Generalizations of orthogonal polynomials

A. Bultheel, Annie Cuyt, W. Van Assche,
M. Van Barel, B. Verdonk

Report TW 375, December 2003

Katholieke Universiteit Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

Generalizations of orthogonal polynomials

A. Bultheel, Annie Cuyt, W. Van Assche,
M. Van Barel, B. Verdonk

Report TW 375, December 2003

Department of Computer Science, K.U.Leuven

We give a survey of recent generalizations for orthogonal polynomials that were recently obtained. It concerns not only multidimensional (matrix and vector orthogonal polynomials) and multivariate versions, or multipole (orthogonal rational functions) variants of the classical polynomials but also extensions of the orthogonality conditions (multiple orthogonality). Most of these generalizations are inspired by the applications in which they are applied. We also give a glimpse of the applications, but they are usually also generalizations of applications where classical orthogonal polynomials play a fundamental role: moment problems, numerical quadrature, rational approximation, linear algebra, recurrence relations, random matrices.

Abstract

Keywords : system identification, orthogonal rational functions, least squares.
AMS(MOS) Classification : Primary : 93B30, Secondary : 42C05, 93E24

Generalizations of orthogonal polynomials ⋆

A. Bultheel [1]

Dept. Computer Science (NALAG), K.U.Leuven,
Celestijnenlaan 200 A, B-3001 Leuven, Belgium.
adhemar.bultheel@cs.kuleuven.ac.be

A. Cuyt

Dept. Mathematics and Computer Science, Universiteit Antwerpen,
Middelheimlaan 1, B-2020 Antwerpen, Belgium.
annie.cuyt@ua.ac.be

W. Van Assche [2]

Dept. Mathematics, K.U.Leuven
Celestijnenlaan 200B, B-3001 Leuven, Belgium.
walter.vanassche@wis.kuleuven.ac.be

M. Van Barel [1]

Dept. Computer Science (NALAG), K.U.Leuven,
Celestijnenlaan 200 A, B-3001 Leuven, Belgium.
marc.vanbarel@cs.kuleuven.ac.be

B. Verdonk

Dept. Mathematics and Computer Science, Universiteit Antwerpen,
Middelheimlaan 1, B-2020 Antwerpen, Belgium.
brigitte.verdonk@ua.ac.be

Abstract

We give a survey of recent generalizations for orthogonal polynomials that were recently obtained. It concerns not only multidimensional (matrix and vector orthogonal polynomials) and multivariate versions, or multipole (orthogonal rational functions) variants of the classical polynomials but also extensions of the orthogonality conditions (multiple orthogonality). Most of these generalizations are inspired by the applications in which they are applied. We also give a glimpse of the applications, but they are usually also generalizations of applications where classical orthogonal polynomials play a fundamental role: moment problems, numerical quadrature, rational approximation, linear algebra, recurrence relations, random matrices.

Key words: Orthogonal polynomials, rational approximation, linear algebra
2000 MSC: 42C10, 33D45, 41A, 30E05, 65D30, 46E35

This paper is dedicated to Olav Njåstad on the occasion of his 70th birthday.

1 Introduction

Since the fundamental work of Szegő [46], orthogonal polynomials have been a basic tool in the analysis of basic problems in mathematics and engineering. For example moment problems, numerical quadrature, rational and polynomial approximation and interpolation, and all the direct or indirect applications of these techniques in engineering and applied problems, they are all indebted to the basic properties of orthogonal polynomials.

The first thing one needs is an inner product defined on the space of polynomials. There are several formalizations of this concept. For example, one can define a positive definite Hermitian linear functional $M[\cdot]$ on the space of polynomials. This means the following. Let Π_n the space of polynomials of degree at most n and Π the space of all polynomials. The dual space of Π_n is Π_{n*}, namely the space of all linear functionals. With respect to a set of basis functions $\{B_0, B_1, \ldots, B_n\}$ that span Π_n for $n = 0, 1, \ldots$, it is clear that a polynomial has a uniquely defined set of coefficients, representing this polynomial. Thus, given a nested basis of Π, we can identify the complex polynomials Π_n with the space $\mathbb{C}^{(n+1)\times 1}$ of complex $(n+1) \times 1$ column vectors.

Suppose the dual space is spanned by a sequence of basic linear functionals $\{L_k\}_{k=0}^{\infty}$, and define $\Pi_{n*} = \text{span}\{L_0, L_1, \ldots, L_n\}$ for $n = 0, 1, 2, \ldots$. Then the dual subspace Π_{n*} can be identified with $\mathbb{C}^{1\times(n+1)}$, the space of complex $1 \times (n+1)$ row vectors. Now, given a sequence of linear functionals $\{L_k\}_{k=0}^{\infty}$, we say that a sequence of polynomials $\{P_k\}_{k=0}^{\infty}$ with $P_k \in \Pi_k$, is orthonormal with respect to the sequence of linear functionals $\{L_k\}_{k=0}^{\infty}$ with $L_k \in \Pi_{k*}$, if

$$L_k(P_l) = \delta_{kl}, \quad k, l = 0, 1, 2 \ldots.$$

Hereby we have to assure some non-degeneracy, which means that the moment matrix of the system is Hermitian positive definite. This moment matrix is defined as follows. Consider the basis $B_0, B_1, \ldots$ in Π and a basis $L_0, L_1, \ldots$ for the dual space Π_*, then the moment matrix is

---

⋆ This work is partially supported by the Fund for Scientific Research (FWO) via the Scientific Research Network "Advanced Numerical Methods for Mathematical Modeling", grant #WO.012.01N.

[1] The work of these authors is partially supported by the Fund for Scientific Research (FWO) via the projects G.0078.01 "SMA: Structured Matrices and their Applications", G.0176.02 "ANCILA: Asymptotic aNalysis of the Convergence behavior of Iterative methods in numerical Linear Algebra", G.0184.02 "CORFU: Constructive study of Orthogonal Functions", G.0455.04 "RHPH: Riemann-Hilbert problems, random matrices and Padé-Hermite approximation", the Research Council K.U.Leuven, project OT/00/16 "SLAP: Structured Linear Algebra Package", and the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. The scientific responsibility rests with the authors.

[2] The work of this author is partially supported by the Fund for Scientific Research (FWO) via the projects G.0184.02 "CORFU: Constructive study of Orthogonal Functions" and G.0455.04 "RHPH: Riemann-Hilbert problems, random matrices and Padé-Hermite approximation".

the infinite matrix

$$M = \begin{bmatrix} m_{00} & m_{01} & m_{02} & \dots \\ m_{10} & m_{11} & m_{12} & \dots \\ m_{20} & m_{21} & m_{22} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad \text{with} \quad m_{ij} = L_i(B_j).$$

It is Hermitian positive definite if $M_{kk} = [m_{ij}]_{i,j=0}^{k}$ is Hermitian positive definite for all $k = 0, 1, \dots$.

In some formal generalizations, positive definiteness may not be necessary; a nondegeneracy condition is then sufficient. In other applications it is even not really necessary to impose this nondegeneracy condition, and in that case there should be some notion of block orthogonality because the existence of an orthonormal set is not guaranteed anymore.

Note that if the coefficients of $P \in \Pi_n$ and $Q_* \in \Pi_{n*}$ are given by $p = [p_0, p_1, \dots]^T$ and $q = [q_0, q_1, \dots]$ respectively, then $Q_*(P) = qMp$.

Classical cases fall into this framework. For example consider a positive measure μ of a finite or infinite interval $\mathbb{I}$ on the real line, a basis $1, x, x^2, \dots$ for the space of real polynomials and a basis of linear functionals $L_0, L_1, \dots$ defined by

$$L_k(P) = \int_{\mathbb{I}} x^k P(x) \mathrm{d}\mu(x),$$

then we can choose L_k as the dual of the polynomial x^k and therefore introduce an inner product in Π as

$$\langle Q, P \rangle = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} q_k p_l \left\langle x^k, x^l \right\rangle = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} q_k p_l L_k(x^l) = Q_*(P),$$

if $Q_* = \sum_{k=0}^{\infty} q_k L_k$, $Q(x) = \sum_{k=0}^{\infty} q_k x^k$, and $P(x) = \sum_{k=0}^{\infty} p_k x^k$. If μ is a positive measure, the moment matrix is guaranteed to be positive definite.

Note that in this case we need to define only one linear functional L on Π to determine the whole moment matrix. It is defined as $L(x^i) = \int_{\mathbb{I}} x^i \mathrm{d}\mu(x)$. The moment matrix is completely defined by the sequence $m_k = L(x^k)$, $k = 0, 1, 2, \dots$.

Another famous case is obtained by orthogonality on the unit circle. Consider $\mathbb{T} = \{t \in \mathbb{C} : |t| = 1\}$ and a positive measure on $\mathbb{T}$. The set of complex polynomials are spanned by $1, z, z^2, \dots$ and we consider linear functionals L_k defined by

$$L_k(z^l) = L(z^{l-k}) = \int_{\mathbb{T}} t^{l-k} \mathrm{d}\mu(t), \quad k, l = 0, 1, 2, \dots.$$

Thus we can again use only one linear functional $L(P(z)) = \int_{\mathbb{T}} P(t)\mathrm{d}\mu(t)$ and define a positive definite Hermitian inner product on the set of complex polynomials by

$$\begin{aligned}
\langle Q, P\rangle &= \left\langle \sum_{k=0}^{\infty} q_k z^k, \sum_{l=0}^{\infty} p_l z^l \right\rangle = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \overline{q}_k p_l \left\langle z^k, z^l \right\rangle \\
&= \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \overline{q}_k p_l L_k(z^l) = \sum_{k=0}^{\infty} \overline{q}_k L_k \left(\sum_{l=0}^{\infty} p_l z^l \right) = Q_*(P) \\
&= \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \overline{q}_k p_l L(z^{l-k}) = \int_{\mathbb{T}} \left(\sum_{k=0}^{\infty} \overline{q}_k t^{-k} \right) \left(\sum_{l=0}^{\infty} p_l t^l \right) \mathrm{d}\mu(t) \\
&= \int_{\mathbb{T}} Q_*(t) P(t) \mathrm{d}\mu(t).
\end{aligned}$$

where we have abused the notation Q_* for both the linear functional $Q_* = \sum_{k=0}^{\infty} \overline{q}_k L_k$ and for the dual polynomial $Q_*(z) = \sum_{k=0}^{\infty} q_k z^{-k}$, which is the dual of $Q(z) = \sum_{k=0}^{\infty} q_k z^k$, and we have set $P(z) = \sum_{l=0}^{\infty} p_l z^l$. Note that here the linear functional L is defined on the space of Laurent polynomials $\Lambda = \operatorname{span}\{z^k : k \in \mathbb{Z}\}$. The moment matrix is completely defined by the one dimensional sequence $m_k = L(z^k)$, $k \in \mathbb{Z}$, and because μ is positive, it is sufficient to give m_k, $k = 0, 1, 2 \ldots$ because $m_{-k} = L(z^{-k}) = \overline{L(z^k)} = \overline{m}_k$.

Note that in the case of polynomials orthogonal on an interval on the real line, the moment matrix $[m_{kl}]$ is real and has a Hankel structure and in the case of orthogonality on the circle, then the moment matrix is complex Hermitian and has a Toeplitz structure. This explains of course why a single sequence defines the whole matrix in both cases.

One of the basic problems that is considered, is the moment problem where it is required to recover from a positive definite moment matrix a representation of the inner product. In the examples above, it is to find the positive measure μ from the (positive) moment sequence $\{m_k\}$.

The relation with structured linear algebra problems has given rise to an intensive research on fast algorithms for the solution of linear systems of equations and other linear algebra problems. The duality between real Hankel and complex Toeplitz is in this context a natural distinction. However, what is possible for one case is usually also true in some form for the other case.

For example, the Hankel structure is at the heart of the famous three-term recurrence relation for orthogonal polynomials. For 3 successive orthogonal polynomials $\phi_n, \phi_{n-1}, \phi_{n-2}$ there are constants A_n, B_n, and C_n with $A_n > 0$ and $C_n > 0$ such that

$$\phi_n(x) = (A_n x + B_n)\phi_{n-1}(x) - C_n \phi_{n-2}(x), \quad n = 2, 3, \ldots$$

Closely related to this recurrence is the Christoffel-Darboux relation which gives a closed form expression for the (reproducing) kernel $k_n(z, w)$

$$k_n(x, y) := \sum_{k=0}^{n} \phi_n(x)\phi_n(y) = \frac{\kappa_n}{\kappa_{n+1}} \frac{\phi_{n+1}(x)\phi_n(y) - \phi_n(x)\phi_{n+1}(y)}{x - y}$$

where κ_n is the highest degree coefficient of ϕ_n. All three items: orthogonality, three-term recurrence, and a Christoffel-Darboux relation are in a sense equivalent. The Favard theorem states that if there is a three-term recurrence relation with certain properties, then the sequence of polynomials that it generates will be a sequence of orthogonal polynomials with respect to some inner product. Brezinski showed [12] that the Christoffel-Darboux relation is equivalent with the recurrence relation.

In the case of the unit circle, another fundamental type of recurrence relation is due to Szegő. The recursion is of the form

$$\phi_{k+1}(z) = c_{k+1}[z\phi_k(z) + \rho_{k+1}\phi_k^*(z)]$$

where for any polynomial P_k of degree k we set

$$P_k^*(z) = z^k P_{k*}(z) = z^k \overline{P_k(1/\overline{z})},$$

so that ϕ_k^* is the reciprocal of ϕ_k, ρ_{k+1} is a Szegő parameter and $c_{k+1} = (1 - |\rho_{k+1}|^2)^{-1/2}$ is a normalizing constant. This recurrence relation plays the same fundamental role as the three-term recurrence relation does for orthogonality on (part of) the real line. There is a related Favard-type theorem and a Christoffel-Darboux-type of relation that now has the complex form

$$k_n(z,w) := \sum_{k=0}^{n} \phi_n(z)\overline{\phi_n(w)} = \frac{\phi_{n+1}^*(z)\overline{\phi_{n+1}^*(w)} - \phi_{n+1}(z)\overline{\phi_{n+1}(w)}}{1 - z\overline{w}}.$$

Another basic aspect of orthogonal polynomials is rational approximation. Rational approximation is given through the fact that truncating a continued fraction gives an approximant for the function to which it converges. The link with orthogonal polynomials is that continued fractions are essentially equivalent with three-term recurrence relations, and orthogonal polynomials on an interval are known to satisfy such a recurrence. In fact if the orthogonal polynomials are solutions of the recurrence with starting values $\phi_{-1} = 0$ and $\phi_0 = 1$, then an independent solution can be obtained as a polynomial sequence $\{\psi_k\}$ by using the initial conditions $\psi_{-1} = -1$ and $\psi_0 = 0$. It turns out that

$$\psi_n(x) = L\left(\frac{\phi_n(x) - \phi_n(y)}{x - y}\right) = \int_{\mathbb{I}} \frac{\phi_n(x) - \phi_n(y)}{x - y} \mathrm{d}\mu(x)$$

where L is the linear functional defining the inner product on $\mathbb{I} \subset \mathbb{R}$. (Note that ψ_n is a polynomial of degree $n-1$.) Therefore, the nth approximant of the continued fraction

$$\frac{1\ |}{|\ A_1 x + B_1} - \frac{C_2\ |}{|\ A_2 z + B_2} - \frac{C_3\ |}{|\ A_3 z + B_3} - \cdots$$

is given by $\psi_n(x)/\phi_n(x)$. The continued fraction converges to the Stieltjes transform or Cauchy transform (note the Cauchy kernel $C(x,y) = 1/(x-y)$)

$$F_\mu(x) = L\left(\frac{1}{x-y}\right) = \int_{\mathbb{I}} \frac{\mathrm{d}\mu(y)}{x-y}.$$

The approximant is a Padé approximant at ∞ because

$$\frac{\psi_n(x)}{\phi_n(x)} = \frac{m_0}{x} + \frac{m_1}{x^2} + \cdots + \frac{m_{2n-1}}{x^{2n}} + \mathcal{O}\left(\frac{1}{x^{2n+1}}\right) = F_\mu(x) + \mathcal{O}\left(\frac{1}{x^{2n+1}}\right), \quad x \to \infty.$$

All the $2n+1$ free parameters in the rational function ψ_n/ϕ_n of degree n are used to fit the first $2n+1$ coefficients in the asymptotic expansion of F_μ at ∞.

Again, there is an analog situation for the unit circle case. Then the function that is approximated is a Riesz-Herglotz transform

$$F_\mu(z) = \int_{\mathbb{T}} \frac{t+z}{t-z} d\mu(t).$$

where now the Riesz-Herglotz kernel $D(t,z) = (t+z)/(t-z)$ is used. This function is analytic in the open unit disk and has a positive real part for $|z| < 1$. It is therefore a Carathéodory function. By the Cayley transform, one can map the right half plane to the unit disk, by which we can transform a Carathéodory function F into a Schur function, since indeed $S(z) = (F_\mu(z) - F_\mu(0))/[z(F_\mu(z) + F_\mu(0))]$ is a function analytic in the unit disk and $|S(z)| < 1$ for $|z| < 1$. It is in this framework that Schur has developed his famous algorithm to check whether a function is in the Schur class. It is based on the simple lemma saying that S is in the Schur class if and only if $|S(0)| < 1$ and $S_1(z) = \frac{1}{z}(S(z) - S(0))/(1 - \overline{S(0)}S(z))$ is in the Schur class. Applying this lemma recursively gives the complete test. This kind of test is closely related to a stability test for polynomials in discrete time linear system theory or the solution of difference equations. It is known as the Jury test. A similar derivation exists for the case of an interval on the real line, which leads to the Routh-Hurwitz test, which is a bit more involved.

Note also that the moments show up as Fourier-Stieltjes coefficients of F_μ because

$$F_\mu(z) = \int_{\mathbb{T}} \left[1 + 2\sum_{k=1}^{\infty} \frac{z^k}{t^k}\right] \mathrm{d}\mu(t) = m_0 + 2\sum_{k=1}^{\infty} m_{-k} z^k.$$

It is again possible to construct a continued fraction whose approximants are alternatingly ψ_n/ϕ_n and ψ_{n*}/ϕ_{n*}, and these are two-point Padé approximants at the origin and infinity for F_μ in a linearized sense.

I.e., one has

$$\begin{aligned} F_\mu(z) + \psi_n(z)/\phi_n(z) &= \mathcal{O}(z^{-n-1}), \quad z \to \infty, \\ F_\mu(z)\phi_n(z) + \psi_n(z) &= \mathcal{O}(z^n), \quad z \to 0, \end{aligned}$$

and

$$F_\mu(z)\phi_{n*}(z) - \psi_{n*}(z) = \mathcal{O}(z^{-n}), \quad z \to \infty,$$
$$F_\mu(z) - \psi_{n*}(z)/\phi_{n*}(z) = \mathcal{O}(z^{n+1}), \quad z \to 0.$$

Here the ψ_n are defined by

$$\psi_n(z) = L\left(D(t,z)[\psi_n(t) - \phi_n(z)]\right) = \int_{\mathbb{T}} \frac{t+z}{t-z}[\phi_n(t) - \phi_n(z)]\mathrm{d}\mu(t).$$

The term two-point Padé approximant is justified by the fact that the interpolation is in the points 0 and ∞ and the number of interpolation conditions equals the degrees of freedom in the rational function of degree n. Since ϕ_{n*}/ψ_{n*} is a rational Carathéodory function, it is a solution of a partial Carathéodory coefficient problem. This is the problem of finding a Carathéodory function with given coefficients for its expansion at the origin. To a large extent the Schur interpolation problem, the Carathéodory coefficient problem and the trigonometric moment problem are all equivalent.

Another important aspect directly related to orthogonal polynomials and the previous approximation properties is numerical quadrature formulas. By a quadrature formula for the integral $I_\mu(f) := \int_{\mathbb{I}} f(x)\mathrm{d}\mu(x)$ is meant a formula of the form $I_n(f) := \sum_{k=1}^n w_{nk} f(\xi_{nk})$. The knots ξ_{nk} should be in distinct points that are preferably in $\mathbb{I}$, the support of the measure μ, and the weights are preferably positive. Both these requirements are met by the Gauss quadrature formulas, i.e., when the ξ_{nk} are chosen as the n zeros of the orthogonal polynomial ϕ_n. The weights or Christoffel numbers are then given by $w_{nk} = 1/k_n(\xi_{nk}, \xi_{nk}) = 1/\sum_{k=0}^n [\phi_n(\xi_{nk})]^2$ and the quadrature formula has the maximal domain of validity in the set of polynomials. This means that $I_n(f) = I_\mu(f)$ for all f that are polynomials of degree at most $2n-1$. It can be shown that there is no n-point quadrature formula that will be exact for all polynomials of degree $2n$, so that the polynomial degree of exactness is maximal.

In the case of the unit circle, the integral $I_\mu(f) := \int_{\mathbb{T}} f(t)\mathrm{d}\mu(t)$ is again approximated by a formula of the form $I_n(f) := \sum_{k=1}^n w_{nk} f(\xi_{nk})$, where now the knots are preferably on the unit circle. However, the zeros of ϕ_n are known to be strictly less than one in absolute value. Therefore, the para-orthogonal polynomials are introduced as

$$Q_n(z,\tau) = \phi_n(z) + \tau\phi_n^*(z), \quad \tau \in \mathbb{T}.$$

It is known that these polynomials have exactly n simple zeros on $\mathbb{T}$ and thus they can be used as knots for a quadrature formula. If the corresponding weights are chosen as before, namely $w_{nk} = 1/k_n(\xi_{nk}, \xi_{nk}) = 1/\sum_{k=0}^n |\phi_n(\xi_{nk})|^2$, then these are obviously positive and the quadrature formula becomes a Szegő formula, again with maximal domain of validity, namely $I_n(f) = I_\mu(f)$ for all f that are in the span of $\{z^{-n+1}, \ldots, z^{n-1}\}$, a subspace of dimension $2n-1$ in the space of Laurent polynomials.

We have just given the most elementary account of what orthogonal polynomials are related to. Many other aspects are not even mentioned: for example the tridiagonal Jacobian operator (real case) or the unitary Hessenberg operator (circle case) which catch the recurrence relation in one operator equation, also the deep studies by Geronimus, Freud and many others to study the asymptotics of orthogonal polynomials, their zero distribution, and many other properties under various assumptions on the weights and/or on the recurrence parameters [25,28,36,37,45], there are the differential equations like Rodrigues formulas and generating functions that hold for so called classical orthogonal polynomials, the whole Askey-Wilson scheme, introducing a wealth of extensions for the two simplest possible schemes that were introduced above.

Also from the application side there are many generalizations, some are formal [16] orthogonality relations inspired by fast algorithms for linear algebra, some are matrix and vector forms of orthogonal polynomials which are often inspired by linear system theory and all kind of generalizations of rational interpolation problems. And so further and so on.

In this paper we want to give a survey of recent achievements about generalizations of orthogonal polynomials. What we shall present is just a sample of what is possible and reflects the interest of the authors. It is far from being a complete survey. Nevertheless, it is an illustration of the diversity of possibilities that are still open for further research.

2 Orthogonal rational functions

One of the recent generalizations of orthogonal polynomials that has emerged during the last decade is the analysis of orthogonal rational functions. They were first introduced by Djrbashian in the 1960's. Most of his papers appeared in the Russian literature. An accessible survey in English can be found in [22]. Details about this section can be found in the book [14]. For a survey about their use in numerical quadrature see the survey paper [15], for another survey and further generalizations see [13]. Several results about matrix-valued orthogonal rational functions are found in [34,26,27].

2.1 Orthogonal rational functions on the unit circle

Some connections between orthogonal polynomials and other related problems were given in the introduction to this paper. The simplest way to introduce the orthogonal rational functions is to look at a slight generalization of the Schur lemma. With the Schur function constructed from a positive measure μ as it was in the introduction, the lemma says that if μ has infinitely many points of increase, then for some $\alpha \in \mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ we have $S(\alpha) \in \mathbb{D}$ and S_1 is again a Schur function if $S_1(z) = S_\alpha(z)/\zeta_\alpha(z)$ with $S_\alpha(z) = (S(z) - S(\alpha))/(1 - \overline{S(\alpha)}S(z))$ and $\zeta_\alpha(z) = (z - \alpha)/(1 - \overline{\alpha}z)$. As in the polynomial case, a recursive application of this lemma leads to some continued fraction-like algorithm that computes for a given sequence of points $\{\alpha_k\}_{k=1}^{\infty} \subset \mathbb{D}$ (with or without repetitions) a sequence of parameters $\rho_k = S_k(\alpha_{k+1})$ that are all in $\mathbb{D}$ and that are generalizations of the Szegő parameters.

Thus instead of taking all the $\alpha_k = 0$, which yields the Szegő polynomials, we obtain a multipoint generalization. The multipoint generalization of the Schur algorithm is the algorithm of Nevanlinna-Pick. It is well known that this algorithm constructs rational approximants of increasing degree that interpolate the original Schur function S in the successive points α_k. Suppose we define the successive Schur functions $S_n(z)$ as the ratio of two functions analytic in $\mathbb{D}$, namely $S_n(z) = \Delta_{n1}(z)/\Delta_{n2}(z)$, then the Schur recursion reads ($\rho_{n+1} = S_n(\alpha_{n+1})$ and $\zeta_n(z) = z_n \frac{z-\alpha_n}{1-\overline{\alpha}_n z}$ with $z_n = 1$ if $\alpha_n = 0$ and $z_n = \overline{\alpha}_n/|\alpha_n|$ otherwise)

$$[\Delta_{n+1,1} \quad \Delta_{n+1,2}] = [\Delta_{n,1} \quad \Delta_{n,2}] \begin{bmatrix} 1 & -\overline{\rho}_{n+1} \\ -\rho_{n+1} & 1 \end{bmatrix} \begin{bmatrix} 1/\zeta_{n+1} & 0 \\ 0 & 1 \end{bmatrix}.$$

This describes the recurrence for the tails. The inverse recurrence is the recurrence for the partial numerators and denominators of the underlying continued fraction:

$$[\phi_{n+1} \quad \phi^*_{n+1}] = [\phi_n \quad \phi^*_n] \begin{bmatrix} \zeta_{n+1} & 0 \\ 0 & 1 \end{bmatrix} \frac{1}{\sqrt{1-|\rho_{n+1}|^2}} \begin{bmatrix} 1 & \overline{\rho}_{n+1} \\ \rho_{n+1} & 1 \end{bmatrix}.$$

When starting with $\phi_0 = \phi^*_0 = 1$, this generates rational functions ϕ_n which are of degree n and which are in certain spaces with poles among the points $\{1/\overline{\alpha}_k\}$

$$\phi_n, \phi^*_n \in \mathcal{L}_n = \text{span}\{B_0, B_1, \ldots, B_n\} = \left\{\frac{p_n}{\pi_n} : p_n \in \Pi_n\right\}$$

where $\pi_n(z) = \prod_{k=1}^n (1 - \overline{\alpha}_j z)$ and the finite Blaschke products are defined by

$$B_0 = 1, \quad B_k = \zeta_1 \zeta_2 \cdots \zeta_k.$$

Moreover, it is easily verified that $\phi_n(z) = B_n(z)\phi_{n*}(z)$ where $\phi_{n*}(z) = \overline{\phi_n(1/\overline{z})}$. This should make clear that the recurrence

$$\phi_{n+1}(z) = c_{n+1}[\zeta_{n+1}(z)\phi_n(z) + \rho_{n+1}\phi^*_n(z)], \quad c_{n+1} = (1 - |\rho_{n+1}|^2)^{-1/2}$$

is a generalization of the Szegő recurrence relation.

Transforming back from the Schur to the Carathéodory domain, the approximants of the Schur function correspond to rational approximants of increasing degree that interpolate the functions F_μ in the points α_k. Defining the rational functions of the second kind ψ_n exactly as in the polynomial case, then we have multipoint Padé approximants since

$$\begin{aligned} &zB_n(z)[F_\mu(z) + \psi_n(z)/\phi_n(z)] \ \text{ is holomorfic in } 1 < |z| \leq \infty \\ &[zB_{n-1}(z)]^{-1}[F_\mu(z)\phi_n(z) + \psi_n(z)] \ \text{ is holomorfic in } 0 \leq |z| < 1 \end{aligned}$$

and

$$\begin{aligned} &[zB_{n-1}(z)][F_\mu(z)\phi_{n*}(z) - \psi_{n*}(z)] \ \text{ is holomorfic in } 1 < |z| \leq \infty \\ &[zB_n(z)]^{-1}[F_\mu(z) - \psi_{n*}(z)/\phi_{n*}(z)] \ \text{ is holomorfic in } 0 \leq |z| < 1. \end{aligned}$$

The ϕ_k correspond to orthogonal rational functions with respect to the Riesz-Herglotz measure of F_μ. They can be obtained by a Gram-Schmidt orthogonalization procedure applied to the sequence $B_0, B_1, \ldots$. If all α_k are zero, the poles are all at infinity and the Szegő polynomials emerge as a special case.

Just as F_μ is a moment generating function by applying the linear functional L to the (formal) expansion of the Riesz-Herglotz kernel, we now have

$$F_\mu(z) = \int_{\mathbb{T}} \left[1 + 2z \sum_{k=1}^{\infty} \frac{\pi_{n-1}^*(z)}{\pi_n^*(t)}\right] \mathrm{d}\mu(t) = m_0 + 2 \sum_{k=1}^{\infty} m_{-k} z \pi_{n-1}^*(z)$$

where the generalized moments are now defined by

$$m_{-k} = \int_{\mathbb{T}} \frac{\mathrm{d}\mu(t)}{\pi_k^*(t)}, \quad \pi_k^*(z) = \prod_{j=1}^{k} (z - \alpha_j).$$

Note that also in this generalized rational case, we can define a linear functional L operating on $\mathcal{L} = \cup_{k=0}^{\infty} \mathcal{L}_k$ via the definition of the moments $L(1) = m_0$ and $L(1/\pi_k^*) = m_{-k}$ for $k = 1, 2, \ldots$. If L is a real functional, then by taking the complex conjugate of the latter and by partial fraction decomposition, it should be clear that the functional is actually defined on the space $\mathcal{L} \cdot \mathcal{L}_*$ where $\mathcal{L}_* = \{f : f_* \in \mathcal{L}\}$. Thus, we can use L to define a complex Hermitian inner product on $\mathcal{L}$ and so the use of the orthogonal rational functions is possible for the solution of the generalized moment problem. The essence of the technique is to note that the quadrature formula whose knots are the zeros $\{\xi_{nk}\}_{k=1}^{n}$ of the para-orthogonal function $Q_n(z, \tau) = \phi_n(z) + \tau \phi_n^*(z)$ (they are all simple and lie on $\mathbb{T}$) and as weights the numbers $1/k_{n-1}(\xi_{nk}, \xi_{nk}) > 0$, then this quadrature formula is exact for all rational functions in $\mathcal{L}_{n-1} \cdot \mathcal{L}_{n-1}$. It then follows that under certain conditions the discrete measure that corresponds to the quadrature formula converges for $n \to \infty$ in a weak sense to a solution of the moment problem. The conditions for this to hold are now involved, not only with the moments, but also with the selection of the sequence of points $\{\alpha_k\}_{k=q}^{\infty} \subset \mathbb{D}$. A typical condition being that $\sum_{k=1}^{\infty}(1 - |\alpha_k|) = \infty$, i.e., the condition that makes the Blaschke product $\prod_{k=1}^{\infty} \zeta_k$ diverge to zero.

2.2 Orthogonal rational functions on the real line

About the same discussion can be given for orthogonal rational functions on the real axis. If however we want the polynomials (which are rational functions with poles at ∞) to come out as a special case, then the natural multipoint generalization is to consider a sequence of points that are all *on* the (extended) real axis $\hat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$. For technical reasons, we have to exclude one point. Without loss of generality, we shall assume this to be the point at infinity. Thus we consider the sequence of points $\{\alpha_k\}_{k=1}^{\infty} \subset \mathbb{R}$ and we define $\pi_n(z) = \prod_{k=1}^{n}(1 - \alpha_k z)$. The spaces of rational functions we shall consider are given by $\mathcal{L}_n = \{p_n/\pi_n : p_n \in \Pi_n\}$. If we define the basis functions $b_0 = 1$, $b_n(z) = z^n/\pi_n(z)$, $k = 1, 2, \ldots$, then orthogonalization of this basis gives the orthogonal rational functions ϕ_n. The inner product can be defined in terms of a positive measure on $\hat{\mathbb{R}}$ via (we assume functions with real coefficients)

$$\langle f, g \rangle = \int_{\mathbb{R}} f(x)g(x)\mathrm{d}\mu(x), \quad f, g \in \mathcal{L},$$

or via some positive linear functional L defined on the space $\mathcal{L} \cdot \mathcal{L}$. Such a linear functional is defined if we know the moments

$$m_{kl} = L(b_k b_l), \quad k, l = 0, 1, \ldots$$

Thus in this case, defining the functional on $\mathcal{L}$ or on $\mathcal{L} \cdot \mathcal{L}$ are two different things. In the first case we only need the moments m_{k0}, in the second case we need a doubly indexed moment sequence. Thus, there are two different moment problems: the one where we look for a representation on $\mathcal{L}$ and the one representing L on $\mathcal{L} \cdot \mathcal{L}$. If all the $\alpha_k = 0$, we get polynomials, and then $\mathcal{L} = \mathcal{L} \cdot \mathcal{L}$ and the two problems are the same. This is the Hamburger moment problem. Also when there is only a finite number of different α_k that are each repeated an infinite number of times, we are in that comfortable situation. An extreme form of the latter is when the only α are 0 and ∞ which leads to (orthogonal) Laurent polynomials, first discussed by Jones, Njåstad and Thron [31].

We also mention here that this (and also the previous) section is related to polynomials orthogonal with respect to varying measures. Indeed if $\phi_n = p_n/\pi_n$, then for $k = 0, 1, \ldots, n-1$

$$0 = \left\langle \phi_n, x^k/\pi_{n-1} \right\rangle = \int_{\mathbb{R}} p_n(x) x^k \mathrm{d}\mu_n(x)$$

where the (in general not positive definite) measure $\mathrm{d}\mu_n(x) = \mathrm{d}\mu(x)/[(1-\alpha_n x)\pi_{n-1}(x)^2]$ depends on n. For polynomials orthogonal w.r.t. varying measures see e.g. [45].

The generalization of the three-term recursion of the polynomials will only exist if some regularity condition holds, namely $p_n(1/\alpha_n) \neq 0$ for all $k = 1, 2, \ldots$. We say that the sequence $\{\phi_n\}$ is regular and it holds then that

$$\phi_n(x) = \left(E_n \frac{x}{1 - \alpha_n z} + B_n \frac{1 - \alpha_{n-1} x}{1 - \alpha_n x} \right) \phi_n(x) - \frac{E_n}{E_{n-1}} \frac{1 - \alpha_{n-2} x}{1 - \alpha_n x} \phi_{n-2}(x)$$

for $n = 1, 2, \ldots$, while the initial conditions are $\phi_{-1} = 0$ and $\phi_0 = 1$. Moreover it holds that $E_n \neq 0$ for all n.

Functions of the second kind can be introduced as in the polynomial case by

$$\psi_n(x) = \int_{\mathbb{R}} \frac{\phi_n(y) - \phi_n(x)}{y - x} \mathrm{d}\mu(y), \quad n = 0, 1, \ldots$$

They also satisfy the same three term recurrence relation with initial conditions $\psi_{-1} = 1$ and $\psi_0 = 0$. The corresponding continued fraction is called a multipoint Padé fraction or MPfraction because its convergents ψ_n/ϕ_n are multipoint Padé approximants of type $[n-1/n]$ to the Stieltjes transform $F_\mu(x) = \int_{\mathbb{R}} (x-y)^{-1} \mathrm{d}\mu(y)$. These rational functions approximate in the sense that for $\alpha \neq 0$ and

$$\lim_{z \to 1/\alpha} \left[\frac{\psi_n(x)}{\phi_n(x)} - F_\mu(x) \right]^{(k)} = 0, \quad k = 0, 1, \ldots, \alpha^{\#} - 1$$

and if $\alpha = 0$ then

$$\lim_{z \to \infty} \left[\frac{\psi_n(x)}{\phi_n(x)} - F_\mu(x) \right] z^{0^{\#}} = 0,$$

where $\alpha \in \{0, \alpha_1, \alpha_1, \ldots, \alpha_{n-1}, \alpha_{n-1}, \alpha_n\}$, and $\alpha^{\#}$ is the multiplicity of α in this set and the limit to $\alpha \in \mathbb{R}$ is nontangential. The MP-fractions are generalizations of the J-fractions to which they are reduced in the polynomial case, i.e., if all the $\alpha_k = 0$.

As for the quadrature formulas, one may consider the rational functions $Q_n(x, \tau) = \phi_n(x) + (1 - \alpha_{n-1}x)/(1 - \alpha_n x) E_n \phi_{n-1}(x)$. If ϕ_n is regular, then except for at most a finite number of $\tau \in \hat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$, these quasi-orthogonal functions have n simple zeros on the real axis that differ from $\{1/\alpha_1, \ldots, 1/\alpha_n\}$. Again, taking these zeros $\{\xi_{nk}\}_{k=1}^n$ as knots and the corresponding weights as $1/k_{n-1}(\xi_{nk}, \xi_{nk}) = 1/\sum_{k=0}^{n-1} |\phi_k(\xi_{nk})|^2 > 0$, we get quadrature formulas that are exact for all rational functions in $\mathcal{L}_{n-1} \cdot \mathcal{L}_{n-1}$. If ϕ_n is regular and $\tau = 0$ is not one of those exceptional values for τ, then the formula is even exact in $\mathcal{L}_n \cdot \mathcal{L}_{n-1}$. Since an orthogonal polynomial sequence is always regular and since there are no exceptional values for τ, one can thus always take the zeros of ϕ_n for the construction of the quadrature formula, so that we are back in the case of Gauss quadrature formulas.

These quadrature formulas, apart from being of practical interest, can be used to find a solution for the moment problem in $\mathcal{L}$. Note that we use orthogonality, thus an inner product so that for the solution of the moment problem in $\mathcal{L}$, we need the linear functional L to be defined on $\mathcal{L} \cdot \mathcal{L}$. It is not known how the problem could be solved using only the moments defining L on $\mathcal{L}$.

2.3 Orthogonal rational functions on an interval

Of course, many of the classical orthogonal polynomials are not defined with respect to a measure on the unit circle or the whole real line, but they are orthogonal over a finite interval or a half-line.

Not much is known about the generalization of these cases to the rational case. There is a high potential in there because the analysis of orthogonal rational functions on the real line suffered from technical difficulties because the poles of the function spaces were in the support of the measure. If the support of the measure is only a finite interval or a half-line, we could easily locate the poles on the real axis, but outside the support of the measure. New intriguing questions about the location of the zeros, the quadrature formulas, the moment problems arise. For further details on this topic we refer to [59,56,57,55,58].

3 Homogeneous orthogonal polynomials

In the presentation of one of the multivariate generalizations of the concept of orthogonal polynomials, we follow the outline of Section 1. An inner product or linear functional is defined, orthogonality relations are imposed on multivariate functions of a specific form, 3-term recurrence relations come into play and some properties of the zeroes of these multivariate orthogonal polynomials are presented. The 3-term recurrence relations link the polynomials to rational approximants and continued fractions. The zero properties allow the development of some new cubature rules.

Without loss of generality we present the results only for the bivariate case.

3.1 Orthogonality conditions

Let us first introduce some notation. We denote by $\mathbb{C}[z]$ the linear space of polynomials in the variable z with complex coefficients, by $\mathbb{C}[\lambda_1, \lambda_2]$ the linear space of bivariate polynomials in λ_1 and λ_2 with complex coefficients, by $\mathbb{C}(\lambda_1, \lambda_2)$ the commutative field of rational functions in λ_1 and λ_2 with complex coefficients, by $\mathbb{C}(\lambda_1, \lambda_2)[z]$ the linear space of polynomials in the variable z with coefficients from $\mathbb{C}(\lambda_1, \lambda_2)$ and by $\mathbb{C}[\lambda_1, \lambda_2][z]$ the linear space of polynomials in the variable z with coefficients from $\mathbb{C}[\lambda_1, \lambda_2]$.

In dealing with bivariate polynomials and functions we shall often switch between the Cartesian and the spherical coordinate system. We introduce the directional vector $\lambda = (\lambda_1, \lambda_2) \in \mathbb{C}^2$, satisfying $||\lambda|| = 1$ and hence belonging to the unit sphere S_2, and write $(x, y) = (\lambda_1 z, \lambda_2 z)$ with $x, y, z \in \mathbb{C}$. This new view on the multivariate problem in which the Cartesian coordinates (x, y) are replaced by the coordinates $\lambda = (\lambda_1, \lambda_2)$ and z, with $||\lambda|| = 1$, will turn out to be a powerful tool in the sequel of the text.

We define a bivariate function $V_m(x, y)$ of the form

$$V_m(x, y) = \mathcal{V}_m(z) = \sum_{i=0}^{m} B_{\Delta_m+m-i}(\lambda) z^i \ , \tag{3.1}$$

$$B_{\Delta_m+m-i}(\lambda) = \sum_{j=0}^{\Delta_m+m-i} b_{\Delta_m+m-i-j,j} \lambda_1^{\Delta_m+m-i-j} \lambda_2^j \ . \tag{3.2}$$

The role of the degree shift Δ_m will become apparent later on. The function $V_m(x, y)$ is a polynomial of degree m in z with polynomial coefficients from $\mathbb{C}[\lambda_1, \lambda_2]$. The coefficients $B_{\Delta_m}, \ldots, B_{\Delta_m+m}$ are homogeneous polynomials in the parameters λ_1 and λ_2. The function $V_m(x, y)$ does itself not belong to $\mathbb{C}[x, y]$ but since $V_m(x, y) = \mathcal{V}_m(z)$, it belongs to $\mathbb{C}[\lambda_1, \lambda_2][z]$.

The form (3.1) has been chosen because, remarkably enough, the bivariate function

$$\tilde{V}_m(x,y) = \tilde{\mathcal{V}}_m(z) = z^{\Delta_m+m}\mathcal{V}_m(z^{-1})$$

belongs to $\mathbb{C}[x,y]$, which proves to be useful later on.

We also introduce the linear functional Γ acting on the variable z, as

$$\Gamma(z^i) = c_i(\lambda)$$

where $c_i(\lambda)$ is a homogeneous expression of degree i in λ_1 and λ_2:

$$c_i(\lambda) = \sum_{j=0}^{i} c_{i-j,j}\lambda_1^{i-j}\lambda_2^{j} \, . \tag{3.3}$$

We then impose the orthogonality conditions

$$\Gamma(z^i\mathcal{V}_m(z)) = 0 \qquad i = 0,\ldots,m-1 \, . \tag{3.4}$$

As in the univariate case the orthogonality conditions (3.4) only determine $\mathcal{V}_m(z)$ up to a kind of normalization: $m+1$ polynomial coefficients $B_{\Delta_m+m-i}(\lambda)$ must be determined from the m parameterized conditions (3.4). How this is solved, is explained below.

With the $c_i(\lambda)$ we define the polynomial Hankel determinants

$$H_m(\lambda) = \begin{vmatrix} c_0(\lambda) & \cdots & c_{m-2}(\lambda) & c_{m-1}(\lambda) \\ \vdots & \iddots & \iddots & c_m(\lambda) \\ c_{m-2}(\lambda) & \iddots & \iddots & \vdots \\ c_{m-1}(\lambda) & c_m(\lambda) & \cdots & c_{2m-2}(\lambda) \end{vmatrix}, \qquad H_0(\lambda) = 1.$$

We call the functional Γ definite if

$$H_m(\lambda) \not\equiv 0 \qquad m \geq 0.$$

In the sequel of the text we assume that $\mathcal{V}_m(z)$ satisfies (3.4) and that Γ is a definite functional. Also we shall assume that $\mathcal{V}_m(z)$ as given by (3.1) is primitive, meaning that its polynomial coefficients $B_{\Delta_m+m-i}(\lambda)$ are relatively prime. This last condition can always be satisfied, because for a definite functional Γ a solution of (3.4) is given by [6]

$$\mathcal{V}_m(z) = \frac{1}{p_m(\lambda)} \begin{vmatrix} c_0(\lambda) & \cdots & c_{m-1}(\lambda) & c_m(\lambda) \\ \vdots & \iddots & & \vdots \\ c_{m-1}(\lambda) & & \cdots & c_{2m-1}(\lambda) \\ 1 & z & \cdots & z^m \end{vmatrix}, \qquad \mathcal{V}_0(z) = 1 \tag{3.5}$$

where the polynomial $p_m(\lambda)$ is a polynomial greatest common divisor of the polynomial coefficients of the powers of z in this determinant expression. Clearly (3.4) completely determines $\mathcal{V}_m(z)$ and consequently $V_m(x,y)$.

The associated polynomials $W_m(x,y)$ defined by

$$W_m(x,y) = \mathcal{W}_m(z) = \Gamma\left(\frac{\mathcal{V}_m(z) - \mathcal{V}_m(u)}{z-u}\right) \tag{3.6}$$

are of the form

$$W_m(x,y) = \mathcal{W}_m(z) = \sum_{i=0}^{m-1} A_{\Delta_m+m-1-i}(\lambda) z^i \tag{3.7}$$

$$A_{\Delta_m+m-1-i}(\lambda) = \sum_{j=0}^{m-1-i} B_{\Delta_m+m-1-i-j}(\lambda) c_j(\lambda). \tag{3.8}$$

The expression $A_{\Delta_m+m-1-i}(\lambda)$ is a homogeneous polynomial of degree $\Delta_m + m - 1 - i$ in the parameters λ_1 and λ_2. Note again that $W_m(x,y)$ does not necessarily belong to $\mathbb{C}[x,y]$ because the homogeneous degree in λ_1 and λ_2 doesn't equal the degree in z. Instead it belongs to $\mathbb{C}[\lambda_1,\lambda_2][z]$. On the other hand, the function

$$\tilde{W}_m(x,y) = \tilde{\mathcal{W}}_m(z) = z^{\Delta_m+m-1}\mathcal{W}_m(z^{-1})$$

belongs to $\mathbb{C}[x,y]$.

3.2 Recurrence relations

In the sequel of the text we use both the notations $V_m(x,y)$ and $\mathcal{V}_m(z)$ interchangeably to refer to (3.1), and analogously for $W_m(x,y)$ and $\mathcal{W}_m(z)$ in (3.7). For simplicity, we also refer to both $V_m(x,y)$ and $\mathcal{V}_m(z)$ as polynomials, and similarly for $W_m(x,y)$ and $\mathcal{W}_m(z)$.

The link between the orthogonal polynomials $V_m(x,y)$, the associated polynomials $W_m(x,y)$ and rational approximation theory is obvious from the following. This relationship also makes it easy to deduce a number of recurrence relations, the proofs of which can be found in [6].

Assume that, from Γ, we construct the bivariate series expansion

$$f(x,y) = \sum_{i,j=0}^{\infty} c_{ij} x^i y^j = \sum_{i,j=0}^{\infty} c_{ij} \lambda_1^i \lambda_2^j z^{i+j} = \sum_{i=0}^{\infty} c_i(\lambda) z^i.$$

Then, with $\Delta_m = (m-1)m$, the polynomials

$$\tilde{V}_m(x,y) = \tilde{\mathcal{V}}_m(z) = z^{\Delta_m+m}\mathcal{V}_m(z^{-1})$$

$$
\begin{aligned}
&= \sum_{i=0}^{m} B_{\Delta_m+i}(\lambda) z^{\Delta_m+i} \\
&= \sum_{i=0}^{m} \sum_{j=0}^{\Delta_m+i} b_{\Delta_m+i-j,j} x^{\Delta_m+i-j} y^j
\end{aligned}
$$

and

$$
\begin{aligned}
\tilde{W}_m(x,y) = \tilde{\mathcal{W}}_m(z) &= z^{\Delta_m+m-1}\mathcal{W}_m(z^{-1}) \\
&= \sum_{i=0}^{m-1} A_{\Delta_m+i}(\lambda) z^{\Delta_m+i} \\
&= \sum_{i=0}^{m-1} \sum_{j=0}^{\Delta_m+i} a_{\Delta_m+i-j,j} x^{\Delta_m+i-j} y^j
\end{aligned}
$$

satisfy the Padé approximation conditions

$$
\begin{aligned}
\left(f\tilde{V}_m - \tilde{W}_m\right)(x,y) &= \left(f\tilde{\mathcal{V}}_m - \tilde{\mathcal{W}}_m\right)(z) \\
&= \sum_{i=\Delta_m+2m}^{\infty} d_i(\lambda) z^i \\
&= \sum_{i=\Delta_m+2m}^{\infty} \left(\sum_{j=0}^{i} d_{i-j,j} x^{i-j} y^j \right)
\end{aligned}
$$

where, as in (3.2), (3.3) and (3.8), the subscripted function $d_i(\lambda)$ is a homogeneous function of degree i in λ_1 and λ_2. The rational function $\tilde{W}_m(x,y)/\tilde{V}_m(x,y)$ is called the homogeneous Padé approximant for $f(x,y)$. More information about these approximants can be found in [20]. It is now easy to give a three-term recurrence relation for the $V_m(x,y)$ and the associated functions $W_m(x,y)$, as well as an identity linking the $V_m(x,y)$ and the $W_m(x,y)$.

Theorem 3.1 *Let the functional Γ be definite and let the polynomials $\mathcal{V}_m(z)$ and $p_m(\lambda)$ be defined as in (3.5). Then the polynomial sequences $\{\mathcal{V}_m(z)\}_m$ and $\{\mathcal{W}_m(z)\}_m$ satisfy the recurrence relations*

$$
\begin{aligned}
V_{m+1}(x,y) &= \alpha_{m+1}(\lambda)\left((z-\beta_{m+1}(\lambda))V_m(x,y) - \gamma_{m+1}(\lambda)V_{m-1}(x,y)\right), \\
&\quad V_{-1}(x,y) = 0, \qquad V_0(x,y) = 1 \\
W_{m+1}(x,y) &= \alpha_{m+1}(\lambda)\left((z-\beta_{m+1}(\lambda))W_m(x,y) - \gamma_{m+1}(\lambda)W_{m-1}(x,y)\right), \\
&\quad W_{-1}(x,y) = -1, \qquad W_0(x,y) = 0
\end{aligned}
$$

with

$$
\begin{aligned}
\alpha_{m+1}(\lambda) &= \frac{p_m(\lambda)}{p_{m+1}(\lambda)} \frac{H_{m+1}(\lambda)}{H_m(\lambda)}, \\
\beta_{m+1}(\lambda) &= \frac{\Gamma\left(z\left[V_m(x,y)\right]^2\right)}{\Gamma\left(\left[V_m(x,y)\right]^2\right)},
\end{aligned}
$$

$$\gamma_{m+1}(\lambda) = \frac{p_{m-1}(\lambda)}{p_m(\lambda)} \frac{H_{m+1}(\lambda)}{H_m(\lambda)}, \qquad \gamma_1(\lambda) = c_0(\lambda).$$

Theorem 3.2 *Let the functional* Γ *be definite and let the polynomial sequences* $\mathcal{V}_m(z)$ *and* $p_m(\lambda)$ *be defined as in (3.5). Then the polynomials* $\mathcal{V}_m(z)$ *and* $\mathcal{W}_m(z)$ *satisfy the identity*

$$\begin{aligned}\mathcal{V}_m(z)\mathcal{W}_{m+1}(z) - \mathcal{W}_m(z)\mathcal{V}_{m+1}(z) &= V_m(x,y)W_{m+1}(x,y) - W_m(x,y)V_{m+1}(x,y) \\ &= \frac{[H_{m+1}(\lambda)]^2}{p_m(\lambda)p_{m+1}(\lambda)}.\end{aligned}$$

The preceding theorem shows that the expression

$$\mathcal{V}_m(z)\mathcal{W}_{m+1}(z) - \mathcal{W}_m(z)\mathcal{V}_{m+1}(z)$$

is independent of z and homogeneous in λ_1 and λ_2. If $p_m(\lambda)$ and $p_{m+1}(\lambda)$ are constants, this homogeneous expression is of degree $2m(m+1)$.

3.3 Factorization of the homogeneous orthogonal polynomials

Let us now show that for a definite functional Γ the orthogonal polynomials $V_m(x,y)$ and $V_{m+1}(x,y)$ have no common factors. The same holds for the associated polynomials $W_m(x,y)$ and $W_{m+1}(x,y)$ and for the polynomials $V_m(x,y)$ and $W_m(x,y)$. The proofs of these results can be found in [6].

We take a closer look at the factorization of the orthogonal polynomials $\mathcal{V}_m(z)$ and their associated polynomials $\mathcal{W}_m(z)$ in irreducible factors. This factorization is unique in $\mathbb{C}[\lambda_1,\lambda_2][z]$ except for multiplicative constants from $\mathbb{C}$ which are the unit multiples in $\mathbb{C}[\lambda_1,\lambda_2]$ and except for the order of the factors. This is because $\mathbb{C}[\lambda_1,\lambda_2][z]$ is a unique factorization domain.

Theorem 3.3 *Let the functional* Γ *be definite and let the polynomials* $\mathcal{V}_m(z)$ *and* $p_m(\lambda)$ *be defined as in (3.5). Let* $\mathcal{W}_m(z)$ *be given by (3.6). Then*

- $\mathcal{V}_m(z)$ *and* $\mathcal{V}_{m+1}(z)$ *have no common factor,*
- $\mathcal{W}_m(z)$ *and* $\mathcal{W}_{m+1}(z)$ *have no common factor,*
- $\mathcal{V}_m(z)$ *and* $\mathcal{W}_m(z)$ *have no common factor.*

If the variables x and y are real and the functional Γ is positive definite, then even more can be said.

Theorem 3.4 *For a positive definite functional* Γ*, the polynomials* $\mathcal{V}_m(z)$ *satisfying (3.4) have no irreducible factors in* $\mathbb{R}[\lambda_1,\lambda_2][z]$ *of multiplicity larger than 1.*

We illustrate this by considering the following positive definite functional

$$
\begin{aligned}
\Gamma(z^i) &= c_i(\lambda) = \sum_{j=0}^{i} c_{i-j,j}\lambda_1^{i-j}\lambda_2^j \ , \\
c_{i-j,j} &= \binom{i}{j} \iint\limits_{||(x,y)||\leq 1} x^{i-j}y^j dxdy \ . \qquad (3.9)
\end{aligned}
$$

For the ℓ_2-norm the expressions $c_i(\lambda)$ equal

$$
c_i(\lambda) = \iint\limits_{||(x,y)||\leq 1} (x\lambda_1 + y\lambda_2)^i \ dx \ dy
$$

which are zero for odd i and respectively π, $\frac{\pi}{4}(\lambda_1^2+\lambda_2^2)$, $\frac{\pi}{8}(\lambda_1^2+\lambda_2^2)^2$, $\frac{5\pi}{64}(\lambda_1^2+\lambda_2^2)^3$, $\frac{7\pi}{128}(\lambda_1^2+\lambda_2^2)^4$ for $i = 0, 2, 4, 6, 8$.

With $V_m(x,y)$ given by (3.1), the orthogonality conditions (3.4) then translate to

$$
\begin{aligned}
\Gamma(z^i\mathcal{V}_m(z)) &= \sum_{k=0}^{m} B_{\Delta_m+m-k}(\lambda)\Gamma(z^{i+k}) \\
&= \iint\limits_{||(x,y)||\leq 1} \sum_{k=0}^{m} B_{\Delta_m+m-k}(\lambda) \left(r\lambda_1 + s\lambda_2\right)^{i+k} \ dr \ ds \qquad i = 0,\ldots,m-1 \\
&= \iint\limits_{||(x,y)||\leq 1} \left(r\lambda_1 + s\lambda_2\right)^i \mathcal{V}_m(r\lambda_1 + s\lambda_2) \ dr \ ds = 0 \qquad i = 0,\ldots,m-1 \ .
\end{aligned}
$$

When introducing a signed distance function

$$
\mathrm{sd}(x,y) = \mathrm{sgn}(x)||(x,y)||
$$

the first few orthogonal polynomials satisfying (3.4), having only simple irreducible factors, can be written as (we use the notation $\mathcal{V}_m(z)$ to designate both $V_m(x,y)$ and $\mathcal{V}_m(z)$):

$$
\begin{aligned}
\mathcal{V}_0(z) &= 1 \\
\mathcal{V}_1(z) &= z = \mathrm{sd}(x,y) \\
\mathcal{V}_2(z) &= z^2 - \tfrac{1}{4}\left(\lambda_1^2+\lambda_2^2\right) \qquad (3.10) \\
&= \left(\mathrm{sd}(x,y) - \tfrac{1}{2}\right)\left(\mathrm{sd}(x,y) + \tfrac{1}{2}\right) \\
\mathcal{V}_3(z) &= z^3 - \tfrac{1}{2}\left(\lambda_1^2+\lambda_2^2\right) z \\
&= \mathrm{sd}(x,y)\left(\mathrm{sd}(x,y) - \tfrac{1}{\sqrt{2}}\right)\left(\mathrm{sd}(x,y) + \tfrac{1}{\sqrt{2}}\right) \\
\mathcal{V}_4(z) &= z^4 - \tfrac{3}{4}\left(\lambda_1^2+\lambda_2^2\right) z^2 + \tfrac{1}{16}\left(\lambda_1^2+\lambda_2^2\right)^2 \\
&= \left(\mathrm{sd}(x,y) - \tfrac{\sqrt{3-\sqrt{5}}}{2\sqrt{2}}\right)\left(\mathrm{sd}(x,y) + \tfrac{\sqrt{3-\sqrt{5}}}{2\sqrt{2}}\right)\left(\mathrm{sd}(x,y) - \tfrac{\sqrt{3+\sqrt{5}}}{2\sqrt{2}}\right)\left(\mathrm{sd}(x,y) + \tfrac{\sqrt{3+\sqrt{5}}}{2\sqrt{2}}\right) \\
\mathcal{V}_5(z) &= z^5 - \left(\lambda_1^2+\lambda_2^2\right) z^3 + \tfrac{3}{16}\left(\lambda_1^2+\lambda_2^2\right)^2 z
\end{aligned}
$$

$$= \mathrm{sd}(x,y)\left(\mathrm{sd}(x,y)-\tfrac{1}{2}\right)\left(\mathrm{sd}(x,y)+\tfrac{1}{2}\right)\left(\mathrm{sd}(x,y)-\tfrac{\sqrt{3}}{2}\right)\left(\mathrm{sd}(x,y)+\tfrac{\sqrt{3}}{2}\right).$$

Let us now fix $\lambda = \lambda^*$ and take a look at the projected functions

$$\begin{aligned} f_{\lambda^*}(z) &= f(\lambda_1^* z, \lambda_2^* z), \\ \mathcal{V}_{m,\lambda^*}(z) &= V_m(\lambda_1^* z, \lambda_2^* z), \\ \mathcal{W}_{m,\lambda^*}(z) &= W_m(\lambda_1^* z, \lambda_2^* z). \end{aligned}$$

If we introduce the linear functional c^* acting on the variable z, as

$$c^*(z^i) = c_i(\lambda^*) = \Gamma(z^i)\mid_{\lambda=\lambda^*} \tag{3.11}$$

then we can prove the following projection property of the $V_m(x,y)$.

In the following result we use the notation $V_m(z)$ to denote the univariate polynomials of degree m orthogonal with respect to the linear functional c^*. The reader should not confuse these polynomials with the $\mathcal{V}_m(z)$ or the $V_m(x,y)$.

Theorem 3.5 *Let the monic univariate polynomials $V_m(z)$ satisfy the orthogonality conditions*

$$c^*(z^i V_m(z)) = 0 \qquad i = 0, \ldots, m-1 .$$

with c^* given by (3.11), and let the polynomials $V_m(x,y) = \mathcal{V}_m(z)$ satisfy the orthogonality conditions (3.4). Then

$$\begin{aligned} H_m(\lambda_1^*, \lambda_2^*) V_m(z) &= p_m(\lambda_1^*, \lambda_2^*) \mathcal{V}_{m,\lambda^*}(z) \\ &= p_m(\lambda_1^*, \lambda_2^*) V_m(\lambda_1^* z, \lambda_2^* z). \end{aligned}$$

If the functional Γ is positive definite, then the zeroes $z_i^{(m)}(\lambda^*)$ of $\mathcal{V}_{m,\lambda^*}(z)$ are real and simple because the functional c^* is positive definite. Then according to the implicit function theorem, there exists for each $z_i^{(m)}(\lambda^*)$ a unique holomorphic function $\phi_i^{(m)}(\lambda_1^*, \lambda_2^*)$ such that in a neighborhood of $z_i^{(m)}(\lambda^*)$,

$$\mathcal{V}_m(z) = 0 \Longleftrightarrow z = \phi_i^{(m)}(\lambda_1^*, \lambda_2^*).$$

Since this is true for each $\lambda = \lambda^*$ because Γ is positive definite, this implies that for each $i = 1, \ldots, m$ the zeroes $z_i^{(m)}$ can be viewed as a holomorphic function of λ, namely $z_i^{(m)} = \phi_i^{(m)}(\lambda_1, \lambda_2)$. Let us denote

$$A_i^{(m)}(\lambda) = \frac{\mathcal{W}_{m,\lambda}(z_i^{(m)})}{\mathcal{V}'_{m,\lambda}(z_i^{(m)})} = \frac{\mathcal{W}_m(\phi_i^{(m)}(\lambda))}{\mathcal{V}'_m(\phi_i^{(m)}(\lambda))}.$$

Then the following cubature formula can rightfully be called a Gaussian cubature formula. The proof of this fact can be found in [5].

Theorem 3.6 *Let $\mathcal{P}(z)$ be a polynomial of degree $2m-1$ belonging to $\mathbb{C}(\lambda_1,\lambda_2)[z]$, the set of polynomials in the variable z with coefficients from the space of bivariate rational functions in λ_1 and λ_2 with complex coefficients. Let the functions $\phi_i^{(m)}(\lambda_1,\lambda_2)$ be given as above and be such that*

$$\forall\lambda\in S_2 : j\neq i \Longrightarrow \phi_j^{(m)}(\lambda)\neq\phi_i^{(m)}(\lambda).$$

Then for $z=\lambda_1 x+\lambda_2 y$ holds

$$\iint_{||(x,y)||\leq 1} \mathcal{P}(\lambda_1 x+\lambda_2 y)dxdy = \sum_{i=1}^{m} A_i^{(m)}(\lambda)\mathcal{P}(\phi_i^{(m)}(\lambda)).$$

Let us illustrate Theorem 3.6 with an example. Take

$$\mathcal{P}(z)=\mathcal{P}(\lambda_1 x+\lambda_2 y)=\sum_{i=0}^{3}\binom{3}{i}\left(\frac{\lambda_1}{\lambda_2}\right)^{3-i}(\lambda_1 x+\lambda_2 y)^i$$

and consider again the ℓ_2-norm. Then

$$\iint_{||(x,y)||\leq 1} \mathcal{P}(\lambda_1 x+\lambda_2 y)dx\,dy = \frac{\pi\lambda_1}{4\lambda_2^3}(4\lambda_1^2+3\lambda_1^2\lambda_2^2+3\lambda_2^4). \tag{3.12}$$

The exact integration rule given in Theorem 3.6, applies to (3.12) with $m=2$. From the orthogonal function $V_2(x,y)=\mathcal{V}_2(z)$ given in (3.10), we obtain the zeroes

$$\phi_1^{(2)}(\lambda) = \frac{1}{2}\sqrt{\lambda_1^2+\lambda_2^2}$$
$$\phi_2^{(2)}(\lambda) = -\frac{1}{2}\sqrt{\lambda_1^2+\lambda_2^2}$$

and the weights

$$A_1^{(2)}(\lambda)=A_2^{(2)}(\lambda)=\frac{\pi}{2}\,.$$

The integration rule

$$A_1^{(2)}\mathcal{P}(\phi_1^{(2)}(\lambda))+A_2^{(2)}\mathcal{P}(\phi_2^{(2)}(\lambda))$$

then yields the same result as (3.12). In fact, the Gaussian m-point cubature formula given in Theorem 3.6 exactly integrates a parameterized family of polynomials, over a domain in $\mathbb{R}^2$, or more generally $\mathbb{R}^n$. The m nodes and weights are themselves functions of the parameters λ_1 and λ_2.

For the ℓ_1- and ℓ_∞-norm similar computations can be performed: after obtaining the $c_i(\lambda)$ for these norms, the orthogonal polynomial $\mathcal{V}_2(z)$ constructed from the $c_i(\lambda)$ delivers all necessary ingredients for the application of the Gaussian cubature rule.

More properties of the homogeneous orthogonal polynomials $V_m(x, y)$ can be proved, such as the fact that they are the characteristic polynomials of certain parametrized tridiagonal matrices [7]. The connection between their theory and the theory of the univariate orthogonal polynomials is very close.

4 Vector and matrix orthogonal polynomials

In this section, we generalize some results of Section 1 on scalar orthogonal polynomials to the vector and matrix case.

Let Π^α be the space of all vector polynomials with α components. Let $\Pi^\alpha_{\vec{n}}$ be the subspace of Π^α of all vector polynomials of degree (elementwise) at most $\vec{n} \in \mathbb{N}^\alpha$. The dimension of this subspace is

$$|\vec{n}| = \sum_{i=1}^{\alpha}(n_i + 1) \qquad \text{with} \qquad \vec{n} = (n_1, n_2, \ldots, n_\alpha).$$

Following the notation of Section 1, we denote a set of basis functions for $\Pi^\alpha_{\vec{n}}$ as

$$\{B_1, B_2, \ldots, B_{|\vec{n}|}\}.$$

In contrast to the scalar case, a nested basis of increasing degree can be chosen in several different ways, e.g., with $\alpha = 2$, a natural choice could be

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} x \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ x \end{bmatrix}, \ldots. \tag{4.1}$$

Another possibility is

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} x \\ 0 \end{bmatrix}, \begin{bmatrix} x^2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} x^3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ x \end{bmatrix}, \ldots.$$

Once, we have chosen a (nested) basis in $\Pi^\alpha_{\vec{n}}$, each element of $\Pi^\alpha_{\vec{n}}$ can be identified by an element of $\mathbb{C}^{|\vec{n}|\times 1}$. Similarly, choosing a basis in the dual space, each linear functional on $\Pi^\alpha_{\vec{n}}$ can be represented by an element of $\mathbb{C}^{1\times|\vec{n}|}$.

Let μ be a matrix-valued measure of a finite or infinite interval $\mathbb{I}$ on the real line. Then, the components of

$$L_k(P) = \int_{\mathbb{I}} x^k d\mu(x) P(x)$$

can be considered as the duals of the vector polynomials

$$\begin{bmatrix} x^k \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ x^k \\ \vdots \\ 0 \end{bmatrix}, \ldots$$

The corresponding inner product for two vector polynomials P and Q is introduced as follows

$$\langle Q, P\rangle = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} q_k^T \langle x^k I_\alpha, x^l I_\alpha\rangle p_l = \sum_{k=0}^{\infty} q_k^T L_l(P)$$

with $Q(x) = \sum_{k=0}^{\infty} q_k x^k$ and $P(x) = \sum_{k=0}^{\infty} p_k x^k$. When we consider the natural nested basis (4.1), the moment matrix is block Hankel and all blocks are completely determined by the matrix-valued function L defined as

$$L(x^i) = \int_{\mathbb{I}} x^i d\mu(x)$$

because the (k, l)th block of the moment matrix equals

$$L_k(x^l) = L(x^{k+l}).$$

In a similar way, we can extend the results for scalar polynomials orthogonal on the unit circle into vector orthogonal polynomials where, then, the moment matrix has a block Toeplitz structure.

Taking the natural nested basis, and taking the vector orthogonal polynomials together in groups of α elements, we derive $\alpha\times\alpha$ matrix orthogonal polynomials $\hat{P}_i$, $i = 0, 1, \ldots$ of increasing degree i satisfying the "matrix" orthogonality relationship

$$\left\langle \hat{P}_i, \hat{P}_j\right\rangle = \delta_{ij} I_\alpha$$

with $\langle\cdot,\cdot\rangle$ defined in an obvious way based on the inner product of vector polynomials. Several other properties of Section 1 can be generalized in the same way for vector and matrix orthogonal polynomials [44,42,43].

Let us consider the following discrete inner product based on the points $z_i \in \mathbb{C}$, $i = 1, 2, \ldots, N$ and the weights (vectors) $F_i \in \mathbb{C}^{\alpha\times 1}$:

$$\langle V, U\rangle = \sum_{i=1}^{N} V(z_i)^H F_i F_i^H U(z_i), \qquad \text{with} \qquad U, V \in \Pi_{\vec{n}}^{\alpha}.$$

Note that this is a true inner product as long as there is no element U from $\Pi_{\vec{n}}^{\alpha}$ such that $\langle U, U\rangle = 0$. To find a recurrence relation for the vector orthogonal polynomials based on the natural nested basis for $\Pi_{\vec{n}}^{\alpha}$, we can solve the following inverse eigenvalue problem. Given z_i, F_i, $i = 1, 2, \ldots, N$, find the upper triangular matrix R and the generalized Hessenberg matrix H such that

$$\left[Q^H F | Q^H \Lambda_z Q\right] = \left[\overline{R}|H\right], \tag{4.2}$$

where the right-hand side matrix has upper triangular structure, the rows of the matrix F are the weights F_i^H, the matrix Q is a unitary $N\times N$ matrix, Λ_z is the diagonal matrix with the points z_i on the diagonal, and $\overline{R}$ is a $N\times\alpha$ matrix which is zero except for the upper $\alpha\times\alpha$ block which is the upper triangular matrix R. Note that because $\overline{H} = \left[\overline{R}|H\right]$ has the upper triangular structure, H is a generalized Hessenberg matrix having α subdiagonals different from zero. Instead of the natural nested basis, we can take a more complicated nested basis. In this case the matrix $\left[\overline{R}|H\right]$ will still have the upper triangular structure, but only after a column permutation.

The columns of the unitary matrix Q are connected to the values of the corresponding vector orthogonal polynomials $\phi_1, \phi_2, \ldots$ as follows

$$Q_{ij} = F_i^H \phi_j(z_i), \qquad \text{with} \qquad i,j = 1,2,\ldots,N.$$

Because the relation (4.2) gives us a recurrence for the columns of Q, we get the corresponding recurrence relation for the vector orthogonal polynomials ϕ_i:

$$\begin{aligned} h_{ii}\phi_i(z) &= e_i - \sum_{j=1}^{i-1} h_{ji} h_{ji} \phi_j(z), \quad i = 1,2,\ldots,\alpha \\ &= z\phi_{i-\alpha} - \sum_{j=1}^{i-1} h_{ji} h_{ji} \phi_j(z), \quad i = \alpha+1, \alpha+2, \ldots, N \end{aligned}$$

where h_{ij} is the (i,j)th element of the upper triangular (rectangular) matrix $\overline{H}$.

For z_i arbitrary chosen in the complex plane, the previous inverse eigenvalue problem requires $O(N^3)$ floating point operations. However, this computational complexity decreases by an order of magnitude in the following two special cases.

(1) *All the points z_i are real and the weights are real vectors*

In this case, all computations can be done using real numbers. Hence, the matrix Q will also be real (orthogonal). Therefore, $H = Q^T Z Q$ will be symmetric and because H is a generalized Hessenberg, it will be a symmetric banded matrix with bandwidth $2\alpha + 1$. Note that the recurrence relation for the vector orthogonal polynomials only involves $2\alpha + 1$ of these polynomials, i.e., for the special case of $\alpha = 1$, we obtain the classical 3-term recurrence relation.

(2) *All the points z_i are on the unit circle*

In this case, H is not only generalized Hessenberg but also unitary. In this case, the matrix H can be written as a product of more simple unitary matrices G_i:

$$H = G_1 G_2 \cdots G_{N-\alpha}$$

where $G_i = I_{i-1} \oplus Q_i \oplus I_{N-i-\alpha-1}$ with Q_i an $\alpha \times \alpha$ unitary matrix. When the inverse eigenvalue problem is solved where H is parameterized in terms of the unitary matrices Q_i, the computational complexity reduces to $O(N^2)$. The recurrence relation for the vector orthonormal polynomials turns out to be a generalization of the classical Szegő relation.

For more details on vector orthogonal polynomials with respect to a discrete inner product, we refer the interested reader to [17,51,53]. These vector and/or matrix orthogonal polynomials can be applied in system identification [40,18], to design fast and accurate algorithms to solve structured systems [52,54].

5 Multiple orthogonality and Hermite-Padé approximation

Hermite-Padé approximation is simultaneous rational approximation to a vector of r functions $f_1, f_2, \ldots, f_r$, which are all given as Taylor series around a point $a \in \mathbb{C}$ and for which we require interpolation conditions at a. We will restrict our attention to Hermite-Padé approximation around infinity and impose interpolation conditions at infinity. Certain polynomials which appear in this rational approximation problem satisfy a number of orthogonality conditions with respect to r measures and hence we call them *multiple orthogonal polynomials*. These polynomials are one-variable polynomials but the degree is a multi-index. A good source for information on Hermite-Padé approximation is the book by Nikishin and Sorokin [38, Chapter 4], where the multiple orthogonal polynomials are called polyorthogonal polynomials. Other good sources of information are the surveys by Aptekarev [2] and de Bruin [21].

Suppose we are given r functions with Laurent expansions

$$f_j(z) = \sum_{k=0}^{\infty} \frac{c_{k,j}}{z^{k+1}}, \qquad j = 1, 2, \ldots, r.$$

There are basically two different types of Hermite-Padé approximation. First we will need multi-indices $\vec{n} = (n_1, n_2, \ldots, n_r) \in \mathbb{N}^r$ and their size $|\vec{n}| = n_1 + n_2 + \cdots + n_r$.

Definition 5.1 (Type I) *Type I Hermite-Padé approximation to the vector $(f_1, \ldots, f_r)$ near infinity consists of finding a vector of polynomials $(A_{\vec{n},1}, \ldots, A_{\vec{n},r})$ and a polynomial $B_{\vec{n}}$, with $A_{\vec{n},j}$ of degree $\leq n_j - 1$, such that*

$$\sum_{j=1}^{r} A_{\vec{n},j}(z) f_j(z) - B_{\vec{n}}(z) = \mathcal{O}\left(\frac{1}{z^{|\vec{n}|}}\right), \qquad z \to \infty. \tag{5.1}$$

In type I Hermite-Padé approximation one wants to approximate a linear combination (with polynomial coefficients) of the r functions by a polynomial. This is often done for the vector of functions $f, f^2, \ldots, f^r$, where f is a given function. The solution of the equation

$$\sum_{j=1}^{r} A_{\vec{n},j}(z) \hat{f}^j(z) - B_{\vec{n}}(z) = 0$$

is an algebraic function and then gives an algebraic approximant $\hat{f}$ for the function f.

Definition 5.2 (Type II) *Type II Hermite-Padé approximation to the vector $(f_1, \ldots, f_r)$ near infinity consists of finding a polynomial $P_{\vec{n}}$ of degree $\leq |\vec{n}|$ and polynomials $Q_{\vec{n},j}$ $(j = 1, 2, \ldots, r)$ such that*

$$\begin{aligned} P_{\vec{n}}(z) f_1(z) - Q_{\vec{n},1}(z) &= \mathcal{O}\left(\frac{1}{z^{n_1+1}}\right), \qquad z \to \infty \\ &\vdots \end{aligned} \tag{5.2}$$

$$P_{\vec{n}}(z)f_r(z) - Q_{\vec{n},r}(z) = \mathcal{O}\left(\frac{1}{z^{n_r+1}}\right), \qquad z \to \infty.$$

Type II Hermite-Padé approximation therefore corresponds to an approximation of each function f_j separately by rational functions *with a common denominator* $P_{\vec{n}}$. Combinations of type I and type II Hermite-Padé approximation also are possible.

5.1 *Orthogonality*

When we consider r Markov functions

$$f_j(z) = \int_{a_j}^{b_j} \frac{d\mu_j(x)}{z-x}, \qquad j = 1, 2, \ldots, r,$$

then Hermite-Padé approximation corresponds to certain orthogonality conditions.

First consider type I approximation. Multiply (5.1) by z^k and integrate over a contour Γ encircling all the intervals $[a_j, b_j]$ in the positive direction, then

$$\frac{1}{2\pi i}\int_\Gamma \sum_{j=1}^r z^k A_{\vec{n},j}(z) f_j(z)\,dz - \frac{1}{2\pi i}\int_\Gamma z^k B_{\vec{n}}(z)\,dz = \sum_{\ell=|\vec{n}|}^\infty b_\ell \frac{1}{2\pi i}\int_\Gamma z^{k-\ell}\,dz.$$

Clearly Cauchy's theorem implies

$$\frac{1}{2\pi i}\int_\Gamma z^k B_{\vec{n}}(z)\,dz = 0.$$

Furthermore, there is only a contribution on the right hand side when $\ell = k+1$, so when $k \le |\vec{n}| - 2$, then none of the terms in the infinite sum have a contribution. Therefore we see that

$$\frac{1}{2\pi i}\int_\Gamma \sum_{j=1}^r z^k A_{\vec{n},j}(z) f_j(z)\,dz = 0, \qquad 0 \le k \le |\vec{n}| - 2.$$

Now each f_j is a Markov function, so by changing the order of integration we get

$$\frac{1}{2\pi i}\int_\Gamma z^k A_{\vec{n},j}(z) f_j(z)\,dz = \int_{a_j}^{b_j} d\mu_j(x)\,\frac{1}{2\pi i}\int_\Gamma \frac{z^k A_{\vec{n},j}(z)}{z-x}\,dz.$$

Since Γ is a contour encircling $[a_j, b_j]$ we have that

$$\frac{1}{2\pi i} \int_{\Gamma} \frac{z^k A_{\vec{n},j}(z)}{z-x} \, dz = x^k A_{\vec{n},j}(x),$$

so that we get the following orthogonality conditions

$$\sum_{j=1}^{r} \int_{a_j}^{b_j} x^k A_{\vec{n},j}(x) \, d\mu_j(x) = 0, \qquad k = 0, 1, \ldots, |\vec{n}| - 2. \tag{5.3}$$

These are $|\vec{n}| - 1$ linear and homogeneous equations for the $|\vec{n}|$ coefficients of the r polynomials $A_{\vec{n},j}$ $(j = 1, 2, \ldots, r)$, so that we can determine these polynomials up to a multiplicative factor, provided that the rank of the matrix in this system is $|\vec{n}| - 1$. If the solution is unique (up to a multiplicative factor), then we say that $\vec{n}$ is a normal index for type I. One can show that this is equivalent with the condition that the degree of each $A_{\vec{n},j}$ is exactly $n_j - 1$. We call the vector $(A_{\vec{n},1}, \ldots, A_{\vec{n},r})$ the *multiple orthogonal polynomials of type I* for $(\mu_1, \ldots, \mu_r)$. Once the polynomial vector $(A_{\vec{n},1}, \ldots, A_{\vec{n},r})$ is determined, we can also find the remaining polynomial $B_{\vec{n}}$ which is given by

$$B_{\vec{n}}(z) = \sum_{j=1}^{r} \int_{a_j}^{b_j} \frac{A_{\vec{n},j}(z) - A_{\vec{n},j}(x)}{z-x} \, d\mu_j(x). \tag{5.4}$$

Indeed, with this definition of $B_{\vec{n}}$ we have

$$\sum_{j=1}^{r} A_{\vec{n},j}(z) f_j(z) - B_{\vec{n}}(z) = \sum_{j=1}^{r} \int_{a_j}^{b_j} \frac{A_{\vec{n},j}(x)}{z-x} \, d\mu_j(x). \tag{5.5}$$

If we use the expansion

$$\frac{1}{z-x} = \sum_{k=0}^{\infty} \frac{x^k}{z^{k+1}},$$

then the right hand side is

$$\sum_{k=0}^{\infty} \frac{1}{z^{k+1}} \sum_{j=1}^{r} \int_{a_j}^{b_j} x^k A_{\vec{n},j}(x) \, d\mu_j(x),$$

and the orthogonality conditions (5.3) show that the sum over k starts with $k = |\vec{n}| - 1$, hence the right hand side is $\mathcal{O}(z^{-|\vec{n}|})$, which is the order given in the definition of type I Hermite-Padé approximation.

Next we consider type II approximation. Multiply (5.2) by z^k and integrate over a contour Γ encircling all the intervals $[a_j, b_j]$, then

$$\frac{1}{2\pi i} \int_\Gamma z^k P_{\vec{n}}(z) f_j(z)\, dz - \frac{1}{2\pi i} \int_\Gamma z^k Q_{\vec{n},j}(z)\, dz = \sum_{\ell=n_j+1}^{\infty} b_\ell \frac{1}{2\pi i} \int_\Gamma z^{k-\ell}\, dz.$$

Cauchy's theorem gives

$$\frac{1}{2\pi i} \int_\Gamma z^k Q_{\vec{n},j}(z)\, dz = 0,$$

and on the right hand side we only have a contribution when $\ell = k+1$. So for $k \leq n_j - 1$ none of the terms in the infinite sum contribute. Hence

$$\frac{1}{2\pi i} \int_\Gamma z^k P_{\vec{n}}(z) f_j(z)\, dz = 0, \qquad 0 \leq k \leq n_j - 1.$$

Interchanging the order of integration on the left hand side gives the orthogonality conditions

$$\begin{aligned} &\int_{a_1}^{b_1} x^k P_{\vec{n}}(x)\, d\mu_1(x) = 0, \qquad k = 0, 1, \ldots, n_1 - 1, \\ &\qquad\qquad \vdots \\ &\int_{a_r}^{b_r} x^k P_{\vec{n}}(x)\, d\mu_r(x) = 0, \qquad k = 0, 1, \ldots, n_r - 1. \end{aligned} \tag{5.6}$$

This gives $|\vec{n}|$ linear and homogeneous equations for the $|\vec{n}|+1$ coefficients of $P_{\vec{n}}$, hence we can obtain the polynomial $P_{\vec{n}}$ up to a multiplicative factor, provided the matrix of coefficients has rank $|\vec{n}|$. In that case we call the index $\vec{n}$ normal for type II. One can show that this is equivalent with the condition that the degree of $P_{\vec{n}}$ is exactly $|\vec{n}|$. We call this polynomial $P_{\vec{n}}$ the *multiple orthogonal polynomial of type II* for $(\mu_1, \ldots, \mu_r)$. Once the polynomial $P_{\vec{n}}$ is determined, we can obtain the polynomials $Q_{\vec{n},j}$ by

$$Q_{\vec{n},j}(z) = \int_{a_j}^{b_j} \frac{P_{\vec{n}}(z) - P_{\vec{n}}(x)}{z - x}\, d\mu_j(x). \tag{5.7}$$

Indeed, with this expression for $Q_{\vec{n},j}$ we have

$$P_{\vec{n}}(z) f_j(z) - Q_{\vec{n},j}(z) = \int_{a_j}^{b_j} \frac{P_{\vec{n}}(x)}{z - x}\, d\mu_j(x), \tag{5.8}$$

and if we expand $1/(z-x)$, then the right hand side is of the form

$$\sum_{k=0}^{\infty} \frac{1}{z^{k+1}} \int_{a_j}^{b_j} x^k P_{\vec{n}}(x)\, d\mu_j(x),$$

and the orthogonality conditions (5.6) show that the infinite sum starts at $k = n_j$, which gives an expression of $\mathcal{O}(z^{-n_j-1})$, which is exactly what is required for type II Hermite-Padé approximation.

5.2 Angelesco systems

An interesting system of functions, which allows detailed analysis, was introduced by Angelesco [1]:

Definition 5.3 *An Angelesco system $(f_1, f_2, \ldots, f_r)$ consists of r Markov functions for which the intervals (a_j, b_j) are pairwise disjoint.*

All multi-indices are normal for type II in an Angelesco system. We will prove this by showing that the multiple orthogonal polynomial $P_{\vec{n}}$ has degree exactly equal to $|\vec{n}|$. In fact more is true, namely

Theorem 5.1 *If $(f_1, \ldots, f_r)$ is an Angelesco system with measures μ_j that have infinitely many point in their support. Then $P_{\vec{n}}$ has n_j simple zeros on (a_j, b_j) for $j = 1, \ldots, r$.*

PROOF. Let $x_1, \ldots, x_m$ be the sign changes of $P_{\vec{n}}$ on (a_j, b_j). Suppose that $m < n_j$ and let $\pi_m(x) = (x-x_1)\cdots(x-x_m)$, then $P_{\vec{n}}\pi_m$ does not change sign on $[a_j, b_j]$. Since the support of μ_j has infinitely many points, we have

$$\int_{a_j}^{b_j} P_{\vec{n}}(x)\pi_m(x)\, d\mu_j(x) \neq 0.$$

However, the orthogonality (5.6) implies that $P_{\vec{n}}$ is orthogonal to all polynomials of degree $\leq n_j - 1$ with respect to the measure μ_j on $[a_j, b_j]$, so that the integral is zero. This contradiction implies that $m \geq n_j$, and hence $P_{\vec{n}}$ has at least n_j zeros on (a_j, b_j). This holds for every j, and since the intervals (a_j, b_j) are disjoint this gives at least $|\vec{n}|$ zeros on the real line. But the degree of $P_{\vec{n}}$ is $\leq |\vec{n}|$, hence $P_{\vec{n}}$ has exactly n_j simple zeros on (a_j, b_j). □

The polynomial $P_{\vec{n}}$ can therefore be factored as

$$P_{\vec{n}}(x) = q_{n_1}(x) q_{n_2}(x) \cdots q_{n_r}(x),$$

where each q_{n_j} is a polynomial of degree n_j with its zeros on (a_j, b_j). The orthogonality (5.6) then gives

$$\int_{a_j}^{b_j} x^k q_{n_j}(x) \prod_{i \neq j} q_{n_i}(x)\, d\mu_j(x) = 0, \qquad k = 0, 1, \ldots, n_j - 1. \tag{5.9}$$

The product $\prod_{i \neq j} q_{n_i}(x)$ does not change sign on (a_j, b_j), hence (5.9) shows that q_{n_j} is an ordinary orthogonal polynomial of degree n_j on the interval $[a_j, b_j]$ with respect to the measure $\prod_{i \neq j} |q_{n_i}(x)|\, d\mu_j(x)$. The measure depends on the multi-index $\vec{n}$.

5.3 Algebraic Chebyshev systems

A Chebyshev system $\{\varphi_1, \ldots, \varphi_n\}$ on $[a, b]$ is a system of n linearly independent functions such that every linear combination $\sum_{k=1}^n a_k \varphi_k$ has at most $n - 1$ zeros on $[a, b]$. This is equivalent with the condition that

$$\det \begin{pmatrix} \varphi_1(x_1) & \varphi_1(x_2) & \cdots & \varphi_1(x_n) \\ \varphi_2(x_1) & \varphi_2(x_2) & \cdots & \varphi_2(x_n) \\ \vdots & \vdots & \cdots & \vdots \\ \varphi_n(x_1) & \varphi_n(x_2) & \cdots & \varphi_n(x_n) \end{pmatrix} \neq 0$$

for every choice of n different points $x_1, \ldots, x_n \in [a, b]$. Indeed, when $x_1, \ldots, x_n$ are such that the determinant is zero, then there is a linear combination of the rows that gives a zero row, but this means that for this linear combination $\sum_{k=1}^n a_k \varphi_k$ has zeros at $x_1, \ldots, x_n$, giving n zeros, which is not allowed.

Definition 5.4 *A system $(f_1, \ldots, f_r)$ is an algebraic Chebyshev system (AT system) for the index $\vec{n}$ if each f_j is a Markov function on the same interval $[a, b]$ with a measure $w_j(x)\, d\mu(x)$, where μ has an infinite support and the w_j are such that*

$$\{w_1, xw_1, \ldots, x^{n_1-1}w_1, w_2, xw_2, \ldots, x^{n_2-1}w_2, \ldots, w_r, xw_r, \ldots, x^{n_r-1}w_r\} \tag{5.10}$$

is a Chebyshev system on $[a, b]$.

Theorem 5.2 *Suppose $\vec{n}$ is a multi-index such that $(f_1, \ldots, f_r)$ is an AT system on $[a, b]$ for every index $\vec{m}$ for which $m_j \leq n_j$ $(1 \leq j \leq r)$. Then $P_{\vec{n}}$ has $|\vec{n}|$ zeros on (a, b) and hence $\vec{n}$ is a normal index for type II.*

PROOF. Let $x_1, \ldots, x_m$ be the sign changes of $P_{\vec{n}}$ on (a, b) and suppose that $m < |\vec{n}|$. We can then find a multi-index $\vec{m}$ such that $|\vec{m}| = m$ and $m_j \leq n_j$ for every $1 \leq j \leq r$ and $m_k < n_k$

for some $1 \leq k \leq r$. Consider the interpolation problem where we want to find a function

$$L(x) = \sum_{j=1}^{r} q_j(x) w_j(x),$$

where q_j is a polynomial of degree $m_j - 1$ if $j \neq k$ and q_k a polynomial of degree m_k, that satisfies

$$\begin{aligned} L(x_j) &= 0, \qquad j = 1, ..., m, \\ L(x_0) &= 1, \qquad \text{for some other point } x_0 \in [a, b], \end{aligned}$$

then this interpolation problem has a unique solution since this involves a Chebyshev system of basis functions. The function L has, by construction, m zeros and the Chebyshev system has $m + 1$ basis functions, so L can have at most m zeros on $[a, b]$ and each zero is a sign change. Hence $P_{\vec{n}} L$ does not change sign on $[a, b]$. Since μ has infinite support, we thus have

$$\int_a^b L(x) P_{\vec{n}}(x) d\mu(x) \neq 0.$$

But the orthogonality (5.6) gives

$$\int_a^b q_j(x) P_{\vec{n}}(x) w_j(x) \, d\mu(x) = 0, \qquad j = 1, 2, \ldots, r,$$

and this contradiction implies that $P_{\vec{n}}$ has $|\vec{n}|$ simple zeros on (a, b). □

We have a similar result for type I Hermite-Padé approximation:

Theorem 5.3 *Suppose $\vec{n}$ is a multi-index such that $(f_1, \ldots, f_r)$ is an AT system on $[a, b]$ for every index $\vec{m}$ for which $m_j \leq n_j$ $(1 \leq j \leq r)$. Then $\sum_{j=1}^{r} A_{\vec{n},j} w_j$ has $|\vec{n}| - 1$ zeros on (a, b) and $\vec{n}$ is a normal index for type I.*

PROOF. Let $x_1, \ldots, x_m$ be the sign changes of $\sum_{j=1}^{r} A_{\vec{n},j} w_j$ on (a, b) and suppose that $m < |\vec{n}| - 1$. Let π_m be the monic polynomial with these points as zeros, then $\pi_m \sum_{j=1}^{r} A_{\vec{n},j} w_j$ does not change sign on $[a, b]$ and hence

$$\int_a^b \pi_m(x) \sum_{j=1}^{r} A_{\vec{n},j}(x) w_j(x) \, d\mu(x) \neq 0.$$

But the orthogonality conditions (5.3) indicate that this integral is zero. This contradiction implies that $m \geq |\vec{n}| - 1$. The sum $\sum_{j=1}^{r} A_{\vec{n},j} w_j$ is a linear combination of the Chebyshev system (5.10), hence it has at most $|\vec{n}| - 1$ zeros on $[a, b]$. Therefore we see that $m = |\vec{n}| - 1$.

To see that the index $\vec{n}$ is normal for type I, we assume that for some k with $1 \leq k \leq r$ the degree of $A_{\vec{n},k}$ is less than $n_k - 1$. Then $\sum_{j=1}^r A_{\vec{n},j} w_j$ is a linear combination of a the Chebyshev system (5.10) from which the function $x^{n_k-1} w_k$ is removed. This is still a Chebyshev system by assumption, and hence this linear combination has at most $|\vec{n}| - 2$ zeros on $[a, b]$. But this contradicts our previous observation that it has $|\vec{n}| - 1$ zeros. Therefore every $A_{\vec{n},j}$ has degree exactly $n_j - 1$, so that the index $\vec{n}$ is normal. □

5.4 Nikishin systems

A special construction, suggested by Nikishin [39], gives an AT system that can be handled in some detail. The construction is by induction. A Nikishin system of order 1 is a Markov function $f_{1,1}$ for a measure μ_1 on the interval $[a_1, b_1]$. A Nikishin system of order 2 is a vector of Markov functions $(f_{1,2}, f_{2,2})$ on $[a_2, b_2]$ such that

$$f_{1,2}(z) = \int_{a_2}^{b_2} \frac{d\mu_2(x)}{z - x}, \quad f_{2,2}(z) = \int_{a_2}^{b_2} f_{1,1}(x) \frac{d\mu_2(x)}{z - x},$$

where $f_{1,1}$ is a Nikishin system of order 1 on $[a_1, b_1]$ and $(a_1, b_1) \cap (a_2, b_2) = \emptyset$. In general we have

Definition 5.5 *A Nikishin system of order r consists of r Markov functions $(f_{1,r}, \ldots, f_{r,r})$ on $[a_r, b_r]$ such that*

$$f_{1,r}(z) = \int_{a_r}^{b_r} \frac{d\mu_r(x)}{z - x}, \tag{5.11}$$

$$f_{j,r}(z) = \int_{a_r}^{b_r} f_{j-1,r-1}(x) \frac{d\mu_r(x)}{z - x}, \qquad j = 2, \ldots, r, \tag{5.12}$$

where $(f_{1,r-1}, \ldots, f_{r-1,r-1})$ is a Nikishin system of order $r - 1$ on $[a_{r-1}, b_{r-1}]$ and $(a_r, b_r) \cap (a_{r-1}, b_{r-1}) = \emptyset$.

For a Nikishin system of order r one knows that the multi-indices $\vec{n}$ with $n_1 \geq n_2 \geq \cdots \geq n_r$ are normal (the system is an AT-system for these indices), but is an open problem whether every multi-index is normal (for $r > 2$; for $r = 2$ it has been proved that every multi-index is normal).

What can be said about type II Hermite-Padé approximation for $r = 2$? Recall (5.8) for the function $f_{1,2}$:

$$P_{n_1,n_2}(y) f_{1,2}(y) - Q_{n_1,n_2;1}(y) = \int_{a_2}^{b_2} \frac{P_{n_1,n_2}(x)}{y - x} \, d\mu_2(x).$$

Multiply both sides by y^k, with $k \leq n_1$, then the right hand side is

$$\int_{a_2}^{b_2} \frac{y^k P_{n_1,n_2}(x)}{y-x}\, d\mu_2(x) = \int_{a_2}^{b_2} \frac{(y^k - x^k) P_{n_1,n_2}(x)}{y-x}\, d\mu_2(x) + \int_{a_2}^{b_2} \frac{x^k P_{n_1,n_2}(x)}{y-x}\, d\mu_2(x).$$

Clearly $(y^k - x^k)/(y-x)$ is a polynomial in x of degree $k-1 \leq n_1 - 1$ hence the first integral on the right vanishes because of the orthogonality (5.6). Integrate over the variable $y \in [a_1, b_1]$ with respect to the measure μ_1, then we find for $k \leq n_1$

$$\int_{a_1}^{b_1} [P_{n_1,n_2}(y) f_{1,2}(y) - Q_{n_1,n_2;1}(y)] y^k\, d\mu_1(y) = \int_{a_1}^{b_1}\int_{a_2}^{b_2} \frac{x^k P_{n_1,n_2}(x)}{y-x}\, d\mu_2(x)\, d\mu_1(y).$$

Change the order of integration on the right hand side, then

$$\int_{a_1}^{b_1} [P_{n_1,n_2}(y) f_{1,2}(y) - Q_{n_1,n_2;1}(y)] y^k\, d\mu_1(y) = -\int_{a_2}^{b_2} x^k P_{n_1,n_2}(x) f_{1,1}(x)\, d\mu_2(x)$$

and this is zero for $k \leq n_2 - 1$. Hence if $n_2 \leq n_1 + 1$ then the expression $P_{n_1,n_2}(y) f_{1,2}(y) - Q_{n_1,n_2;1}(y)$ is orthogonal to all polynomials of degree $\leq n_2 - 1$ on $[a_1, b_1]$. This implies that $P_{n_1,n_2}(y) f_{1,2}(y) - Q_{n_1,n_2;1}(y)$ has at least n_2 zeros on (a_1, b_1) using an argument similar to what we have been using earlier. Let R_{n_2} be the monic polynomial with n_2 of these zeros on (a_1, b_1), then $[P_{n_1,n_2}(y) f_{1,2}(y) - Q_{n_1,n_2;1}(y)]/R_{n_2}(y)$ is an analytic function on $\mathbb{C} \setminus [a_2, b_2]$, which has the representation

$$\frac{P_{n_1,n_2}(y) f_{1,2}(y) - Q_{n_1,n_2;1}(y)}{R_{n_2}(y)} = \frac{1}{R_{n_2}(y)} \int_{a_2}^{b_2} \frac{P_{n_1,n_2}(x)}{y-x}\, d\mu_2(x).$$

Multiply both sides by y^k and integrate over a contour Γ encircling the interval $[a_2, b_2]$ in the positive direction, but with all the zeros of R_{n_2} outside Γ, then

$$\frac{1}{2\pi i} \int_{\Gamma} y^k \frac{P_{n_1,n_2}(y) f_{1,2}(y) - Q_{n_1,n_2;1}(y)}{R_{n_2}(y)}\, dy = \frac{1}{2\pi i} \int_{\Gamma} \frac{y^k}{R_{n_2}(y)} \frac{P_{n_1,n_2}(x)}{y-x}\, d\mu_2(x)\, dy.$$

If we interchange the order of integration on the right hand side and use Cauchy's theorem, then this gives the integral

$$\int_{a_2}^{b_2} x^k P_{n_1,n_2}(x) \frac{d\mu_2(x)}{R_{n_2}(x)}.$$

By the interpolation condition (5.2) the integrand on the left hand side is of the order $\mathcal{O}(y^{k-n_1-n_2-1})$, so if we use Cauchy's theorem for the exterior of Γ, then the integral vanishes for $k \leq n_1 + n_2 - 1$.

Hence we get

$$\int_{a_2}^{b_2} x^k P_{n_1,n_2}(x)\,\frac{d\mu_2(x)}{R_{n_2}(x)} = 0, \qquad k = 0, 1, \ldots n_1 + n_2 - 1. \tag{5.13}$$

This shows that P_{n_1,n_2} is an ordinary orthogonal polynomial on $[a_2, b_2]$ with respect to the measure $d\mu_2(x)/R_{n_2}(x)$. Observe that $(a_1, b_1) \cap (a_2, b_2) = \emptyset$ implies that R_{n_2} does not change sign on $[a_2, b_2]$. Finally we have

$$\begin{aligned}\int_{a_2}^{b_2} \frac{P_{n_1,n_2}^2(x)}{y-x}\,\frac{d\mu_2(x)}{R_{n_2}(x)} &= \int_{a_2}^{b_2} P_{n_1,n_2}(x)\frac{P_{n_1,n_2}(x) - P_{n_1,n_2}(y)}{y-x}\,\frac{d\mu_2(x)}{R_{n_2}(x)} \\ &\quad + P_{n_1,n_2}(y)\int_{a_2}^{b_2} \frac{P_{n_1,n_2}(x)}{y-x}\,\frac{d\mu_2(x)}{R_{n_2}(x)} \\ &= P_{n_1,n_2}(y)\int_{a_2}^{b_2} \frac{P_{n_1,n_2}(x)}{y-x}\,\frac{d\mu_2(x)}{R_{n_2}(x)},\end{aligned}$$

since $[P_{n_1,n_2}(y) - P_{n_1,n_2}(x)]/(y-x)$ is a polynomials in x of degree $n_1 + n_2 - 1$ and because of the orthogonality (5.13). Hence

$$P_{n_1,n_2}(y) f_{1,2}(y) - Q_{n_1,n_2;1}(y) = \frac{R_{n_2}(y)}{P_{n_1,n_2}(y)} \int_{a_2}^{b_2} \frac{P_{n_1,n_2}^2(x)}{y-x}\,\frac{d\mu_2(x)}{R_{n_2}(x)}. \tag{5.14}$$

Both sides of the equation have zeros at the zeros of R_{n_2}, but there will not be any other zeros on $[a_1, b_1]$ since the integral on the right hand side has constant sign.

5.5 Some applications

Many of the classical orthogonal polynomials have been extended to this multiple orthogonality setting: the Jacobi, Laguerre and Hermite polynomials have multiple extensions worked out in [3], [32], [47], [48]. Discrete multiple orthogonal polynomials have been found in [4] and [41]. New special polynomials corresponding to orthogonality measures involving Bessel functions were found in [19] and [50]. Many of the properties of the classical orthogonal polynomials have nice extensions in this multiple orthogonality setting: there will be a higher order linear recurrence relation, there are nice differential or difference properties, such a linear differential equation (of higher order) and Rodrigues-type formulas. The weak asymptotics (and the asymptotic distribution of the zeros) has been worked out by means of an equilibrium problem for vector potentials [29] and recently a matrix Riemann-Hilbert problem was found for multiple orthogonal polynomials [49] which will be very useful for obtaining strong asymptotics, uniformly in the complex plane.

5.5.1 Irrationality and transcendence

Hermite-Padé approximation finds its origin in number theory. Hermite's proof of the transcendence of e is based on Hermite-Padé approximation of $(e^x, e^{2x}, \dots, e^{rx})$ at $x = 0$. Many proofs of irrationality are also based on Hermite-Padé approximation, even though this is often not explicit in the proof. Apéry's proof that $\zeta(3)$ is irrational can be reduced to Hermite-Padé approximation to three functions

$$f_1(z) = \int_0^1 \frac{dx}{z-x}, \quad f_2(z) = -\int_0^1 \log x \frac{dx}{z-x}, \quad f_3(z) = \frac{1}{2}\int_0^1 \log^2 x \frac{dx}{z-x},$$

which form an AT-system. The proof uses a mixture of type I and type II Hermite-Padé approximation: find polynomials (A_n, B_n) (both of degree n) and polynomials C_n and D_n such that

$$\begin{aligned} A_n(1) &= 0 \\ A_n(z)f_1(z) + B_n(z)f_2(z) - C_n(z) &= \mathcal{O}(1/z^{n+1}), \qquad z \to \infty \\ A_n(z)f_2(z) + 2B_n(z)f_3(z) - D_n(z) &= \mathcal{O}(1/z^{n+1}), \qquad z \to \infty. \end{aligned}$$

Observe that $f_3(1) = \zeta(3)$, hence if we evaluate the approximations at $z = 1$, then we see that $2B_n(1)\zeta(3) - D_n(1)$ will be small and $D_n(1)/(2B_n(1))$ is a good rational approximation to $\zeta(3)$. In fact, asymptotic analysis of the error and of the denominator $B_n(1)$ and some simple number theory show that this rational approximation is better than order 1, which implies that $\zeta(3)$ is irrational. See [47] for details.

For another example we consider the two Markov functions

$$f_1(z) = \int_0^1 \frac{dx}{z-x}, \quad f_2(z) = \int_{-1}^0 \frac{dx}{z-x},$$

which form an Angelesco system. Some straightforward calculus gives

$$f_1(i) = -\frac{1}{2}\log 2 - \frac{i\pi}{4}, \quad f_2(i) = \frac{1}{2}\log 2 - \frac{i\pi}{4},$$

hence the sum gives $f_1(i) + f_2(i) = -i\pi/2$. The type II Hermite-Padé approximants for f_1 and f_2 will give approximations to π. Recall that

$$\begin{aligned} P_{n,n}(z)f_1(z) - Q_{n,n;1}(z) &= \int_0^1 \frac{P_{n,n}(x)}{z-x}\,dx \\ P_{n,n}(z)f_2(z) - Q_{n,n;2}(z) &= \int_{-1}^0 \frac{P_{n,n}(x)}{z-x}\,dx. \end{aligned}$$

Summing both equations gives

$$P_{n,n}(z)[f_1(z)+f_2(z)] - [Q_{n,n;1}(z)+Q_{n,n;2}(z)] = \int_{-1}^{1} \frac{P_{n,n}(x)}{z-x}\,dx.$$

So the fact that we are using a common denominator comes in very handy here. Then we evaluate these expressions at $z=i$ and hope that $P_{n,n}(i)$ and $Q_{n,n;1}(i)+Q_{n,n;2}(i)$ are (up to the factor i) integers or rational numbers with simple denominators. Asymptotic properties of the Hermite-Padé approximants and the multiple orthogonal polynomials then gives useful quantitative information about the order of rational approximation to π. For this particular case the type II multiple orthogonal polynomials are given by a Rodrigues formula

$$P_{n,n}(x) = \frac{d^n}{dx^n}\left(x^n(1-x^2)^n\right),$$

and these polynomials are known as Legendre-Angelesco polynomials. They have been studied in detail by Kalyagin [32] (see also [47]). The Rodrigues formula in fact simplifies the asymptotic analysis, since integration by parts now gives

$$\int_{-1}^{1} \frac{P_{n,n}(x)}{z-x}\,dx = \int_{-1}^{1} (-1)^n n! \frac{x^n(1-x^2)^n}{(z-x)^{n+1}}\,dx,$$

which can be handled easily. Some trial and error show that one gets better results by taking $2n$ instead of n, and by differentiating n times extra:

$$\begin{aligned} r\,\frac{d^n}{dz^n}&\left(P_{2n,2n}(z)[f_1(z)+f_2(z)] - [Q_{2n,2n;1}(z)+Q_{2n,2n;2}(z)]\right)_{z=i} \\ &= (3n)!(-i)^{n+1}\int_{-1}^{1} \frac{x^{2n}(1-x^2)^{2n}}{(1+ix)^{3n+1}}\,dx. \end{aligned} \tag{5.15}$$

This gives rational approximants to π of the form

$$\pi = \frac{b_n}{a_n c_n} + \frac{K_n}{a_n},$$

where a_n, b_n, c_n are explicitly known integers and K_n is the integral on the right hand side of (5.15). The rational approximants show that π is irrational (which was shown already in 1761 by Lambert), but they even show that you can't approximate π by rational at order greater than 23.271 (Beukers [8]). This upper bound for the order of approximation can be reduced to 8.02 (Hata [30]) by considering Markov functions f_1 and f_3, with

$$f_3(z) = \int_{-i}^{0} \frac{dx}{z-x}.$$

This f_3 is now over a complex interval, and then Theorem 5.1 about the location of the zeros no longer holds, and the asymptotic behavior will have to be handled with another method.

5.5.2 Random matrices

Multiple orthogonal polynomials appear in certain problems in the theory of random matrices. The connection between eigenvalues of random matrices and orthogonal polynomials is well known: if we define a matrix ensemble by giving the joint probability density function for its eigenvalues as

$$P(x_1, \ldots, x_N) = \prod_{i=1}^{N} f(x_i) \prod_{1 \leq i < j \leq N} (x_i - x_j)^2,$$

then the eigenvalues density σ_N is given by

$$\sigma_n(x) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} P(x, x_2, \ldots, x_N)\, dx_2 \cdots dx_N = \frac{1}{N} \sum_{j=0}^{N-1} p_j^2(x),$$

where the p_n are the orthonormal polynomials with weight function f. Furthermore the n-point correlation function is given in terms of the Christoffel-Darboux kernel

$$\sum_{j=0}^{N-1} p_j(x) p_j(y).$$

(see, e.g., [35, §19.3]). The Gaussian unitary ensemble corresponds to Hermite polynomials.

Recently a random matrix ensemble with an external source was considered by Brézin and Hikami [11] and Zinn-Justin [60]. The joint probability density function of the matrix elements of the random Hermitian matrix M is of the form

$$\frac{1}{Z_N} e^{-\mathrm{Tr}(M^2 - AM)}\, dM$$

where A is a fixed $N \times N$ Hermitian matrix (the external source). Bleher and Kuijlaars [9] observed that the average characteristic polynomial $P_N(z) = \mathrm{E}[\det(zI - M)]$ can be characterized by the property

$$\int_{-\infty}^{\infty} P_N(z) x^k e^{-(x^2 - a_j x)}\, dx = 0, \qquad k = 0, 1, \ldots, N_j - 1,$$

where N_j is the multiplicity of the eigenvalue a_j of A. This means that P_N is a multiple Hermite polynomial of type II with multi-index $(N_1, \ldots, N_r)$ when A has r distinct eigenvalues $a_1, \ldots, a_r$ with multiplicities $N_1, \ldots, N_r$ respectively. These multiple Hermite polynomials were investigated in [3]. The eigenvalue correlations and the eigenvalues density can be written in terms of the kernel

$$\sum_{k=0}^{N-1} P_k(x)Q_k(y),$$

where the Q_k are basically the type I multiple Hermite polynomials and the P_k are the type II multiple Hermite polynomials. The asymptotic analysis of the eigenvalues and their correlations and universality questions can therefore be handled using asymptotic analysis of multiple Hermite polynomials.

Another application is in the theory of coupled random matrices [23] [24] [33]. The two-matrix model deals with pairs of random matrices (M_1, M_2) which are both $N \times N$ Hermitian matrices with joint density function

$$\frac{1}{Z_N} e^{-\mathrm{Tr}(M_1^4+M_2^4-\tau M_1 M_2)}\, dM_1\, dM_2.$$

The statistical relevant quantities for the eigenvalues of M_1 and M_2 can be expressed in terms of biorthogonal polynomials p_k and q_k which satisfy

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_k(x)q_j(y)e^{-x^4-y^4+\tau xy}\, dx\, dy = \delta_{k,j}. \tag{5.16}$$

Due to the symmetry we have that $p_k = q_k$. Consider the functions

$$w_k(y) = \int_{-\infty}^{\infty} x^k e^{-x^4+\tau xy}\, dx,$$

then a simple integration by parts shows that

$$w_{k+3}(y) = \frac{k}{4} w_{k-1}(y) - \frac{\tau y}{4} w_k(y),$$

so that each w_k is a linear combination of w_0, w_1, w_2 with polynomial coefficients, in particular

$$\begin{aligned}
w_{3k}(y) &= a_k(y)w_0(y) + b_{k-1}(y)w_1(y) + c_{k-1}(y)w_2(y) \\
w_{3k+1}(y) &= \hat{a}_k(y)w_0(y) + \hat{b}_k(y)w_1(y) + \hat{c}_{k-1}(y)w_2(y) \\
w_{3k+2}(y) &= \tilde{a}_k(y)w_0(y) + \tilde{b}_k(y)w_1(y) + \tilde{c}_k(y)w_2(y)
\end{aligned}$$

where $a_k, \hat{b}_k, \tilde{c}_k$ are polynomials of degree k. This means that

$$Q_n(y) = \int_{-\infty}^{\infty} p_j(y)e^{-x^4+\tau xy}\, dx$$

is a linear combination of w_0, w_1, w_2 with polynomials coefficients, in particular

$$\begin{aligned}
Q_{3n}(y) &= A_n(y)w_0(y) + B_{n-1}(y)w_1(y) + C_{n-1}(y)w_2(y) \\
Q_{3n+1}(y) &= \widehat{A}_n(y)w_0(y) + \widehat{B}_n(y)w_1(y) + \widehat{C}_{n-1}(y)w_2(y) \\
Q_{3n+2}(y) &= \widetilde{A}_n(y)w_0(y) + \widetilde{B}_n(y)w_1(y) + \widetilde{C}_n(y)w_2(y).
\end{aligned}$$

It turns out that (A_n, B_{n-1}, C_{n-1}), $(\widehat{A}_n, \widehat{B}_n, \widehat{C}_{n-1})$, and $(\widetilde{A}_n, \widetilde{B}_n, \widetilde{C}_n)$ are multiple orthogonal polynomials of type I for the densities $e^{-x^4}w_0(x), e^{-x^4}w_1(x), e^{-x^4}w_2(x)$ with multi-indices $(n+1, n, n)$, $(n+1, n+1, n)$ and $(n+1, n+1, n+1)$ respectively, and p_{3n}, p_{3n+1} and p_{3n+2} are multiple orthogonal polynomials of type II with multi-indices (n, n, n), $(n+1, n, n)$ and $(n+1, n+1, n)$ respectively. The multiple orthogonality conditions (5.3) and (5.6) then lead to the biorthogonality (5.16). Note that w_0 and w_2 are positive densities but w_1 changes sign at the origin.

5.5.3 *Simultaneous Gauss quadrature*

In a number of applications we need to approximate several integrals of the same function, but with respect to different measures. The following example comes from [10]. Suppose that g is the spectral distribution of light in the direction of the observer and w_1, w_2, w_3 are weight functions describing the profiles for red, green and blue light. Then the integrals

$$\int_0^{2\pi} g(x)w_1(x)\,dx, \quad \int_0^{2\pi} g(x)w_2(x)\,dx, \quad \int_0^{2\pi} g(x)w_3(x)\,dx$$

give the amount of light after passing through the filters for red, green and blue. In this case we need to approximate three integrals of the same function g. We would like to use as few function evaluations as possible, but the integrals should be accurate for polynomials g of degree as high as possible. If we use Gauss quadrature with n nodes for each integral, then we require $3n$ function evaluations and all integrals will be correct for polynomials of degree $\leq 2n-1$ (a space of dimension 2n). This gives an efficiency of 3/2. In fact, with $3n$ function evaluations we can double the dimension of the space in which the formula is exact. Consider the Markov functions

$$f_j(z) = \int_a^b \frac{w_j(x)\,dx}{z-x}, \qquad j = 1, 2, 3$$

and the type II Hermite-Padé approximation problem

$$f_j(z) - \frac{Q_{n,n,n;j}(z)}{P_{n,n,n}(z)} = \mathcal{O}(z^{-4n-1}), \qquad z \to \infty.$$

Now we can multiply by a polynomial π_{4n-1} of degree at most $4n-1$, and integrate along contour Γ encircling $[a,b]$ in the positive direction, to obtain

$$\int_a^b \pi_{4n-1}(x) w_j(x),\ dx = \sum_{k=1}^{3n} \lambda_{k,n;j} g(x_{k,n}), \qquad j = 1,2,3 \tag{5.17}$$

where $x_{k,n}$ are the zeros of $P_{n,n,n}$ and $\lambda_{k,n;j}$ are the residues of $Q_{n,n,n;j}/P_{n,n,n}$ at the zero $x_{k,n}$:

$$\lambda_{k,n;j} = \frac{Q_{n,n,n;j}(x_{k,n})}{P'_{n,n,n}(x_{k,n})}.$$

Therefore the three integrals will be evaluated exactly by the three sums in (5.17) for polynomials of degree $\leq 4n-1$. The convergence is somewhat more difficult to handle, since we do not have a general result that the quadrature coefficients $\lambda_{k,n;j}$ are positive. The positivity has to be investigated separately for Angelesco and Nikishin systems.

References

[1] M.A. Angelesco. Sur deux extensions des fractions continues algébriques. *C.R. Acad. Sci. Paris*, 18:262–263, 1919.

[2] A.I. Aptekarev. Multiple orthogonal polynomials. *J. Comput. Appl. Math.*, 99:423–447, 1998.

[3] A.I. Aptekarev, A. Branquinho, and W. Van Assche. Multiple orthogonal polynomials for classical weights. *Trans. Amer. Math. Soc.*, 355(10):3887–3914, 2003.

[4] J. Arvesú, J. Coussement, and W. Van Assche. Some discrete multiple orthogonal polynomials. *J. Comput. Appl. Math.*, 153:19–45, 2003.

[5] B. Benouahmane and A. Cuyt. Multivariate orthogonal polynomials, homogeneous Padé approximants and Gaussian cubature. *Numer. Algorithms*, 24:1–15, 2000.

[6] B. Benouahmane and A. Cuyt. Properties of multivariate homogeneous orthogonal polynomials. *J. Approx. Theory*, 113:1–20, 2001.

[7] B. Benouahmane, A. Cuyt, and B. Verdonk. On the solution of parameterized (smallest, largest or multiple) eigenvalue problems. Technical report, University of Antwerp (UA), 2004.

[8] F. Beukers. A rational approach to π. *Nieuw Arch. Wisk.*, 1:372–379, 2000.

[9] P.M. Bleher and A.B.J. Kuijlaars. Random matrices with external source and multiple orthogonal polynomials. *Internat. Math. Res. Notices*, 4:109–129, 2004.

[10] C.F. Borges. On a class of Gauss-like quadrature rules. *Numer. Math.*, 67:271–288, 1994.

[11] E. Brézin and S. Hikami. Level spacing of random matrices in an external source. *Phys. Rev. E*, 58:7176–7185, 1998.

[12] C. Brezinski. A direct proof of the Christoffel-Darboux identity and its equivalence to the recurrence relationship. *J. Comput. Appl. Math.*, 32:17–25, 1990.

[13] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Numerical quadratures on the unit circle: a survey with some new results. In M. Alfaro, A. García, C. Jagels, and F. Marcellan, editors, *Orthogonal Polynomials on the Unit Circle: Theory and Applications*, pages 1–20. Universidad Carlos III de Madrid, 1994.

[14] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. *Orthogonal rational functions*, volume 5 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, 1999.

[15] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Quadrature and orthogonal rational functions. *J. Comput. Appl. Math.*, 127(1-2):67–91, 2001. Invited paper.

[16] A. Bultheel and M. Van Barel. Formal orthogonal polynomials and Hankel/Toeplitz duality. *Numer. Algorithms*, 10:289–335, 1995.

[17] A. Bultheel and M. Van Barel. Vector orthogonal polynomials and least squares approximation. *SIAM J. Matrix Anal. Appl.*, 16(3):863–885, 1995.

[18] A. Bultheel, M. Van Barel, and P. Van gucht. Orthogonal bases in discrete least squares rational approximation. *J. Comput. Appl. Math.*, 2003. Accepted. Invited paper ICCAM 2002 Conference.

[19] E. Coussement and W. Van Assche. Multiple orthogonal polynomials associated with the modified Bessel functions of the first kind. *Constr. Approx.*, 19:237–263, 2003.

[20] A. Cuyt. How well can the concept of Padé approximant be generalized to the multivariate case. *J. Comput. Appl. Math.*, 105:25–50, 1999.

[21] M.G. de Bruin. Simultaneous Padé approximation and orthogonality. In C. Brezinski, A. Draux, A.P. Magnus, P. Maroni, and A. Ronveaux, editors, *Proc. Polynômes Orthogoneaux et Applications, Bar-le-Duc, 1984*, volume 1171 of *Lecture Notes in Math.*, pages 74–83. Springer, 1985.

[22] M.M. Djrbashian. A survey on the theory of orthogonal systems and some open problems. In P. Nevai, editor, *Orthogonal polynomials: Theory and practice*, volume 294 of *Series C: Mathematical and Physical Sciences*, pages 135–146, Boston, 1990. NATO-ASI, Kluwer Academic Publishers.

[23] N. Ercolani and K.T-R. McLaughlin. Asymptotics and integrable structures for biorthogonal polynomials associated to a random two-matrix model. *J. Phys. D*, 152/153:232–268, 2001.

[24] B. Eynard and M.L. Mehta. Matrices coupled in a chain: eigenvalue correlations. *J. Phys. A*, 31:4449–4456, 1998.

[25] G. Freud. *Orthogonal polynomials*. Pergamon Press, Oxford, 1971.

[26] B. Fritzsche and B. Kirstein amd A. Lasarow. Orthogonal rational matrix-valued functions on the unit circle. *Math. Nachr.*, 2002. To appear.

[27] B. Fritzsche and B. Kirstein amd A. Lasarow. Further aspects of the theory of orthogonal rational matrix-valued functions on the unit circle. *Math. Nachr.*, 2003. Submitted.

[28] Ya. Geronimus. *Polynomials orthogonal on a circle and interval*. International Series of Monographs in Pure and Applied Mathematics. Pergamon Press, Oxford, 1960.

[29] A.A. Gonchar, E.A. Rakhmanov, and V.N. Sorokin. Hermite-Padé approximants for systems of Markov-type functions. *Math. Sb.*, 188:38–58, 1997. (Russian), *Russian Acad. Sb. Math.* **188** (1997), 671–696.

[30] M. Hata. Rational approximations to π and some other numbers. *Acta Arith.*, 63(4):335–349, 1993.

[31] W.B. Jones, O. Njåstad, and W.J. Thron. Moment theory, orthogonal polynomials, quadrature and continued fractions associated with the unit circle. *Bull. London Math. Soc.*, 21:113–152, 1989.

[32] V.A. Kalyagin (Kaliaguine). On a class of polynomials defined by two orthogonality conditions. *Math. Sb.*, 110:609–627, 1979. (Russian), *Math. USSR Sb.* **38** (1981), 563–580.

[33] A.B.J. Kuijlaars and K.T-R. McLaughlin. A Riemann-Hilbert problem for biorthogonal polynomials, 2003. `arXiv:math.CV/0310204`.

[34] A. Lasarow. *Aufbau einer Szegő-Theorie für rationale Matrixfunktionen.* PhD thesis, Universität Leipzig, Fak. Mathematik Informatik, 2000.

[35] M.L. Mehta. *Random Matrices* (revised and enlarged second edition). Academic Press, Boston, 1991.

[36] H.N. Mhaskar. *Weighted polynomial approximation.* World Scientific, 1996.

[37] P. Nevai. Géza Freud, orthogonal polynomials and Christoffel functions. A case study. *J. Approx. Theory*, 48:3–167, 1986.

[38] E.M. Nikishin and V.N. Sorokin. *Rational approximation and orthogonality*, volume 92 of *Transl. Math. Monographs.* Amer. Math. Soc., 1991.

[39] E.M. Nikišin. On simultaneous Padé approximants. *Math. Sb.*, 113:499–519, 1980. (Russian); *Math. USSR Sb.* **41** (1982), 409–425.

[40] R. Pintelon, Y. Rolain, A. Bultheel, and M. Van Barel. Numerically robust frequency domain identification of multivariable systems. In P. Sas and B. Van Hal, editors, *Proceedings of the 2002 International Conference on Modal Analysis, Noise and Vibration Engineering, Leuven 2002, September 16-18*, pages 1316–1321, 2002.

[41] K. Postelmans and W. Van Assche. Multiple little q-jacobi polynomials. *J. Comput. Appl. Math.*, 2004. To appear.

[42] A. Sinap. Gaussian quadrature, for matrix valued functions on the unit circle. *Electron. Trans. Numer. Anal.*, 3:95–115, 1995.

[43] A. Sinap and W. Van Assche. Gaussian quadrature for matrix valued functions on the real line. *J. Comput. Appl. Math.*, 65:369–385, 1995.

[44] A. Sinap and W. Van Assche. Orthogonal matrix polynomials and applications. *J. Comput. Appl. Math.*, 66:27–52, 1996.

[45] H. Stahl and V. Totik. *General orthogonal polynomials.* Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1992.

[46] G. Szegő. *Orthogonal polynomials*, volume 33 of *Amer. Math. Soc. Colloq. Publ.* Amer. Math. Soc., Providence, Rhode Island, 3rd edition, 1967. First edition 1939.

[47] W. Van Assche. Multiple orthogonal polynomials, irrationality and transcendence. In *Continued fractions: from analytic number theory to constructive approximation (Columbia, MO, 1998)*, volume 236 of *Contemp. Math.*, pages 325–342, 1999.

[48] W. Van Assche and E. Coussement. Some classical multiple orthogonal polynomials. *Numerical analysis 2000, Vol. V, Quadrature and orthogonal polynomials, J. Comput. Appl. Math.*, 127(1-2):317–347, 2001.

[49] W. Van Assche, J.S. Geronimo, and A.B.J. Kuijlaars. Riemann-Hilbert problems for multiple orthogonal polynomials. In J. Bustoz, M. Ismail, and S. Suslov, editors, *Special Functions 2000: Current Perspective and Future Directions*, volume 30 of *NATO Science Series II. Mathematics, Physics and Chemistry*, pages 23–59, Dordrecht, 2001. Kluwer.

[50] W. Van Assche and S.B. Yakubovich. Multiple orthogonal polynomials associuated with Macdonald functions. *Integral Transforms Special Funct.*, 9:229–244, 2000.

[51] M. Van Barel and A. Bultheel. Orthogonal polynomial vectors and least squares approximation for a discrete inner product. *Electron. Trans. Numer. Anal.*, 3:1–23, 1995.

[52] M. Van Barel and A. Bultheel. Look-ahead schemes for block Toeplitz systems and formal orthogonal matrix polynomials. In A. Draux and V. Kaliaguine, editors, *Orthogonal polynomials: the non-definite case. Proceedings workshop, Rouen, France, April 24-26, 1995*, Actes de l'atelier de Rouen, pages 93–112. INSA de Rouen, Lab. de Mathématique, 1997.

[53] M. Van Barel and A. Bultheel. Updating and downdating of orthogonal polynomial vectors and some applications. In V. Olshevsky, editor, *Structured Matrices in Mathematics, Computer Science, and Engineering II: Part II*, volume 281 of *Contemp. Math.*, pages 145–162. Amer. Math. Soc., 2001.

[54] M. Van Barel, G. Heinig, and P. Kravanja. An algorithm based on orthogonal polynomial vectors for Toeplitz least squares problems. In *Numerical Analysis and Its Applications*, volume 1988 of *Lecture Notes in Computer Science*, pages 27–34. Springer-Verlag, 2001.

[55] J. Van Deun and A. Bultheel. An interpolation algorithm for orthogonal rational functions. *J. Comput. Appl. Math.*, 2002. Accepted.

[56] J. Van Deun and A. Bultheel. Orthogonal rational functions and quadrature on an interval. *J. Comput. Appl. Math.*, 153(1-2):487–495, 2003.

[57] J. Van Deun and A. Bultheel. Ratio asymptotics for orthogonal rational functions on an interval. *J. Approx. Theory*, 123(2):162–172, 2003.

[58] J. Van Deun and A. Bultheel. A weak-star convergence result for orthogonal rational functions. *J. Comput. Appl. Math.*, 2003.

[59] P. Van gucht and A. Bultheel. A relation between orthogonal rational functions on the unit circle and the interval $[-1, 1]$. *Comm. Anal. Th. Continued Fractions*, 8:170–182, 2000.

[60] P. Zinn-Justin. Random Hermitian matrices in an external field. *Nuclear Phys. B*, 497:725–732, 1997.