag{17}$$

Let us define $C_n = (\mathscr{H}_n / \mathscr{H}_{n+1})^{1/(p+q)}$, $a_n = C_n^p$, $\bar{\alpha}_n = C_n^q$, then from (13)

$$(a_n^{(q)} \cdots a_{n+\ -p-1}^{(q)})^q = \left[\frac{\mathscr{H}_n \mathscr{H}_{n+p+q}}{\mathscr{H}_{n+p} \mathscr{H}_{n+q}} \right]^{pq/(p+q)}.$$

Similarly with the $a_k^{(-p)}$, it follows

$$(a_{n+p}^{(-p)} \cdots a_{n+p+q-1}^{(-p)})^p = \left[\frac{\mathscr{H}_{n+q}\mathscr{H}_{n+p}}{\mathscr{H}_n\mathscr{H}_{n+p+q}}\right]^{p^2/(p+q)} \left[\frac{\mathscr{H}_n\mathscr{H}_{n+p+q}}{\mathscr{H}_{n+p}\mathscr{H}_{n+q}}\right]^p$$

$$= \left[\frac{\mathscr{H}_n\mathscr{H}_{n+p+q}}{\mathscr{H}_{n+p}\mathscr{H}_{n+q}}\right]^{pq/(p+q)},$$

which ends the proof.

6. MATRIX HERMITE–PADÉ APPROXIMANTS

From the solution of (2), it is possible to construct the matrices of dimension respectively $p \times q$, $q \times q$, $p \times q$

$$\mathbb{R}_n = (R^{(n)}, R^{(n+1)}, ..., R^{(n+q-1)})$$

$$\mathbb{Q}_n = (H^{(n)}, H^{(n+1)}, ..., H^{(n+q-1)})$$

$$\mathbb{P}_n = (K^{(n)}, K^{(n+1)}, ..., K^{(n+q-1)})$$

with which (2) can be rewritten as

$$\mathbb{R}_n = \mathbb{F}\mathbb{Q}_n - \mathbb{P}_n = \mathbb{O}(1/z^n)$$

where $\mathbb{O}(1/z^n)$ is a matrix of terms $O(1/z^k)$, where the powers of a column or of a row are regular multi-indices (decreasing in the columns and increasing in the rows), and so $\mathbb{P}_n\mathbb{Q}_n^{-1}$ is a "matrix Hermite–Padé approximant" of $\mathbb{F}$. The classical cases are exactly recovered, i.e., the scalar case if $p = q = 1$, the Hermite–Padé approximation of the first or the second kind if $q = 1$ (vector case [12]) or $p = 1$ (which follows from Mahler's connection between the first and second type of approximation [9]).

7. SHOHAT–FAVARD THEOREM

The orthogonality implies a recurrence relation satisfied by the vector polynomials $H^{(n)}$. The Shohat–Favard theorem is the converse of this property. The result for orthogonal polynomials can be put in two forms. The first one involves only the "right" polynomials $H^{(n)}$, and is written taking care only of the zero conditions of orthogonality. The theorem is as follows

THEOREM 2. *If the sequence of vector polynomials* $H^{(n)} = (H_1^{(n)}, ..., H_q^{(n)})^t$ *is defined by*

$$xH^{(n)} = \sum_{k=-p}^{q} a_n^{(k)} H^{(n+k)}, \qquad a_n^{(q)} \neq 0, \quad n \geqslant 0,$$

with the initial conditions $H^{(-k)} = 0$, $k > 0$, $H^{(k)}$, $k = 0, ..., q-1$ *arbitrary, of maximum order* k, *then there exists a* $p \times q$ *matrix* Θ *of linear functionals* $\Theta_{i,j}$, $i = 1, ..., p$, $j = 1, ..., q$ *such that if* n *defines the regular multi-index* $(n_1, ..., n_p)$

$$\sum_{j=1}^{q} \Theta_{i,j}(H_j^{(n)} x^\nu) = 0, \qquad \nu = 0, ..., n_i - 1, \quad i = 1, ..., p.$$

Proof. The first step of the proof is to determine Θ by its moments $\Theta_{i,j}^k$, $i = 1, ..., p$, $j = 1, ..., q$, $k \geqslant 0$ using the orthogonality conditions for $\nu = 0$, which gives

$$\sum_{j=1}^{q} \Theta_{i,j}(H_j^{(n)}) = 0, \qquad i = 1, ..., p, \quad n \geqslant i.$$

The functionals are defined row by row. We define $\Theta_{1,1}(1) = \Theta_{1,1}^0 \neq 0$, then for $n \geqslant 1$ the first equation defines successively $\Theta_{1,2}^0, ..., \Theta_{1,q}^0, \Theta_{1,1}^1, ...$ in terms of $\Theta_{1,1}^0$, so all the moments $\Theta_{1,j}^k$ are defined for $j = 1, ..., q$ and all k positive from the first one $\Theta_{1,1}^0$ (if this one is zero, then all the forms $\Theta_{1,j}$, $j = 1, ..., q$ are identically zero). Similarly the same equation written for the second row, begins at index $n = 2$ and so all the moments $\Theta_{2,j}^k$ are defined from the two first ones $\Theta_{2,1}^0$, $\Theta_{2,2}^0$.

For the row i ($i \geqslant 1$), all the moments are obtained from the first i ones, i.e., if $i = \pi q + \rho$, $0 \leqslant \rho < q$, $\Theta_{i,1}^0, ..., \Theta_{i,q}^0, \Theta_{i,1}^1, ..., \Theta_{i,\rho}^\pi$. The space of solutions Θ is a vector space of dimension $(p(p+1))/2$ as in the vector case of dimension p ([12], where unfortunately $p!$ was written for the sum $1 + \cdots + p$).

The second step of the proof is, Θ being defined by its moments, to verify the orthogonality relations. By definition

$$xH^{(n)} = \sum_{k=-p}^{q} a_n^{(k)} H^{(n+k)},$$

so it follows that

$$\sum_{j=1}^{q} \Theta_{i,j}(xH_j^{(n)}) = \sum_{k=-p}^{q} a_n^{(k)} \left(\sum_{j=1}^{q} \Theta_{i,j}(H_j^{(n+k)}) \right),$$

and all the terms of the sum on the right hand side are zero if $n \geqslant i+p$, which is the right condition to be satisfied. All the orthogonality conditions are similarly satisfied.

A second form of the theorem can be given in a bilinear setting. The orthogonality conditions are completely written

$$\langle e_m, H^{(n)} \rangle_\Theta = 0, \qquad m = 0, ..., n-1, \quad \langle e_n, H^{(n)} \rangle_\Theta \neq 0,$$

or equivalently, the $E^{(m)}$ being of maximum order,

$$\langle E^{(m)}, H^{(n)} \rangle_\Theta \begin{cases} = 0 & \text{if} \quad m \neq n, \\ \neq 0 & \text{if} \quad m = n. \end{cases}$$

The bilinear form of the theorem is the following

Theorem 3. *If $H^{(n)} = (H_1^{(n)}, ..., H_q^{(n)})^t$ and $E^{(m)} = (E_1^{(m)}, ..., E_p^{(m)})^t$ are sequences of vector polynomials of size q and p respectively, the first q (resp. p) of them have maximum order, defined by the recurrence relations*

$$xH^{(n)} = \sum_{k=-p}^{q} a_n^{(k)} H^{(n+k)}, \qquad H^{(-k)} = 0, \quad k > 0, \quad H^{(0)}, ..., H^{(q-1)} \textit{ fixed}$$

$$xE^{(n)} = \sum_{k=-q}^{p} \bar{a}_{n+k}^{(-k)} E^{(n+k)}, \qquad E^{(-k)} = 0, \quad k > 0, \quad E^{(0)}, ..., E^{(p-1)} \textit{ fixed},$$

where $a_n^{(q)} \neq 0$ and $a_n^{(-p)} \neq 0$ for all $n \geqslant 0$, then there exists a unique $p \times q$ matrix $\Theta = (\Theta_{i,j})_{i=1,...,p;\, j=1,...,q}$ of linear functionals such that

$$\langle E^{(m)}, H^{(n)} \rangle_\Theta = \delta_{n,m} \qquad n, m \geqslant 0.$$

Proof. Let us look at the values $\langle E^{(n)}, H^{(n)} \rangle_\Theta$. We have $E^{(n)} = \alpha_n e_n + \cdots$, as previously denoted, so if $n = \nu p + \nu_0$, $\nu_0 < p$, then

$$\begin{aligned} \langle E^{(n)}, H^{(n)} \rangle_\Theta &= \bar{\alpha}_n \langle e_n, H^{(n)} \rangle_\Theta \\ &= \bar{\alpha}_n \langle e_{n-p}, xH^{(n)} \rangle_\Theta \\ &= \bar{\alpha}_n a_n^{(-p)} \langle e_{n-p}, H^{(n-p)} \rangle_\Theta \\ &= \bar{\alpha}_n a_n^{(-p)} a_{n-p}^{(-p)} \cdots a_{n-(\nu-1)p}^{(-p)} \langle e_{\nu_0}, H^{(\nu_0)} \rangle_\Theta. \end{aligned}$$

The α_n can be expressed through the coefficients of the recurrence relation, identifying the leading coefficient

$$xE^{(n)} = \sum_{k=-q}^{p} \bar{a}_{n+k}^{(-k)} E^{(n+k)}$$

$$\alpha_n = \bar{a}_{n+p}^{(-p)} \alpha_{n+p}$$

$$\alpha_{v_0} = \bar{a}_{n-(v-1)p}^{(-p)} \cdots \bar{a}_n^{(-p)} \alpha_n$$

so finally

$$\langle E^{(n)}, H^{(n)} \rangle_\Theta = \langle E^{(v_0)}, H^{(v_0)} \rangle_\Theta.$$

If the initial conditions ensure $\langle E^{(n)}, H^{(n)} \rangle_\Theta = 1$ for $n = 0, ..., p-1$, then this relation is satisfied for all n.

The initial conditions are known from the $p(p+1)/2$ values $\langle e_i, h_j \rangle_\Theta$, $0 \leqslant j \leqslant i \leqslant p-1$, which is equivalent to knowing the $p(p+1)/2$ values $\langle E^{(i)}, H^{(j)} \rangle_\Theta$, $0 \leqslant j \leqslant i \leqslant p-1$. Finally to know that these last ones are the Kronecker symbol, defines a unique matrix of linear forms Θ such that

$$\forall m, n \geqslant 0, \qquad \langle E^{(m)}, H^{(n)} \rangle_\Theta = \delta_{n,m}. \tag{18}$$

8. CHRISTOFFEL–DARBOUX FORMULA

In the scalar case, for a family of polynomials $(P_k)_{k \geqslant 0}$ orthogonal with respect to a linear functional c, where $P_k(x) = t_k x^k + \cdots$ and $h_k = c(P_k^2)$ the Christoffel–Darboux formula is as follows

$$\sum_{k=0}^{n} \frac{1}{h_k} P_k(x).P_k(y) = \frac{t_n}{t_{n+1} h_n} \frac{P_{n+1}(x)\, P_n(y) - P_{n+1}(y)\, P_n(x)}{x - y}$$

The left hand side of this formula is to be considered as a reproducing kernel

$$K_n(x, y) = \sum_{k=0}^{n} \frac{1}{h_k} P_k(x).P_k(y)$$

$$c(K_n(x, y)\, P_k(y)) = P_k(x), \qquad k = 0, ..., n.$$

The generalization of the Christoffel–Darboux formula needs the definition of a kind of reproducing kernel ([1]) for the vector polynomials $H^{(n)}$ and $E^{(n)}$ which would be supposed to satisfy (18). We consider the following product, for vectors respectively of $H \in \mathbb{C}^q$ and $E \in \mathbb{C}^p$ as the $q \times p$ matrix, denoted $H\bar{E}$ (to avoid too many indices)

$$H\bar{E} = \begin{pmatrix} H_1 \\ \vdots \\ H_q \end{pmatrix} (\bar{E}_1 \cdots \bar{E}_p) = (H_i \bar{E}_j)_{i=1,\dots,q,\, j=1,\dots,p}$$

and then define $K_n(x, y)$ as the $q \times p$ matrix

$$K_n(x, y) = \sum_{m=0}^{n} H^{(m)}(x)\, \bar{E}^{(m)}(y). \tag{19}$$

Then

$$\begin{cases} K_n(x, y)\, \Theta H^{(k)}(y) = \sum_{m=0}^{n} H^{(m)}(x) \langle E^{(m)}(y), H^{(k)}(y) \rangle_\Theta = H^{(k)}(x) \\ \qquad\qquad k = 0, \dots, n \\ (K_n(x, y))^*\, \Theta^* E^{(m)}(x) = E^{(m)}(y) \qquad m = 0, \dots, n \end{cases}$$

To obtain a kind of Christoffel–Darboux formula, we now study

$$\begin{aligned} &(x-y)\, K_n(x, y) \\ &\quad = \sum_{m=0}^{n} (xH^{(m)}(x)\, \bar{E}^{(m)}(y) - H^{(m)}(x)\, y\bar{E}^{(m)}(y)) \\ &\quad = (H^{(0)}(x), \dots, H^{(n+q)}(x)) \begin{pmatrix} a_0^{(0)} & \cdots & a_p^{(-p)} & 0 & \\ \vdots & \ddots & & \ddots & \\ a_0^{(q)} & & \ddots & & a_n^{(-p)} \\ 0 & \ddots & & \ddots & \\ & & \ddots & & a_n^{(0)} \\ & & & \ddots & \\ & & & & a_n^{(q)} \end{pmatrix} \begin{pmatrix} \bar{E}^{(0)}(y) \\ \vdots \\ \bar{E}^{(n)}(y) \end{pmatrix} \\ &\qquad - (H^{(0)}(x), \dots, H^{(n)}(x)) \begin{pmatrix} a_0^{(0)} & \cdots & a_p^{(-p)} & 0 & & & \\ \vdots & \ddots & & \ddots & & & \\ a_0^{(q)} & & \ddots & & \ddots & & \\ 0 & \ddots & & \ddots & & \ddots & \\ & & a_{n-q}^{(q)} & \cdots & a_n^{(0)} & \cdots & a_{n+p}^{(-p)} \end{pmatrix} \\ &\qquad \times \begin{pmatrix} \bar{E}^{(0)}(y) \\ \vdots \\ \bar{E}^{(n+p)}(y) \end{pmatrix}, \end{aligned} \tag{20}$$

and finally, we get for $n \geqslant q-1$ and $n \geqslant p-1$ (the $H^{(n)}$ are taken in x and the $E^{(n)}$ in y)

$$(x-y)\, K_n(x, y)$$

$$= (H^{(n-p+1)}, ..., H^{(n+q)}) \begin{pmatrix} & & & -a_{n+1}^{(-p)} & & \\ & \mathcal{O}_{p,q} & & \vdots & \ddots & \\ & & & -a_{n+1}^{(-1)} & \cdots & -a_{n+p}^{(-p)} \\ a_{n-q+1}^{(q)} & \cdots & a_n^{(1)} & & & \\ & \ddots & \vdots & & \mathcal{O}_{q,p} & \\ & & a_{n+q}^{(q)} & & & \end{pmatrix}$$

$$\times \begin{pmatrix} \bar{E}^{(n-q+1)} \\ \vdots \\ \bar{E}^{(n+p)} \end{pmatrix}. \tag{21}$$

This formula (21) is the generalized Christoffel–Darboux identity. It is valid for $n \geqslant 0$, the polynomials of negative index being zero. In the scalar case where $p = q = 1$ the formula becomes

$$(x-y) \sum_{m=0}^{n} H^{(m)}(x)\, \bar{E}^{(m)}(y)$$
$$= (H^{(n)}(x), H^{(n+1)}(x)) \begin{pmatrix} 0 & -a_{n+1}^{(-1)} \\ a_n^{(1)} & 0 \end{pmatrix} \begin{pmatrix} \bar{E}^{(n)}(y) \\ \bar{E}^{(n+1)}(y) \end{pmatrix}$$

and the usual formula is recovered if moreover the matrix A is real and symmetric.

In the scalar case, Brezinski ([3]) has proved that the Christoffel–Darboux identity is equivalent to the three-term recurrence formula. A similar result can be proved here. Let two families of polynomials be given, respectively in $\mathbb{C}^q$ and $\mathbb{C}^p$ satisfying, for all n, the generalized Christoffel–Darboux formula (21) for given constants $a_i^{(j)}$.

We must first remark that the matrix A of the previous part is completely known except the main diagonal $(a_n^{(0)})_{n \geqslant 0}$. So we will consider it with arbitrary $a_n^{(0)}$ (we have put them equal to zero for simplicity, but these values would disappear anyway in the computation). The formula (20) is recovered and we write it in block form: the matrix in the middle is part of A^t, so B_n contains the n first rows and columns of A^t, γ_n the first n columns and the rows from index $n+1$ to $n+q$ and similarly for the other terms in such a way that the multiplications are possible (the $H^{(n)}$ are taken in x and the $E^{(n)}$ in y)

$$
\begin{aligned}
&(x-y)\,K_n(x,y)\\
&\quad=((H^{(0)}\cdots H^{(n-1)}),\,(H^{(n)}\cdots H^{(n+q-1)}),\,H^{(n+q)})\\
&\qquad\times\begin{pmatrix} B_{n-1} & c_n\\ \gamma_{n-1} & \gamma'_n\\ \mathcal{O}_{1,n-1} & a_n^{(q)}\end{pmatrix}\begin{pmatrix}\begin{pmatrix}\bar{E}^{(0)}\\ \vdots\\ \bar{E}^{(n-1)}\end{pmatrix}\\ \bar{E}^{(n)}\end{pmatrix}\\
&\qquad-((H^{(0)}\cdots H^{(n-1)}),\,H^{(n)})\begin{pmatrix} B_{n-1} & \beta_{n-1} & \mathcal{O}_{n-1,1}\\ l_n & \beta'_n & a_{n+p}^{(-p)}\end{pmatrix}\begin{pmatrix}\begin{pmatrix}\bar{E}^{(0)}\\ \vdots\\ \bar{E}^{(n-1)}\end{pmatrix}\\ \begin{pmatrix}\bar{E}^{(n)}\\ \vdots\\ \bar{E}^{(n+p-1)}\end{pmatrix}\\ \bar{E}^{(n+p)}\end{pmatrix}.
\end{aligned}
$$

From this last one, $(x-y)\,H^{(n)}(x)\,\bar{E}^{(n)}(y)$ is obtained as

$$
\begin{aligned}
&(x-y)\,H^{(n)}(x).\bar{E}^{(n)}(y)\\
&\quad=((H^{(0)}\cdots H^{(n-1)})\,c_n+(H^{(n)}\cdots H^{(n+q-1)})\,\gamma'_n+a_n^{(q)}H^{(n+q)})\,\bar{E}^{(n)}\\
&\qquad-H^{(n)}(l_n(\bar{E}^{(0)}\cdots\bar{E}^{(n-1)})^t+\beta'_n(\bar{E}^{(n)}\cdots\bar{E}^{(n+p-1)})^t+a_{n+p}^{(-p)}\bar{E}^{(n+p)})
\end{aligned}
$$

Separating the terms in x and the terms in y, it follows that

$$
\begin{aligned}
&\left(xH^{(n)}(x)-\sum_{k=-p}^{q}a_n^{(k)}H^{(n+k)}(x)\right)\bar{E}^{(n)}(y)\\
&\qquad=H^{(n)}(x)\left(y\bar{E}^{(n)}(y)-\sum_{k=-q}^{p}a_{n+k}^{(-k)}\bar{E}^{(n+k)}(y)\right).
\end{aligned}
$$

This equality can be written more compactly as $M(x)\,\bar{E}^{(n)}(y)=H^{(n)}(x)\,\bar{N}(y)$, which is equivalent, from the definition of the product to the equality for the components

$$M_i(x)\,\bar{E}_j^{(n)}(y)=H_i^{(n)}(x)\,\bar{N}_j(y),\qquad i=1,...,q,\quad j=1,...,p$$

so there exists a constant λ_n independent of i, j, x, y such that

$$M(x)=\lambda_n H^{(n)}(x),\qquad \bar{N}(y)=\lambda_n\bar{E}^{(n)}(y)$$

which is the same as the required recurrence relations with now $a_n^{(0)} = \lambda_n$

$$\begin{cases} xH^{(n)}(x) = \sum\limits_{k=-p}^{q} a_n^{(k)} H^{(n+k)}(x) \\[2ex] yE^{(n)}(y) = \sum\limits_{k=-q}^{p} \bar{a}_{n+k}^{(-k)} E^{(n+k)}(y) \end{cases}$$

From the previous generalization of the Shohat-Favard theorem, it follows that there exists a matrix of linear forms Θ such that

$$\langle E^{(n)}, H^{(m)} \rangle_\Theta = \delta_{n,m},$$

this leads to the expression of $a_n^{(0)}$, the only coefficient not given as data

$$a_n^{(0)} = \langle E^{(n)}, xH^{(n)} \rangle_\Theta$$

and this formula is in fact true for all the coefficients because of the link between $a_n^{(k)}$ and Θ

$$a_n^{(k)} = \langle E^{(n+k)}, xH^{(n)} \rangle_\Theta, \qquad k = -p, ..., q.$$

9. MATRIX POLYNOMIALS

In [6] orthogonal matrix polynomials, i.e., polynomials with square matrix coefficients or equivalently square matrices of polynomials are considered and their recurrence relations studied. For the sake of simplicity, we suppose in the sequel that $p \geqslant q$, in the other case, we would have to consider the polynomials $E^{(n)}$ of size p. Let us rewrite in parallel the recurrence relations satisfied by the vector polynomials $H^{(n)}$ for d consecutive indices

$$x(H^{(n)}, ..., H^{(n+d-1)}) = (H^{(n-p)}, ..., H^{(n+d-1+q)}) \begin{pmatrix} a_n^{(-p)} & & \\ \vdots & \ddots & a_{n+d-1}^{-p)} \\ a_n^{(0)} & \cdots & a_{n+d-1}^{(-q)} \\ \vdots & \ddots & \\ a_n^{(q)} & \cdots & a_{n+d-1}^{(0)} \\ & \ddots & \vdots \\ & & a_{n+d-1}^{(q)} \end{pmatrix}$$

If d is taken as the greatest common divisor (g.c.d.) of p and q, then we set $p=\alpha d$ and $q=\beta d$. The matrix on the right hand side can be put in blocks of size $d\times d$ and becomes

$$x(H^{(n)}, ..., H^{(n+d-1)}) = \sum_{k=-\alpha}^{\beta} (H^{(n+kd)}, ... H^{(n+kd+d-1)})\Gamma^k$$

where the first and the last matrices $\Gamma^{-\alpha}$ and Γ^{β} are triangular and invertible.

Now we define the matrix polynomial of size $q\times d$ by $Q_N=(H^{(Nd)}, ..., H^{(Nd+d-1)})$ and the preceding recurrence relation for the $H^{(n)}$ is equivalent to the relation with square matrix coefficients Γ of size $d\times d$

$$\begin{cases} xQ_N = \sum_{k=-\alpha}^{\beta} Q_{N+k}.\Gamma_N^k, & \\ \Gamma_N^{-\alpha}, \Gamma_N^{\beta} & \text{resp. lower and upper triangular, invertible} \\ \alpha, \beta \in \mathbb{N} & \text{relatively prime} \end{cases}$$

If $q=1$ the relation is the relation characterizing the vector orthogonality of dimension p, if $p=q$, a three term recurrence relation is obtained as in [6] where moreover $\Theta=\Theta^*$ or $A=A^*$ and so $\Gamma_N^{-\alpha}=(\Gamma_{N+1}^{\beta})^*$.

Conversely, suppose a recurrence relation with square matrix coefficients $d\times d$ is given for matrix polynomials of size $q\times d$

$$xQ_N = \sum_{k=-\alpha}^{\beta} Q_{N+k}.\Gamma_N^k, \qquad N\geqslant 0,$$

with the first and the last matrices $\Gamma_N^{-\alpha}$ and Γ_N^{β} invertible. By changing d to $d'=\delta d$ (δ is the g.c.d. of α, β), and Q_N to $(Q_N, ..., Q_{N+\delta-1})$, the recurrence relation becomes a recurrence relation of the same kind where the sum is from $-\alpha' = -\alpha/\delta$ to $\beta'=\beta/\delta$ which are now relatively prime.

With the method explained in [6], it can be assumed without loss of generality that one of the matrices $\Gamma_N^{-\alpha}, \Gamma_N^{\beta}$ is triangular, but without the assumption $A=A^*$, the other one is not triangular.

Let us suppose that $\Gamma_N^{-\alpha}$ is lower triangular for all N. The columns of Q_N are denoted by $H^{(Nd+i)}$, for i between 0 and $d-1$, so the sequence of vectors of size $q(H^{(n)})_n$ is defined and the recurrence for the matrix polynomials give the following recurrence relation for the $H^{(n)}$

$$\begin{aligned} xH^{(Nd+i)} &= a_{Nd+i}^{(-p)}H^{(Nd+i-p)} + \cdots + a_{Nd+i}^{(q)}H^{(Nd+i+q)} \\ &\quad + \cdots + a_{Nd+i}^{(q+d-1-i)}H^{(Nd+q+d-1)}. \end{aligned}$$

If the initial conditions are of maximum order, then it is clear that all the vector polynomials are of maximum order only if the matrices Γ_N^{β} are triangular.

It is now natural to define “matrix polynomials of maximum order” as matrices of polynomials where all the columns are of maximum order, which means that, expanded as a polynomial with matrix coefficients, Q_N has as coefficient of x^N a triangular invertible matrix.

If the study of orthogonal matrix polynomials is limited to matrix polynomials of maximum order, this study is completely done through the study of orthogonal vector polynomials defined very classically by orthogonality with respect to a bilinear form, i.e., by a matrix of linear forms defining a formal inner product (formal because it is not defined by a symmetric, positive definite matrix of forms) and characterized either by the recurrence formula, or by the Christoffel–Darboux identity.

ACKNOWLEDGMENT

Sincere thanks are due to B. Beckermann for fruitful discussions.

REFERENCES

1. B. Beckermann, Nonsymmetric difference operators and polynomials being orthogonal with respect to rectangular matrix valued measures, preprint ANO 335, Lille, 1995.
2. C. Brezinski, “Padé-Type Approximation and General Orthogonal Polynomials,” International Series of Numerical Mathematics, Vol. 50, Birkhauser Verlag, Basel, 1980.
3. C. Brezinski, A direct proof of the Christoffel–Darboux identity and its equivalence to the recurrence relationship, *J. Comput. Appl. Math.* **32** (1990), 17–25.
4. C. Brezinski, “Biorthogonality and Its Applications to Numerical Analysis,” Pure and Applied Math., Dekker, New York, 1992.
5. A. Bultheel and M. Van Barel, Formal orthogonal polynomials and Hankel/Toeplitz duality, *Numer. Algorithms* **10** (1995), 289–335.
6. A. J. Durán and W. Van Assche, Orthogonal matrix polynomials and higher order recurrence relations, *Linear Algebra Appl.* **219** (1995), 261–280.
7. G. Labahn and B. Beckermann, A uniform approach for Hermite–Padé and simultaneous Padé approximants and their matrix type generalization, *Numer. Algorithms* **3** (1992), 45–54.
8. B. Beckermann and G. Labahn, A uniform approach for the fast computation of Matrix Padé Approximants, *SIAM J. Matrix Anal. Appl.* **15** (1994), 804–823.
9. K. Mahler, Perfect systems, *Compositio Math.* **19** (1968), 95–166.
10. A. M. Nikishin and V. N. Sorokin, “Rational Approximations and Orthogonality,” Transl. of Math. Monographs, Vol. 92, Amer. Math. Society, Providence, RI, 1991.
11. V. N. Sorokin, On simultaneous approximations of few linear forms, *Vestnik Moskov. Univ. Ser. I Mat. Mekh.* **1** (1983), 44–47 [in Russian].
12. J. Van Iseghem, “Approximants de Padé Vectoriels,” thesis, Univ. de Lille, France, 1987.
13. J. Van Iseghem, Vector orthogonal relations, Vector *QD*-algorithm, *J. Comput. Appl. Math.* **19**, No. 1 (1987), 141–150.