RACSAM
Rev. R. Acad. Cien. Serie A. Mat.
VOL. **96** (2), 2002, pp. 157–173
Matemática Aplicada / Applied Mathematics
Artículo panorámico / Survey

An introduction to formal orthogonality and some of its applications

C. Brezinski

Abstract. This paper is an introduction to formal orthogonal polynomials and their applications to Padé approximation, Krylov subspace methods for the solution of systems of linear equations, and convergence acceleration methods. Some more general formal orthogonal polynomials, and the concept of biorthogonality and their applications are also discussed.

Introducción a la ortogonalidad formal y a algunas de sus aplicaciones

Resumen. Presentamos una introducción a los polinomios ortogonales formales y a sus aplicaciones a la aproximación de Padé, al método de subespacios de Krylov para la resolución de sistemas de ecuaciones lineales y a los métodos de aceleración de la convergencia. Se estudian también polinomios ortogonales formales más generales, el concepto de biortogonalidad y sus aplicaciones.

1. Introduction

Orthogonal polynomials are well–known and they are widely used in numerical analysis. In particular, they have applications in interpolation, approximation (including least squares and moment preserving splines), quadrature, acceleration of slowly convergent series, linear algebra, wavelets, methods for ordinary differential equations, Toda lattices, etc.; see [57].

In this paper, we will review some of the formal generalizations of orthogonal polynomials and discuss their applications to Padé approximation, Krylov subspace methods for the solution of systems of linear equations, and convergence acceleration methods.

A paper as this one could either be a survey, as complete and detailed as possible, or serve only as an introduction to the subject and a pointer to the literature. We chose this second option and tried to give an overview of the theory and its applications without entering into too many technical details. Citations to the literature have been kept to a minimum, that is to references which have an historical interest, to those which are not so well known, and to those where other references could be found. So, citations are not necessarily given to original works. For more references, see [13].

The reader is assumed to have some knowledge of usual orthogonal polynomials. For an introduction to the topic, see [57]. A more complete treatment is given in [44] and [92]. Chebyshev polynomials are studied in [82]. Orthogonal polynomials of a discrete variable are discussed in [76].

Presentado por Mariano Gasca.
Recibido: 4 de Abril de 2002. *Aceptado*: 5 de Junio de 2002.
Palabras clave / Keywords: orthogonal polynomials, biorthogonality, Padé approximation, Krylov subspace methods, convergence acceleration.
Mathematics Subject Classifications: 42C05, 41A21, 65F10, 65B05

2. Formal orthogonal polynomials

In this Section, formal orthogonal polynomials will be introduced and some of their properties will be presented.

For more results on formal orthogonal polynomials, see [9, Chap. 2].

2.1. Basic definitions

Let c be a linear functional on the vector space of complex polynomials. It is completely defined by its *moments*

$$c_i = c(x^i), \quad i = 0, 1, \ldots$$

The polynomials $\{P_k\}$ are said to form the family of *formal orthogonal polynomials* with respect to c if, for all k, the degree of P_k is at most k, and if it satisfies the *orthogonality conditions*

$$c(x^i P_k(x)) = 0, \quad i = 0, \ldots, k-1. \tag{1}$$

Their name comes out from the orthogonality property with respect to a linear form. For a rigorous justification of this notation and an extensive study of the algebra of linear functionals on the vector space of polynomials, see [69, 70, 71, 72]; see also [83].

Setting $P_k(x) = a_k + a_{k-1}x + \cdots + a_0 x^k$ (where the coefficients depend on k), these conditions are equivalent to

$$a_k c_i + a_{k-1} c_{i+1} + \cdots + a_0 c_{i+k} = 0, \quad i = 0, \ldots, k-1, \tag{2}$$

which is a system of k equations in $k+1$ unknowns. Its solution is completely determined once a supplementary condition has been added (assuming that the determinant of the augmented system is different from zero).

From (2), we have the determinantal representation

$$P_k(x) = D_k \begin{vmatrix} c_0 & \cdots & c_k \\ \vdots & & \vdots \\ c_{k-1} & \cdots & c_{2k-1} \\ 1 & \cdots & x^k \end{vmatrix} \tag{3}$$

where $D_k \neq 0$ is the *normalization factor* since, as explained above, the conditions (1) (or (2)) define P_k apart a multiplying factor. This normalization factor is determined by the supplementary condition added to the system (2), see [27].

We see that P_k has exact degree k if and only if the *Hankel determinant* $H_k(c_0)$ is different from zero, where, for an arbitrary sequence (u_n), we set (a notation useful in the sequel)

$$H_k(u_n) = \begin{vmatrix} u_n & \cdots & u_{n+k-1} \\ \vdots & & \vdots \\ u_{n+k-1} & \cdots & u_{n+2k-2} \end{vmatrix}.$$

If $D_k = 1/H_k(c_0)$, the polynomial P_k is called *monic*, which means that the coefficient of x^k equals 1. Once D_k has been chosen (or defined by the supplementary condition added to the system (2)), P_k is uniquely determined by (2) or (3). However, P_k may not exist for some choices of D_k (for example, the monic polynomial P_k does not exist if $H_k(c_0) = 0$). If not stated differently, we will assume in the sequel that $\forall k, H_k(c_0) \neq 0$. In that case, we say that c is *definite*, and we have $c(pP_k) \neq 0$ for any polynomial p of degree k.

Obviously, from (1), the orthogonality conditions can also be written as $c(pP_k) = 0$ where p is any polynomial of degree at most $k-1$, or as $c(P_n P_k) = 0$ for $n \neq k$.

If the linear functional c is defined by

$$c(x^i) = \int_a^b x^i \, d\alpha(x), \quad i = 0, 1, \ldots \tag{4}$$

where α is bounded and nondecreasing in $[a, b]$, the usual orthogonal polynomials with respect to α are recovered and $\forall k, P_k$ exists and has exact degree k. Moreover, for all k, $H_k(c_0) > 0$.

We will also consider the *associated polynomials* defined by

$$Q_k(z) = c\left(\frac{P_k(x) - P_k(z)}{x - z}\right),$$

where c acts on x, and z is a parameter. The polynomial Q_k has degree $k-1$ at most and it is given by

$$Q_k(z) = (c_0 a_{k-1} + c_1 a_{k-2} + \cdots + c_{k-1} a_0) + \cdots + (c_0 a_1 + c_1 a_0) z^{k-2} + c_0 a_0 z^{k-1}, \tag{5}$$

and it has the representation

$$Q_k(x) = D_k \begin{vmatrix} c_0 & c_1 & c_2 & \cdots & c_k \\ \vdots & \vdots & \vdots & & \vdots \\ c_{k-1} & c_k & c_{k+1} & \cdots & c_{2k-1} \\ 0 & c_0 & c_0 x + c_1 & \cdots & c_0 x^{k-1} + \cdots + c_{k-2} x + c_{k-1} \end{vmatrix}.$$

2.2. Properties

When they exist, formal orthogonal polynomials have the same algebraic properties as the usual orthogonal polynomials except some properties concerning their zeros [9, Chap. 2]. In particular, from the computational point of view, the most interesting property is their three–term recurrence relationship

Property 1

If, for all k, P_k exists and has exact degree k, then

$$P_{k+1}(x) = (A_{k+1} x + B_{k+1}) P_k(x) - C_{k+1} P_{k-1}(x), \quad k = 0, 1, \ldots$$

with $P_{-1}(x) = 0, P_0(x) = A_0 \neq 0$ and

$$A_{k+1} = \frac{t_{k+1}}{t_k}, \quad B_{k+1} = -\frac{\alpha_k t_{k+1}}{h_k t_k}, \quad C_{k+1} = \frac{t_{k-1} t_{k+1}}{t_k^2} \frac{h_k}{h_{k-1}},$$

where t_k is the coefficient of x^k in P_k, and

$$\alpha_k = c(x P_k^2), \quad h_k = c(P_k^2).$$

The associated polynomials Q_k satisfy the same recurrence relationship with the initializations $Q_{-1}(x) = -1, Q_0(x) = 0$, and with $C_1 = A_1 c(P_0)$.

The coefficient t_{k+1} is determined once the normalization factor has been chosen.

The case where some of the polynomials do not exist was extensively treated in [48]; see also [27].

A reciprocal of the preceding result is the Shohat–Favard theorem which says that if a family of polynomials satisfies a three–term recurrence of the form of that of Property 1, then they are orthogonal with respect to a linear functional whose moments can be computed. Moreover, if $\forall k, C_{k+1} > 0$, this functional can be represented as in (4).

A consequence of the three–term recurrence relationship is the Christoffel–Darboux identity

Property 2
If, for all k, P_k exists and has exact degree k, then

$$\frac{t_k}{h_k t_{k+1}}[P_{k+1}(x)P_k(t) - P_{k+1}(t)P_k(x)] = (x-t)\sum_{i=0}^{k} h_i^{-1} P_i(x)P_i(t).$$

This identity is usually proved from the recurrence relationship satisfied by the family $\{P_k\}$. However, a direct proof was given in [11]. It follows that the converse of this identity is also true: if a family of polynomials satisfies a relation of the form of the Christoffel–Darboux identity, then it also satisfies a three–term relationship and, so, it is a family of formal orthogonal polynomials [11]. The Christoffel–Darboux identity also holds for the polynomials Q_k, and there exists a variant mixing the polynomials and their associated ones.

We also have the

Property 3
If, for all k, P_k exists and has exact degree k, then

$$P_k(x)Q_{k+1}(x) - Q_k(x)P_{k+1}(x) = A_{k+1}h_k.$$

This relation shows that, if c is definite, then P_k and P_{k+1} have no common zero, and similarly for Q_k and Q_{k+1}, and for P_k and Q_k. This is the only property that can be proved on the zeros in this general context.

The sum in the right hand side of the Christoffel–Darboux identity is called the *reproducing kernel* of order k of the family of formal orthogonal polynomials

$$K_k(x,t) = \sum_{i=0}^{k} h_i^{-1} P_i(x)P_i(t).$$

It is a symmetric function in x and t, and the following property justifies its name

Property 4
For any polynomial p of degree at most k

$$c(p(x)K_k(x,t)) = p(t).$$

If the degree of p is strictly less that k, then $c(p(x)(x-t)K_k(x,t)) = 0$.

It follows that

$$\begin{array}{ll} c(K_k(x,t)P_n(x)) = P_n(t), & n = 0,\ldots,k, \\ c(K_k(x,t)K_n(x,u)) = K_n(t,u), & n = 0,\ldots,k, \\ c(K_k^2(x,t)) = K_k(t,t). & \end{array}$$

We also have

$$H_{k+1}(c_0)K_k(x,t) = \begin{vmatrix} c_0 & c_1 & \cdots & c_k & 1 \\ c_1 & c_2 & \cdots & c_{k+1} & t \\ \vdots & \vdots & & \vdots & \vdots \\ c_k & c_{k+1} & \cdots & c_{2k} & t^k \\ 1 & x & \cdots & x^k & 0 \end{vmatrix}.$$

On reproducing kernel and their applications, see [97].

2.3. Adjacent families

We consider the linear functionals $c^{(n)}$, $n = 0, 1, \ldots$, defined by

$$c^{(n)}(x^i) = c_{n+i}, \quad i = 0, 1, \ldots$$

We assume that, for all n, $c^{(n)}$ is definite. Let $\{P_k^{(n)}\}$ be the family of formal orthogonal polynomials with respect to $c^{(n)}$, that is such that

$$c^{(n)}(x^i P_k^{(n)}(x)) = 0, \quad i = 0, \ldots, k-1.$$

Thus the polynomials $P_k^{(0)}$ are identical to the polynomials P_k defined above. These families of polynomials are called *adjacent families* of formal orthogonal polynomials. The polynomials of each of these families satisfy a three–term recurrence relationship similar to that of Property 1 (assuming that, for all n, $c^{(n)}$ is definite). Moreover, there exists recurrence relations between polynomials belonging to adjacent families. For example, we have, for monic polynomials,

$$\begin{aligned} P_k^{(n+1)}(x) &= P_k^{(n)}(x) - e_k^{(n)} P_{k-1}^{(n+1)}(x), \\ P_{k+1}^{(n)}(x) &= x P_k^{(n+1)}(x) - q_{k+1}^{(n)} P_k^{(n)}(x) \end{aligned}$$

with

$$\begin{aligned} e_k^{(n)} &= \frac{H_{k-1}(c_{n+1}) H_{k+1}(c_n)}{H_k(c_n) H_k(c_{n+1})} = \frac{c^{(n)}(x^k P_k^{(n)}(x))}{c^{(n+1)}(x^{k-1} P_{k-1}^{(n+1)}(x))}, \\ q_{k+1}^{(n)} &= \frac{H_{k+1}(c_{n+1}) H_k(c_n)}{H_{k+1}(c_n) H_k(c_{n+1})} = \frac{c^{(n+1)}(x^k P_k^{(n+1)}(x))}{c^{(n)}(x^k P_k^{(n)}(x))}. \end{aligned}$$

These quantities are related by the qd–algorithm

$$\begin{aligned} q_{k+1}^{(n)} + e_{k+1}^{(n)} &= q_{k+1}^{(n+1)} + e_k^{(n+1)} \\ e_k^{(n)} q_{k+1}^{(n)} &= e_k^{(n+1)} q_k^{(n+1)} \end{aligned}$$

with $\forall n, q_0^{(n)} = e_0^{(n)} = 0$ and $q_1^{(n)} = c_{n+1}/c_n$. Moreover, by elimination, the three–term recurrence relation of Property 1 can be recovered, and it can be written as

$$P_{k+1}^{(n)}(x) = (x - q_{k+1}^{(n)} - e_k^{(n)}) P_k^{(n)}(x) - q_k^{(n)} e_k^{(n)} P_{k-1}^{(n)}(x), \quad k = 0, 1, \ldots$$

with $P_{-1}^{(n)}(x) = 0$ and $P_0^{(n)}(x) = 1$. A similar relation is satisfied by the associated polynomials

$$Q_k^{(n)}(z) = c^{(n)} \left(\frac{P_k^{(n)}(x) - P_k^{(n)}(z)}{x - z} \right),$$

with $Q_{-1}^{(n)}(z) = -1$ and $Q_0^{(n)}(z) = 0$.

The qd–algorithm was introduced by Rutishauser [84] and it gave rise to the LR–algorithm for the computation of the eigenvalues of a matrix [85]. It is strongly connected to Padé approximants, continued fractions and the ε–algorithm that will be studied in Section 3., see [61, pp. 608–621], [9, pp. 83–84, 94–97, 148–151, 161, 166-167].

3. Applications

We will now have a look at some topics in numerical analysis where formal orthogonal polynomials play a central role. Another application, which is not discussed in this paper, concerns methods for computing all zeros of analytic functions [64].

3.1. Padé approximation

We consider the formal power series

$$f(z) = c_0 + c_1 z + c_2 z^2 + \cdots$$

A *Padé approximant* of f is a rational function whose expansion in ascending powers of z agrees with f as far as possible. More precisely, the Padé approximant denoted by $[p/q]_f(z)$ is a rational function with a numerator N of degree at most p and a denominator D of degree at most q such that

$$N(z) - f(z)D(z) = \mathcal{O}(z^{p+q+1}), \quad (z \to 0). \tag{6}$$

These approximants are put into a double entry array called the *Padé table*.

The coefficients of the polynomials N and D can be obtained by identification of the coefficients of the powers of z in both sides of (6). Let us consider, for simplicity, the Padé approximant $[k-1/k]$. We set

$$\begin{aligned} N(z) &= b_0 + b_1 z + \cdots + b_{k-1} z^{k-1} \\ D(z) &= a_0 + a_1 z + \cdots + a_k z^k. \end{aligned}$$

By identification of the coefficients, we obtain

$$\begin{array}{llllllll} \text{degree } 0 & \Longrightarrow & b_0 & = & c_0 a_0 & & & \\ \text{degree } 1 & \Longrightarrow & b_1 & = & c_0 a_1 & + & c_1 a_0 & \\ & & & \vdots & & & & \\ \text{degree } k-1 & \Longrightarrow & b_{k-1} & = & c_0 a_{k-1} & + & \cdots & + \; c_{k-1} a_0 \\ \text{degree } k & \Longrightarrow & 0 & = & c_0 a_k & + & \cdots & + \; c_k a_0 \\ & & & \vdots & & & & \\ \text{degree } 2k-1 & \Longrightarrow & 0 & = & c_{k-1} a_k & + & \cdots & + \; c_{2k-1} a_0. \end{array}$$

So, comparing the last k equations with (2), we see that $D(z) = \widetilde{P}_k(z) = z^k P_k(z^{-1})$ where P_k is the monic formal orthogonal polynomial of degree at most k with respect to the linear functional c defined by $c(x^i) = c_i$ for $i = 0, 1, \ldots$. Moreover, the first k preceding equations show that $N(z) = \widetilde{Q}_k(z) = z^{k-1} Q_k(z^{-1})$ where Q_k is the associated polynomial given by (5). From the determinantal expressions of P_k and Q_k we see that $[k-1/k]$ can be written as the ratio of two determinants and, more generally, we have

$$[p/q]_f(z) = \begin{vmatrix} z^q f_{p-q}(z) & z^{q-1} f_{p-q+1}(z) & \cdots & f_p(z) \\ c_{p-q+1} & c_{p-q+2} & \cdots & c_{p+1} \\ \vdots & \vdots & & \vdots \\ c_p & c_{p+1} & \cdots & c_{p+q} \end{vmatrix} \Bigg/ \begin{vmatrix} z^q & z^{q-1} & \cdots & 1 \\ c_{p-q+1} & c_{p-q+2} & \cdots & c_{p+1} \\ \vdots & \vdots & & \vdots \\ c_p & c_{p+1} & \cdots & c_{p+q} \end{vmatrix},$$

with $c_i = c(x^i) = 0$ for $i < 0$, $f_k(z) = c_0 + \cdots + c_k z^k$ for $k \geq 0$, and $f_k(z) = 0$ for $k < 0$.

This determinantal expression shows that the other Padé approximants $[p/q]$ are related to the adjacent families of formal orthogonal polynomials. The relations between these polynomials can be used for the recursive computation of any sequence of Padé approximants as described in [9, pp. 135–147]. When some of the Hankel determinants $H_k(c_n)$ are zero, some adjacent approximants in the Padé table become identical by cancellation of a common factor between the numerator and the denominator. This phenomenon is known as the *block structure* of the Padé table [79]. In that case, the recurrence relationships have to be replaced by more complicated ones. This problem was completely solved in [48].

Formal orthogonal polynomials allow to answer several other questions related to Padé approximants. An expression for the relative error $(f(z) - [k-1/k]_f(z))/f(z)$ was given in [24]. However, in practice, it is often necessary to estimate the error. Such estimates are based on the interpretation of Padé approximants as formal Gaussian quadratures. Indeed, it formally holds (c acts on x)

$$f(z) = c\left(\frac{1}{1-xz}\right).$$

As in a Gaussian quadrature method, let us replace the function $1/(1-xz)$ by its Hermite interpolation polynomial R_k at the zeros of P_k given by

$$R_k(x) = \frac{1}{1-xz}\left(1 - \frac{P_k(x)}{P_k(z^{-1})}\right).$$

Then, it holds [38, 39]

$$[k-1/k]_f(z) = c(R_k).$$

So, Padé approximants can be seen as formal Gaussian quadratures. For usual definite integrals, the error of a Gaussian quadrature method can be estimated by Kronrod procedure (see [74] for a review). It consists in comparing the results obtained by the Gaussian quadrature with another one, built on the same nodes (the k zeros of P_k) plus $k+1$ additional nodes chosen in an optimal way (that is so that the new quadrature formula be exact for polynomials of highest possible degree). These additional nodes are the zeros of the *Stieltjes polynomial* V_{k+1} which is defined by

$$c(x^i P_k(x) V_{k+1}(x)) = 0, \quad i = 0, \ldots, k.$$

Kronrod's idea can be extended to Padé approximants since they are Gaussian quadratures. Several procedures for estimating the error in Padé approximation are based on this remark [38]. Formal Stieltjes polynomials are studied in [81].

The most important and difficult problem about Padé approximants is the convergence, in some sense, of a given sequence of approximants to the function f. This problem strongly depends on the asymptotic behavior of formal orthogonal polynomials, a question mostly studied in the case where a representation of the form (4) holds [92, 90]. An overview can be found in [38, pp. 93–129]. Let us mention that counterexamples to the Baker–Gammel–Wills conjecture [3] (see also [4, p. 332]) on the uniform convergence of a subsequence of diagonal approximants for a function meromorphic in the unit disc [68] or holomorphic [43] were recently obtained. In some cases, Padé approximants can be used for the analytic continuation of divergent series, see [6, pp. 383–410] and [98].

On Padé approximants and their connections to formal orthogonal polynomials, see [4, 9, 23, 38, 39]. Padé approximants are related to continued fractions (see [63, 67]), one of the oldest topics in mathematics [12]. Padé approximants have many applications in numerical analysis, special functions, number theory, applied mathematics, physics, chemistry, etc. Some of them are described in [4, 38, 39].

3.2. Krylov subspace methods

We consider the system of (real, for simplicity) linear equations

$$A\widetilde{x} = b.$$

Lanczos method [66] for solving this system consists in constructing a sequence of vectors (x_k) defined by the two conditions

$$\begin{aligned} x_k - x_0 &\in K_k(A, r_0) = \text{span}(r_0, Ar_0, \ldots, A^{k-1}r_0), \\ r_k = b - Ax_k &\perp K_k(A^T, y) = \text{span}(y, A^T y, \ldots, (A^T)^{k-1} y), \end{aligned}$$

where x_0 is arbitrarily chosen, $r_0 = b - Ax_0$, and $y \neq 0$ is a given vector. A subspace of the form K_k is called a *Krylov* subspace and Lanczos method belongs to the class of *Krylov subspace methods*, a particular case of projection methods [17, 56, 60].

The vectors x_k are completely determined by the two preceding conditions. Indeed, the first condition writes

$$x_k - x_0 = -a_{k-1}r_0 - a_{k-2}Ar_0 - \cdots - a_0A^{k-1}r_0.$$

Multiplying both sides by A, adding and subtracting b, leads to

$$r_k = r_0 + a_{k-1}Ar_0 + \cdots + a_0A^k r_0 = P_k(A)r_0$$

with $P_k(x) = 1 + a_{k-1}x + \cdots + a_0x^k$. The second condition gives

$$(r_k, (A^T)^i y) = (P_k(A)r_0, (A^T)^i y) = (A^i P_k(A)r_0, y) = 0, \quad i = 0, \ldots, k-1,$$

that is

$$(A^i r_0 + a_{k-1}A^{i+1}r_0 + \cdots + a_0A^{i+k}r_0, y) = 0, \quad i = 0, \ldots, k-1.$$

Let c be the linear functional defined by

$$c(x^i) = (A^i r_0, y).$$

Thus, since c is linear, $c(p) = (y, p(A)r_0)$ for any polynomial p, and the preceding conditions are equivalent to

$$c_i + a_{k-1}c_{i+1} + \cdots + a_0c_{i+k} = 0, \quad i = 0, \ldots, k-1,$$

which is exactly (2). So, these conditions can be written as

$$c(x^i P_k(x)) = 0, \quad i = 0, \ldots, k-1,$$

which shows that, if it exists (that is, if $H_k(c_1) \neq 0$), P_k is the formal orthogonal polynomial of degree at most k with respect to c, normalized by the condition $P_k(0) = 1$. Thus, it follows

$$r_k = \begin{vmatrix} r_0 & Ar_0 & \cdots & A^k r_0 \\ c_0 & c_1 & \cdots & c_k \\ \vdots & \vdots & & \vdots \\ c_{k-1} & c_k & \cdots & c_{2k-1} \end{vmatrix} \Big/ H_k(c_1),$$

where the determinant in the numerator is the linear combination of the vectors in its first row, obtained by the usual rules for expanding a determinant.

Thanks to the recurrence relationships between adjacent families of formal orthogonal polynomials, the residuals r_k and the iterates x_k can be recursively computed, thus leading to several algorithms for implementing Lanczos method [36]. Among these procedures is the well–known biconjugate gradient algorithm [55] which reduces to the conjugate gradient algorithm when A is symmetric and $y = r_0$ [62], see also [53]. New algorithms can also be deduced [36].

Due to the normalization of the formal orthogonal polynomials involved in Lanczos method, some underlying Hankel determinants can be zero. Moreover, some recurrence relations may be impossible to use. In these cases, a division by zero, called a *breakdown*, arises in the algorithms. Such breakdowns can be avoided by jumping over the polynomials which do not exist and/or over those that cannot be computed by the recurrence relationship under consideration. In this case, more complicated recurrences, based on those given in [48], have to be used. The problem of *near–breakdown*, due to division by a quantity close to zero (which amplifies rounding errors), can be treated by similar techniques; see [33] and the references quoted there. Of course, a linear algebra problem can be studied by purely linear algebra techniques. However, as shown in [34], the theory of formal orthogonal polynomials greatly simplifies the treatment of breakdowns and near–breakdowns. Breakdowns can also be cured using the connection between Lanczos method and Padé approximants [59].

Other Krylov subspace methods can be related to formal orthogonality [23, pp. 164–168], for example, the CGS [88] and the BiCGSTAB [93]. These two methods, and others, were created to avoid matrix–vector products by A^T, a nontrivial task due to the indirect addressing used for storing large sparse matrices [31]. A multiparameter generalization of Lanczos method [20], which can be implemented by an algorithm similar to the block conjugate gradient algorithm [78], can also be interpreted in terms of formal orthogonal polynomials [21]. Breakdowns and near–breakdowns in the corresponding algorithms can be handled by techniques issued from formal orthogonal polynomials [35].

Formal orthogonal polynomials are also involved in Lanczos tridiagonalization (or biorthogonalization) method [65] for the computation of eigenvalues of a matrix [32]. This method, which is related to Padé approximants, has applications for model reduction in linear control theory [2, 52].

3.3. Convergence acceleration

In numerical analysis, many algorithms produce sequences converging to a limit. When the convergence is too slow, the sequence can be transformed, by a *sequence transformation*, into another one converging faster to the same limit (under some conditions). The most well–known sequence transformation is Aitken Δ^2 process. It was generalized by Shanks [87]. *Shanks transformation* consists in transforming a sequence (S_n) into a set of sequences (indexed either by k or by n) defined by

$$e_k(S_n) = H_{k+1}(S_n)/H_k(\Delta^2 S_n), \quad k, n = 0, 1, \ldots \tag{7}$$

These quantities can be recursively computed by the ε–algorithm of Wynn [100]. We set $\varepsilon_{-1}^{(n)} = 0, \varepsilon_0^{(n)} = S_n, n = 0, 1, \ldots$ The rules of the ε–algorithm are

$$\varepsilon_{k+1}^{(n)} = \varepsilon_{k-1}^{(n+1)} + \frac{1}{\varepsilon_k^{(n+1)} - \varepsilon_k^{(n)}}, \quad k, n = 0, 1, \ldots$$

and it holds

$$\varepsilon_{2k}^{(n)} = e_k(S_n), \quad k, n = 0, 1, \ldots$$

When $k = 1$, Aitken process is recovered.

An important concept about a sequence transformation is its *kernel*. It is the set of sequences (S_n) which are transformed into a constant sequence (S). The kernel of Shanks transformation is the set of sequences satisfying, for all n,

$$a_0(S_n - S) + \cdots + a_k(S_{n+k} - S) = 0,$$

with $a_0 a_k \neq 0$ and $a_0 + \cdots + a_k \neq 0$. It is a necessary and sufficient condition for having $e_k(S_n) = S$, $n = 0, 1, \ldots$, [26, pp. 79–80].

Shanks transformation and the ε–algorithm are related to Padé approximants and, thus, to formal orthogonal polynomials. Indeed, if $S_n = c_0 + c_1 z + \cdots + c_n z^n$ is the nth partial sum of a series f, then $\varepsilon_{2k}^{(n)} = e_k(S_n) = [n + k/k]_f(z)$.

More generally, let $P_k^{(n+1)}(x) = x^k \widetilde{P}_k^{(n+1)}(x^{-1})$ be the orthogonal polynomial with respect to the linear functional $c^{(n+1)}$ defined from the sequence of moments $c_0 = S_0$ and $c_i = \Delta S_{i-1}$ for $i \geq 1$, and $Q_k^{(n+1)}(x) = x^{k-1} \widetilde{Q}_k^{(n+1)}(x^{-1})$ be its associated polynomial. Then

$$\varepsilon_{2k}^{(n)} = S_n + \widetilde{Q}_k^{(n+1)}(1)/\widetilde{P}_k^{(n+1)}(1),$$

An error formula can be derived from this interpretation, and Kronrod procedure can be applied for estimating the error $\varepsilon_{2k}^{(n)} - S$ [9, pp. 171–176].

Shanks transformation can also be related to the formal orthogonal polynomials with respect to the functional defined by $c_i = \Delta S_i$, and normalized by the condition $P_k^{(n)}(1) = 1$. Indeed, from (7), we have

$$e_k(S_n) = \begin{vmatrix} S_n & \cdots & S_{n+k} \\ \Delta S_n & \cdots & \Delta S_{n+k} \\ \vdots & & \vdots \\ \Delta S_{n+k-1} & \cdots & \Delta S_{n+2k-1} \end{vmatrix} \Bigg/ \begin{vmatrix} 1 & \cdots & 1 \\ \Delta S_n & \cdots & \Delta S_{n+k} \\ \vdots & & \vdots \\ \Delta S_{n+k-1} & \cdots & \Delta S_{n+2k-1} \end{vmatrix}. \qquad (8)$$

If we define the linear functional s by $s(x^i) = S_i$, then we see that

$$e_k(S_n) = s(x^n P_k^{(n)}(x)),$$

where $\{P_k^{(n)}\}$ is the family of formal orthogonal polynomials with respect to the linear functional $c^{(n)}(x^i) = c_{n+i} = \Delta S_{n+i}$. This point of view allows to obtain several algorithms for the implementation of Shanks transformation and the computation of Padé approximants [8].

The G–algorithm, another convergence acceleration method, is also related to formal orthogonal polynomials. This method consists in replacing ΔS_{n+i} in (8) by g_{n+i} for $i = 0, \ldots, 2k-1$, where (g_n) is a given auxiliary sequence of parameters, see [9, p. 169], [26, pp. 95–101, 224, 264] and [28].

The topological ε–algorithm is a vector generalization of the ε–algorithm. It is related to formal orthogonal polynomials, see [9, pp. 178–184] and [26, pp. 220–227]. It also has connections with Lanczos method for a system of linear equations [9, pp. 184–189].

On convergence acceleration methods, consult [26, 98, 99]. The recent history of the domain is reported in [22].

4. Biorthogonal polynomials

If we define the linear functionals L_i by $L_i(x^j) = c_{i+j}$, $i, j = 0, 1, \ldots$, then the orthogonality conditions (1) for the polynomial P_k write $L_i(P_k) = 0$ for $i = 0, \ldots, k-1$. In that case, the linear functionals L_i are related by the condition $L_i(x^{j+1}) = L_{i+1}(x^j)$.

Let us now consider the more general case where the functionals L_i are defined by

$$L_i(x^j) = c_{ij}, \quad i, j = 0, 1, \ldots$$

where the c_{ij}'s are given complex numbers. Let $\{P_k\}$ be the family of polynomials such that

$$L_i(P_k) = 0, \quad i = 0, \ldots, k-1.$$

These polynomials are called *biorthogonal* with respect to the family of linear functionals $\{L_i\}$ [14, 104–113]. We have

$$P_k(x) = D_k \begin{vmatrix} c_{00} & \cdots & c_{0k} \\ \vdots & & \vdots \\ c_{k-1,0} & \cdots & c_{k-1,k} \\ 1 & \cdots & x^k \end{vmatrix}.$$

The recursive computation of these polynomials was considered in [45, 46], and properties of their zeros were discussed in [30]. Particular cases are orthogonal polynomials on an algebraic variety [16], and orthogonal polynomials involving a Sobolev–type inner product [51].

The concept of biorthogonality was introduced by Banach [5, Chap. VII]. Let E be a vector space and E^* its dual (the vector space of linear functionals on E). The families $\{x_i\}$ of linearly independent elements of E, and $\{L_i\}$ of linearly independent elements of E^* are said to be biorthogonal (or forming *dual basis*)

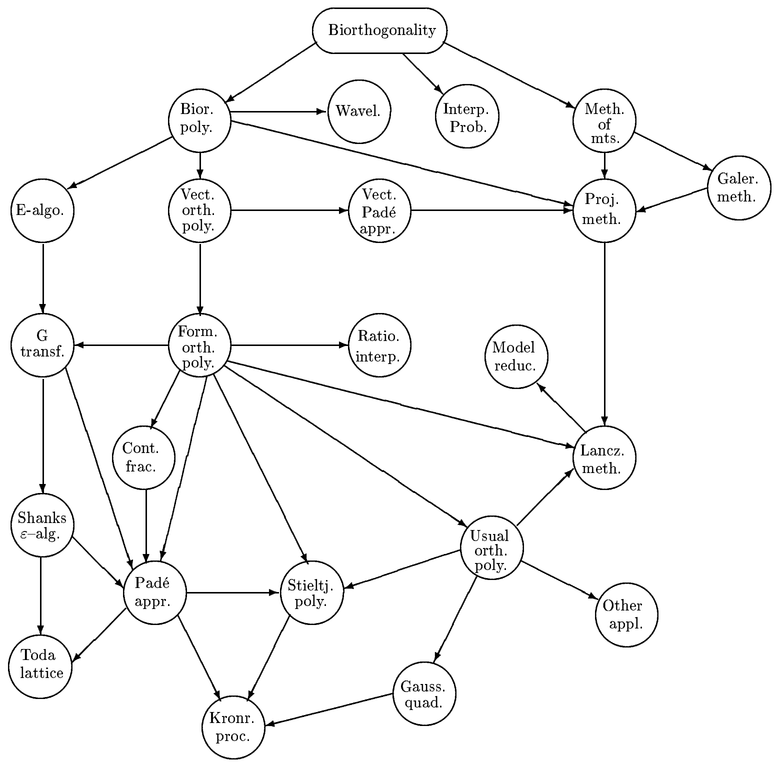

Figure 1. Biorthogonality and its applications.

if $L_i(x_j) =< L_i, x_j >= 0$ if $i \neq j$. Biorthogonality is connected to many topics in numerical analysis and it has many applications as shown in Figure 1 (updated from [17, p. 63]) . Let us comment on it.

The general interpolation problem, as stated in [47, pp. 26–45], enters into this framework. For rational interpolation, see [50, 58]. The E–algorithm, which is the most general convergence acceleration algorithm known so far, and other scalar and vector sequence transformations are related to biorthogonal polynomials [14, pp. 79–92]. Vector and matrix sequence transformations can be defined from biorthogonality [18, 29], in particular the RPA and the H–algorithm [26, pp. 233–244] which have applications in the solution of systems of linear and nonlinear equations. The connection to the method of moments of Vorobyev [96] is explained in [17, pp. 154ff.]. Biorthogonality have applications to multistep methods for ordinary differential equations [14, pp. 92–97], and to some topics in statistics and least squares [14, pp. 114–126]. Biorthogonal polynomials lead to a unified framework for the derivation of conjugate–gradient type algorithms for the solution of systems of linear equations [17, Chap. 4]. Their connection to subspace methods and Padé approximants was described in [19], see also [42]. They also have applications in the solution of linear systems by the bordering method for matrices with a special structure [17, pp. 70–81]. Biorthogonal polynomials are instrumental in the definition of Padé approximants for series of functions [14, pp. 97–104], and other generalizations [77]. Finally, biorthogonal polynomials were recently connected to the relativistic Toda molecule equation [73]. Integrable lattices and discrete soliton equations also have intriguing relationships with convergence acceleration methods [75, 80].

The concept of biorthogonality was extensively studied in [14]. References to the related domains of Figure 1 can be traced back from those given at the end of this paper. Obviously, it is not possible, in this paper, to enter into the details of these topics. However, let us mention some particular cases related to the applications described in Section 3.

Vector orthogonal polynomials of dimension $d > 0$ were introduced in [95]. They correspond to the case where the linear functionals L_i are related by

$$L_i(x^{j+1}) = L_{i+d}(x^j),$$

that is

$$L_i(x^{j+n}) = L_{i+nd}(x^j).$$

Such polynomials satisfy a $d+2$ term recurrence relationship of the form

$$P_{k+1}(x) = (A_{k+1}x + B_{k+1})P_k(x) - \sum_{i=1}^{d} C_{k+1}^{(i)} P_{k-i}(x), \quad k = 0, 1, \ldots.$$

They also satisfy a Shohat–Favard type theorem, an identity similar to Christoffel–Darboux's, and their reciprocals. They have applications in vector Padé approximation for approximating simultaneously a series with vector coefficients by several rational functions with a common denominator [94], see also [38, pp. 169–176] and [39, pp. 81–87]. When $d = 1$, the formal orthogonal polynomials discussed above are recovered.

Vector orthogonal polynomials of dimension $-d < 0$ can also be defined [15, 37]. In that case, the linear functionals L_i are related by

$$L_i(x^{j+1}) = L_{i-d}(x^j),$$

and the polynomials satisfy a $d+2$ term recurrence relationship of the form

$$P_{k+1}(x) = (A_{k+1}x + B_{k+1})P_k(x) - x\sum_{i=1}^{d} C_{k+1}^{(i)} P_{k-i}(x), \quad k = 0, 1, \ldots.$$

They also satisfy a Shohat–Favard type theorem. Formal orthogonal polynomials of dimension -1 can be considered as a generalization of orthogonal polynomials on the unit circle, and they have applications in Padé approximation of Laurent series [40, 41]. The case of multipoint Padé approximants is also of interest [38].

Orthogonal polynomials with (rectangular) matrix coefficients were also studied. They have applications in matrix Padé approximation and matrix continued fractions [89].

Orthogonal polynomials in several variables, and their applications to Padé approximation for multivariate series, are discussed, for example, in [7]. On these polynomials, see [49, 91].

Formal orthogonal polynomials in the least squares sense can be defined as the solution in the least squares sense of the system (2) for $i = 0, \ldots, m-1$ with $m \geq k$ [25]. These polynomials are used to construct least–squares Padé approximants which are less sensitive than Padé approximants to perturbations in the coefficients of the series to be approximated.

Many projection methods for the solution of systems of linear equations are related to biorthogonal polynomials as explained in [17, pp. 141–154]. Among them, are generalizations of Lanczos method [17, pp. 171–180], the method of Arnoldi (FOM) [1], and GMRES of Saad and Schultz [86]; see [19].

As mentioned above, the E–algorithm is the most general sequence transformation known so far. Its kernel is the set of sequences of the form

$$S_n = S + a_1 g_1(n) + \cdots + a_k g_k(n), \quad n = 0, 1, \ldots,$$

where the $(g_i(n))$ are auxiliary known sequences. The new sequences obtained by this transformation are defined by

$$E_k^{(n)} = \begin{vmatrix} S_n & \cdots & S_{n+k} \\ g_1(n) & \cdots & g_1(n+k) \\ \vdots & & \vdots \\ g_k(n) & \cdots & g_k(n+k) \end{vmatrix} \Bigg/ \begin{vmatrix} 1 & \cdots & 1 \\ g_1(n) & \cdots & g_1(n+k) \\ \vdots & & \vdots \\ g_k(n) & \cdots & g_k(n+k) \end{vmatrix}.$$

Let s be the linear functional defined by $s(x^i) = S_i$. We see that $E_k^{(n)} = s(x^n P_k^{(n)}(x))$, where $P_k^{(n)}$ is the formal biorthogonal polynomial, normalized by the condition $P_k^{(n)}(1) = 1$, such that $L_i(x^n P_k^{(n)}(x)) = 0$ for $i = 0, \ldots, k-1$, where $L_i(x^j) = g_{i+1}(j)$.

The preceding ratios of determinants can be recursively computed by the E–algorithm

$$\begin{aligned} E_{k+1}^{(n)} &= E_k^{(n)} - \frac{E_k^{(n+1)} - E_k^{(n)}}{g_{k-1,k}^{(n+1)} - g_{k-1,k}^{(n)}} g_{k-1,k}^{(n)} \\ g_{k+1,i}^{(n)} &= g_{k,i}^{(n)} - \frac{g_{k,i}^{(n+1)} - g_{k,i}^{(n)}}{g_{k-1,k}^{(n+1)} - g_{k-1,k}^{(n)}} g_{k-1,k}^{(n)} \end{aligned}$$

with $E_0^{(n)} = S_n$ and $g_{0,i}^{(n)} = g_i(n)$. Shanks transformation is recovered for the choice $g_i(n) = \Delta S_{n+i-1}$, and the G–transformation corresponds to $g_i(n) = g_{n+i-1}$. On the E–algorithm, see [26, pp. 55–72].

As seen above, ratios of determinants play an important role in biorthogonality where they often appear. Such ratios are related to *Schur complements*, as explained in [10]. It has been proved that quantities computed by a triangular recursive scheme as those used above can be expressed as ratios of determinants, and vice versa, and as a contour integral [26, pp. 21–26]. This framework also includes B–splines, Bernstein polynomials, generalized divided differences, etc., which opens new territories for biorthogonality.

Formal orthogonality may have other applications, for example, in wavelets [54].

Acknowledgement. I would like to thank Michela Redivo Zaglia for several pertinent comments, and Mariano Gasca who provided me the opportunity of writing this paper.

References

[1] Arnoldi, W. E.(1951). The principle of minimized iterations in the solution of the matrix eigenvalue problem, *Quart. Appl. Math.*, **9**, 17–29.

[2] Bai, Z. J. and Freund, R. W.(2001). A partial Padé–via–Lanczos method for reduced–order modeling, *Linear Algebra Appl.*, **332**, 139–164.

[3] Baker Jr., G. A., Gammel, J. L. and Wills, J. G. (1961). An investigation of the applicability of the Padé approximant method, *J. Math. Anal. Appl.*, **2**, 405–418.

[4] Baker Jr., G. A. and Graves–Morris, P. R. (1996). *Padé Approximants*, 2nd edition, Cambridge University Press, Cambridge.

[5] Banach, S. (1932). *Théorie des Opérations Linéaires*, Monografie Matematyczne, Vol. 1, Warzawa.

[6] Bender, C. M. and Orszag, S. A. (1978). *Advanced Mathematical Methods for Scientists and Engineers*, McGraw–Hill, New York.

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

[8] Brezinski, C. (1976). Computation of Padé approximants and continued fractions, *J. Comput. Appl. Math.*, **2**, 113–123.

[9] Brezinski, C. (1980). *Padé–Type Approximation and General Orthogonal Polynomials*, Birkhäuser, Basel.

[10] Brezinski, C. (1988). Other manifestations of the Schur complement, *Linear Algebra Appl.*, **111**, 231–247.

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

[12] Brezinski, C. (1991). *History of Continued Fractions and Padé Approximants*, Springer–Verlag, Berlin.

[13] Brezinski, C. (1991). *A Bibliography on Continued Fractions, Padé Approximation, Sequence Transformation, and Related Subjects*, Prensas Universitarias de Zaragoza, Zaragoza.

[14] Brezinski, C. (1992). *Biorthogonality and its Applications to Numerical Analysis*, Marcel Dekker, New York.

[15] Brezinski, C. (1992). A unified approach to various orthogonalities, *Ann. Fac. Sci. Toulouse, sér. 3*, **1**, (3), 277 – 292.

[16] Brezinski, C. (1995). Formal orthogonality on an algebraic curve, *Annals Numer. Math.*, **2**, 21–33.

[17] Brezinski, C. (1997). *Projection Methods for Systems of Equations*, North–Holland, Amsterdam.

[18] Brezinski, C. (1998). Vector sequence transformations: methodology and applications to linear systems, *J. Comput. Appl. Math.*, **98**, 149–175.

[19] Brezinski, C. (1999). Krylov subspace methods, biorthogonal polynomials and Padé–type approximants, *Numer. Algorithms*, **21**, 97–107.

[20] Brezinski, C. (1999). Multiparameter descent methods, *Linear Algebra Appl.*, **296**, 113–141.

[21] Brezinski, C. (2000). The block Lanczos and Vorobyev methods, *C.R. Acad. Sci. Paris, Série I*, **331**, 137–142.

[22] Brezinski, C. (2000). Convergence acceleration during the 20th century, *J. Comput. Appl. Math.*, **122**, 1–21.

[23] Brezinski, C. (2002). *Computational Aspects of Linear Control*, Kluwer, Dordrecht.

[24] Brezinski, C. and Maroni, P. (1996). The algebra of linear functionals on polynomials, with applications to Padé approximation, *Numer. Algorithms*, **11**, 25–33.

[25] Brezinski, C. and Matos, A. C. (1993). Least–squares orthogonal polynomials, *J. Comput. Appl. Math.*, **46**, 229–239.

[26] Brezinski, C. and Redivo Zaglia, M. (1991). *Extrapolation Methods. Theory and Practice*, North–Holland, Amsterdam.

[27] Brezinski, C. and Redivo Zaglia, M. (1994). Breakdowns in the computation of orthogonal polynomials, in *Nonlinear Numerical Methods and Rational Approximation, II*, A. Cuyt ed., Kluwer, Dordrecht, 49–59.

[28] Brezinski, C. and Redivo Zaglia, M. (1996). A look–ahead strategy for the implementation of some old and new extrapolation methods, *Numer. Algorithms*, **11**, 35–55.

[29] Brezinski, C. and Redivo Zaglia, M. (1996). Vector and matrix sequence transformations based on biorthogonality, *Appl. Numer. Math.*, **21**, 353–373.

[30] Brezinski, C. and Redivo Zaglia, M. (1997). On the zeros of various kinds of orthogonal polynomials, *Annals Numer. Math.*, **4**, 67–78.

[31] Brezinski, C. and Redivo Zaglia, M. (1998). Transpose–free Lanczos–type algorithms for nonsymmetric linear systems, *Numer. Algorithms*, **17**, 67–103.

[32] Brezinski, C. and Redivo Zaglia, M. (2000). Variations on Lanczos' tridiagonalization process, *Calcolo*, **37**, 159–179.

[33] Brezinski, C., Redivo Zaglia, M. and Sadok, H. (1999). New look–ahead Lanczos–type algorithms for linear systems, *Numer. Math.*, **83**, 53–85.

[34] Brezinski, C., Redivo Zaglia, M. and Sadok, H. (2000). The matrix and polynomial approaches to Lanczos–type algorithms, *J. Comput. Appl. Math.*, **123**, 241–260.

[35] Brezinski, C., Redivo Zaglia, M. and Sadok, H. (2002). A review of formal orthogonality in Lanczos–based methods, *J. Comput. Appl. Math.*, **140**, 81–98.

[36] Brezinski, C. and Sadok, H. (1993). Lanczos type methods for solving systems of linear equations, *Appl. Numer. Math.*, **11**, 443–473.

[37] Brezinski, C. and Van Iseghem, J. (1994). Vector orthogonal polynomials of dimension $-d$, in *Approximation and Computation*, R.V.M. Zahar ed., Birkhäuser, Basel, 29–39.

[38] Brezinski, C. and Van Iseghem, J. (1994). Padé approximations, in *Handbook of Numerical Analysis*, Vol. III, P.G. Ciarlet and J.L. Lions eds., North–Holland, Amsterdam, 47–222.

[39] Brezinski, C. and Van Iseghem, J. (1995). A taste of Padé approximation, in *Acta Numerica 1995*, A. Iserles ed., Cambridge University Press, Cambridge, 53–103.

[40] Bultheel, A. (1987). *Laurent Series and their Padé Approximations*, Birkhäuser, Basel.

[41] Bultheel, A., González–Vera, P., Hendriksen, E. and Njåstad, O. (1999). *Orthogonal Rational Functions*, Cambridge University Press, Cambridge.

[42] Bultheel, A. and Van Barel, M. (1997). *Linear Algebra, Rational Approximation and Orthogonal Polynomials*, North–Holland, Amsterdam.

[43] Buslaev, V. I. (2001). Simple counterexample to the Baker–Gammel–Wills conjecture, *East J. Approx.*, **7**, 515–517.

[44] Chihara, T. S. (1978). *An Introduction to Orthogonal Polynomials*, Gordon and Breach, New York.

[45] Da Rocha, Z. (1992). Implementation of the recurrence relations of biorthogonality, *Numer. Algorithms,* **3**, 173–184.

[46] Da Rocha, Z. (1999). QD–algorithms and recurrence relations for biorthogonal polynomials, *J. Comput. Appl. Math.*, **107**, 53–72.

[47] Davis, P. J. (1963). *Interpolation and Approximation*, Blaisdell, Waltham, (reprint by Dover, New York, 1975).

[48] Draux, A. (1983). *Polynômes Orthogonaux Formels. Applications*, Lecture Notes in Mathematics, Vol. **974**, Springer-Verlag, Berlin.

[49] Dunkl, C. F. and Xu, Y. (2001). *Orthogonal Polynomials of Several Variables*, Cambridge University Press, Cambridge.

[50] Eğecioğlu, Ö. and . Koç, Ç (1989). A fast algorithm for rational interpolation via orthogonal polynomials, *Math. Comput.*, **53**, 249–264.

[51] Evans, W. D., Littlejohn, L. L., Marcellán, F., Markett, C. and Ronveaux, A. (1995). On recurrence relations for Sobolev orthogonal polynomials, *SIAM J. Math. Anal.*, **26**, 446–467.

[52] Feldmann, P. and Freund, R. W. (1995). Efficient linear circuit analysis by Padé approximation via the Lanczos process, *IEEE Trans. Comput.–Aided Design*, **14**, 639–649.

[53] Fischer, B. (1996). *Polynomial Based Iteration Methods for Symmetric Linear Systems*, John Wiley, Chichester and B.G. Teubner, Stuttgart.

[54] Fischer, B. and Prestin, J. (1997). Wavelets based on orthogonal polynomials, *Math. Comput.*, **66**, 1593–1618.

[55] Fletcher, R. (1976). Conjugate gradient methods for indefinite systems, in *Numerical Analysis, Dundee 1975*, G.A. Watson ed., Lecture Notes in Mathematics, Vol. **506**, Springer–Verlag, Berlin, 73–89.

[56] Freund, R. W., Golub, G. H.and Nachtigal, N. M. (1992). Iterative solution of linear systems, in *Acta Numerica 1992*, A. Iserles ed., Cambridge University Press, Cambridge, 57–100.

[57] Gautschi, W. (1996). Orthogonal polynomials: applications and computation, in *Acta Numerica 1996*, A. Iserles ed., Cambridge University Press, Cambridge, 45–119.

[58] Gemignani, L. (1993). Rational interpolation via orthogonal polynomials, *Comput. Math. Appl.*, **26**, 27–34.

[59] Graves–Morris, P. R. (1997). A "look–around Lanczos" algorithm for solving a system of linear equations, *Numer. Algorithms*, **15**, 247–274.

[60] Gutknecht, M- H. (1997). Lanczos–type solvers for nonsymmetric linear systems of equations, in *Acta Numerica 1997*, A. Iserles ed., Cambridge University Press, Cambridge, 271–397.

[61] Henrici, P. (1974). *Applied and Computational Complex Analysis*, Vol. 1, Wiley, New York.

[62] Hestenes, M. R. and Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems, *J. Res. Natl. Bur. Stand.*, **49**, 409–436.

[63] Jones, W. B. and Thron, W. J. (1980). *Continued Fractions. Analytic Theory and Applications*, Addison–Wesley, Reading.

[64] Kravanja, P. and Van Barel, M. (2000). *Computing the Zeros of Analytic Functions*, Lecture Notes in Mathematics, Vol. **1727**, Springer–Verlag, Berlin.

[65] Lanczos, C. (1950). An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, *J. Res. Natl. Bur. Stand.*, **45**, 255–282.

[66] Lanczos, C. (1952). Solution of systems of linear equations by minimized iterations, *J. Res. Natl. Bur. Stand.*, **49**, 33–53.

[67] Lorentzen, L. and Waadeland, H. (1992). *Continued Fractions with Applications*, North-Holland, Amsterdam.

[68] Lubinsky, D. S. Rogers–Ramanujan and the Baker–Gammel–Wills (Padé) conjecture, *Ann. Math.*, to appear.

[69] Maroni, P. (1985). Sur quelques espaces de distributions qui sont des formes linéaires sur l'espace vectoriel des polynômes, in *Polynômes Orthogonaux et Applications*, C. Brezinski et al. eds., Lecture Notes in Mathematics, Vol. **1171**, Springer–Verlag, Berlin, 184–194.

[70] Maroni, P. (1987). Prolégomènes à l'étude des polynômes orthogonaux, *Ann. Mat. Pura ed Appl.*, **149**, 165–184.

[71] Maroni, P. (1988). Le calcul des formes linéaires et les polynômes orthogonaux semi–classiques, in *Orthogonal Polynomials and their Applications*, M. Alfaro et al. eds., Lecture Notes in Mathematics, Vol. **1329**, Springer–Verlag, Berlin, 279–288.

[72] Maroni, P. (1991). Une théorie algébrique des polynômes orthogonaux. Application aux polynômes orthogonaux semi–classiques, in *Orthogonal Polynomials and their Applications*, C. Brezinski et al. eds., J.C. Baltzer, Basel, 95–130.

[73] Minesaki, Y. and Nakamura, Y. (2001). The discrete relativistic Toda molecule equation and a Padé approximation algorithm, *Numer. Algorithms*, **27**, 219–235.

[74] Monegato, G. (2001). An overview of the computational aspects of Kronrod quadrature rules, *Numer. Algorithms*, **26**, 173–196.

[75] Nagai, A., and Satsuma, J. (1995). Discrete soliton equations and convergence acceleration algorithms, *Phys. Letters*, **A209**, 305–312.

[76] Nikiforov, A. F., Suslov, S. K. and Uvarov, V. B. (1991). *Classical Orthogonal Polynomials of a Discrete Variable*, Springer–Verlag, Berlin.

[77] Nikishin, E. M. and Sorokin, V. N. (1991). *Rational Approximations and Orthogonality*, American Mathematical Society, Providence.

[78] O'Leary, D. P. (1980). The block conjugate gradient algorithm and related methods, *Linear Algebra Appl.*, **29**, 293–322.

[79] Padé, H. (1984). *Œuvres*, C. Brezinski éd., Librairie Scientifique et Technique A. Blanchard, Paris.

[80] Papageorgiou, V., Grammaticos, B. and Ramani, A. (1993). Integrable lattices and convergence acceleration algorithms, *Phys. Letters*, **A179**, 111–115.

[81] Prévost, M. (1988). Stieltjes– and Geronimus–type polynomials, *J. Comput. Appl. Math.*, **21**, 133–144.

[82] Rivlin, T. J. (1990). *Chebyshev Polynomials. From Approximation Theory to Algebra and Number Theory*, 2nd edition, John Wiley, New York.

[83] Roman, S. (1984). *The Umbral Calculus*, Academic Press, Orlando.

[84] Rutishauser, H. (1957). *Der Quotienten–Differenzen Algorithmus*, Birkhäuser–Verlag, Basel.

[85] Rutishauser, H. (1958). Solution of eigenvalue problems with the LR–transformation, *Natl. Bur. Stand. Appl. Math. Series*, **49**, 47–81.

[86] Saad, Y., Schultz, M. H., GMRES (1986). A generalized minimal residual algorithm for solving nonsymmetric linear systems, *SIAM J. Sci. Statist. Comput.*, **7**, 856–869.

[87] Shanks, D. (1955). Non linear transformations of divergent and slowly convergent sequences, *J. Math. Phys.*, **34**, 1–42.

[88] Sonneveld, P. (1989). A fast Lanczos–type solver for nonsymmetric linear systems, *SIAM J. Sci. Stat. Comput.*, **10**, 35–52.

[89] Sorokin, V. N. and Van Iseghem, J. (1999). Matrix continued fractions, *J. Approx. Theory*, **96**, 237–257.

[90] Stahl, H. and Totik, V. (1992). *General Orthogonal Polynomials*, Cambridge University Press, Cambridge.

[91] Suetin, P. K. (1999). *Orthogonal Polynomials in Two Variables*, Gordon and Breach, Amsterdam.

[92] Szegö, G. (1939). *Orthogonal Polynomials*, American Mathematical Society, Providence.

[93] Van Der Vorst, H. A. (1992). Bi–CGSTAB: a fast and smoothly converging variant of Bi–CG for the solution of nonsymmetric linear systems, *SIAM J. Sci. Stat. Comput.*, **13**, 631–644.

[94] Van Iseghem, J. (1985). Vector Padé approximants, in *Numerical Mathematics and Applications*, R. Vichnevetsky and J. Vignes eds., North–Holland, Amsterdam, 73–77.

[95] Van Iseghem, J. (1987). Vector orthogonal relations. Vector *qd*–algorithm, *J. Comput. Appl. Math.*, **19**, 141–150.

[96] Vorobyev, Yu, V. (1965). *Method of Moments in Applied Mathematics*, Gordon and Breach, New York.

[97] Weinert H. L. ed. (1982). *Reproducing Kernel Hilbert Spaces. Applications in Statistical Signal Processing*, Hutchinson Ross Publishing Company, Stroudsburg.

[98] Weniger, E. J. (1989). Nonlinear sequence transformations for the acceleration of convergence and the summation of divergent series, *Comput. Physics Reports*, **10**, 189–373.

[99] Wimp, J. (1981). *Sequence Transformations and their Applications*, Academic Press, New York.

[100] Wynn, P. (1956). On a device for computing the $e_m(S_n)$ transformation, *MTAC*, **10**, 91–96.

C. Brezinski
Laboratoire de Mathématiques Appliquées, FRE CNRS 2222,
UFR de Mathématiques Pures et Appliquées,
Université des Sciences et Technologies de Lille,
59655–Villeneuve d'Ascq cedex, France.
Claude.Brezinski@univ-lille1.fr