Fields Institute Communications
Volume **00**, 0000

# On Orthogonal Polynomials in Several Variables

**Yuan Xu**
Department of Mathematics
University of Oregon
Eugene, Oregon 97403-1222

**Abstract.** We report on the recent development on the general theory of orthogonal polynomials in several variables, in which results parallel to the theory of orthogonal polynomials in one variable are established using a vector-matrix notation.

## 1 Introduction

The main theme of this paper is the *general* theory of orthogonal polynomials in several variables, which deals with those properties shared by all systems of orthogonal polynomials. Thus, we shall not discuss the theory of particular systems and their applications.

Let $\Pi^d$ be the set of polynomials in $d$ variables on $\mathbb{R}^d$, and let $\Pi_n^d$ be the subset of polynomials of total degree at most $n$. A linear functional $\mathcal{L}$ defined on $\Pi^d$ is called square positive if $\mathcal{L}(p^2) > 0$ for all $p \in \Pi^d$, $p \neq 0$. Such a linear functional $\mathcal{L}$ induces an inner product on $\Pi^d$, defined by $\langle P, Q\rangle = \mathcal{L}(PQ)$. Two polynomials are said to be orthogonal with respect to $\mathcal{L}$ if their inner product is zero.

Given a square positive linear functional, we may apply the Gram-Schmidt orthogonalization process on the multiple sequence of monomials $x_1^{\alpha_1} \dots x_d^{\alpha_d}$, $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$ and $\alpha_i \in \mathbb{N}_0$, to generate a sequence of orthogonal polynomials. In order to do so, however, it is necessary to order the multiple sequence in a simple one. There are many ways to do so; in general, different orderings will lead to different orthogonal systems. Thus, there is no unique system of orthogonal polynomials in several variables. Moreover, any system of orthogonal polynomials obtained by an ordering of the monomials is necessarily unsymmetric in the variables $x_1, \dots, x_d$. These are essential difficulties in the study of orthogonal polynomials in several variables; they are mentioned in the introductory section of Chapter XII in the *Higher Transcendental Functions* (Erdélyi, et al. [1953, p. 265]). On the same page of the book it is also stated that "there does not seem to be an extensive general theory of

---

1991 *Mathematics Subject Classification*. Primary 33C50, 42C05, 65D32; Secondary 33D45, 41A10, 41A63.

Supported in part by NSF Grant #9302721 and #9500532.

orthogonal polynomials in several variables". A brief account of the general properties of orthogonal polynomials known up to then is summarized in Chapt. XII of Erdélyi, et al. [1953].

More recent accounts of general theory of orthogonal polynomials in several variables can be found in Krall and Sheffer [1967], Mysovskikh [1976] and Suetin [1988]. We will report on progress made in the past few years, from which a possible extensive general theory may emerge. In the center of this new development is a vector-matrix notation which helps us to overcome the difficulty caused by the non-uniqueness of orthogonal polynomials. For each $n > 0$ the set of polynomials of degree $n$ that are orthogonal to all polynomials of lower degree froms a vector space $V_n$ of dimension greater than one. The non-uniqueness of orthogonal polynomials means that there is no unique way of choosing a basis for $V_n^d$. In their work of extending the characterization of the classical orthogonal polynomials as eigenfunctions of second order differential operators from one to two variables, Krall and Sheffer [1967] suggested one way to overcome the difficulty; they remarked that if the results can be stated "in terms of $V_0^d, V_1^d, \ldots, V_n^d, \ldots$ rather than in terms of a particular basis in each $V_n^d$, a degree of uniqueness is restored". Kowalski [1982a] and [1982b] used a vector-matrix notation to prove an analog of Favard's theorem, but he seemed unaware of the paper of Krall and Sheffer. His theorem is simplified in Xu [1993a] using a modified vector-matrix notation. The latter notation is inspired by and very much reflects the principle of Krall and Sheffer; it has been adopted in further studies of the author Xu [1993b], [1994a-e] and [1994h], upon which the present report is based.

The paper is organized as follows. The next section is devoted to the notations and preliminaries. In Section 3, we discuss the three-term relation satisfied by vectors of orthogonal polynomials and the results around it. The three-term relation leads us naturally to define a family of block Jacobi matrices, which can be studied as a commuting family of self-adjoint operators; this connection is covered in Section 4. The zeros of orthogonal polynomials are defined as common zeros of a family of polynomials, they are related to numerical cubature formulae. In Section 5 we discuss the basic properties of common zeros and cubature formulae; deeper results are covered in Section 6. Finally, in Section 7 we discuss the preliminary results dealing with asymptotics and Fourier orthogonal expansion involving orthogonal polynomials in several variables.

## 2 Notations and preliminaries

We use the standard multiindex notation. Let $\mathbb{N}_0$ be the set of nonnegative integers. For $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$ and $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$ we write $\mathbf{x}^\alpha = x_1^{\alpha_1} \cdots x_d^{\alpha_d}$. The number $|\alpha| = \alpha_1 + \cdots + \alpha_d$ is called the total degree of $\mathbf{x}^\alpha$. For $n \in \mathbb{N}_0$, we denote by $\Pi^d$ the set of all polynomials in $d$ variables and by $\Pi_n^d$ the subset of polynomials of total degree at most $n$; *i.e.*,

$$\Pi_n^d = \Big\{ \sum_{|\alpha| \le n} a_\alpha \mathbf{x}^\alpha : \quad a_\alpha \in \mathbb{R}, \quad \mathbf{x} \in \mathbb{R}^d \Big\}.$$

We denote by $r_n^d$ the number of monomials of degree exactly $n$ which is equal to the cardinality of the set $\{\alpha \in \mathbb{N}_0^d : |\alpha| = n\}$; it follows that

$$\dim \Pi_n^d = \binom{n+d}{d} \quad \text{and} \quad r_n^d = \binom{n+d-1}{n}.$$

For $n \in \mathbb{N}_0$ and $\mathbf{x} \in \mathbb{R}^d$, we denote by $\mathbf{x}^n = (\mathbf{x}^\alpha)_{|\alpha|=n}$ a vector of size $r_n^d$, where the monomials are arranged according to the lexicographical order of $\{\alpha \in \mathbb{N}_0^d : |\alpha| = n\}$.

A multisequence $s : \mathbb{N}_0^d \mapsto \mathbb{R}$ will be written in the form $s = \{s_\alpha\}_{\alpha \in \mathbb{N}_0^d}$. Such a sequence is called *positive definite* if for every tuple $(\beta^{(1)}, \ldots, \beta^{(r)})$ of distinct multiindices $\beta^{(j)} \in \mathbb{N}_0^d$, $1 \le j \le r$, the matrix $(s_{\beta^{(i)}+\beta^{(j)}})_{i,j=1,\ldots,r}$ has positive determinant. With each real $d$-sequence $s$ one can associate a linear functional on $\Pi^d$ defined by $\mathcal{L}(\mathbf{x}^\alpha) = s_\alpha$. If $s$ is positive definite, we call the associated linear functional *square positive*. Since

$$\mathcal{L}(P^2) = \sum_{\alpha,\beta} s_{\alpha+\beta} a_\alpha a_\beta \qquad \text{where} \qquad P(\mathbf{x}) = \sum_\alpha a_\alpha \mathbf{x}^\alpha,$$

it follows that $\mathcal{L}$ is square positive if, and only if, $\mathcal{L}(P^2) > 0$ for all $P \in \Pi^d$ not identically zero. We remark that we do not deal with *semi square positive* which requires only $\mathcal{L}(P^2) \ge 0$; see Fuglede [1983] and Berg [1987] for discussion. Let $\mathcal{M} = \mathcal{M}(\mathbb{R}^d)$ denote the set of nonnegative Borel measures $\mu$, defined on the $\sigma$-algebra of Borel sets, with an infinite support on $\mathbb{R}^d$ such that

$$\int_{\mathbb{R}^d} |\mathbf{x}^\alpha| \, d\mu(\mathbf{x}) < +\infty \ , \qquad \forall \ \alpha \in \mathbb{N}_0^d \ .$$

For any $\mu \in \mathcal{M}$ the moments of $\mu$ are defined as the real numbers $s_\alpha = \int_{\mathbb{R}^d} \mathbf{x}^\alpha d\mu$. A $d$-sequence $s = (s_\alpha)$ is called a moment sequence if there exists at least one measure $\mu$ whose moments are equal to $s_\alpha$ for all $\alpha \in \mathbb{N}_0^d$. If $s$ is a moment sequence, the corresponding linear functional $\mathcal{L}$ is called a moment functional, which has an integral representation

$$\mathcal{L}(f) = \int_{\mathbb{R}^d} f(\mathbf{x}) \, d\mu(\mathbf{x}) \ , \qquad \mu \in \mathcal{M} \ . \tag{2.1}$$

Two measures in $\mathcal{M}$ are called equivalent if they have the same moments. If the equivalent class of measures having the same moments as $\mu$ consists of $\mu$ only, the measure $\mu$ is called determinate. If $\mathcal{L}$ is a moment functional of a determined measure, then the integral representation is unique. It is known that $\mathcal{L}$ is a moment functional if it is positive, which means that $\mathcal{L}(P) \ge 0$ whenever $P \ge 0$. For $d = 1$, $\mathcal{L}$ being positive is equivalent to $\mathcal{L}(P^2) \ge 0$. For $d > 1$, however, they are no longer equivalent, which is, in fact, the cause of many problems in the multidimensional moment problem (cf. Berg [1983], Berg and Thill [1991] and Fuglede [1983]).

A square positive linear functional $\mathcal{L}$ induces an inner product $\langle \cdot, \cdot \rangle$ on $\Pi^d$ defined by

$$\langle P, Q \rangle = \mathcal{L}(PQ), \qquad P, Q \in \Pi^d.$$

For convenience, we shall always assume that $\mathcal{L}(1) = 1$. Two polynomials $P$ and $Q$ are said to be orthogonal with respect to $\mathcal{L}$, if $\langle P, Q \rangle = 0$. With respect to such an $\mathcal{L}$ we apply the Gram-Schmidt orthogonalization process on the monomials $\{\mathbf{x}^\alpha\}$ arranged as $\{\mathbf{x}^n\}_{n=0}^\infty$ to derive a sequence of orthonormal polynomials, denoted by $\{P_k^n\}_{k=1}^{r_n^d}{}_{n=0}^{\infty}$, where the superscript $n$ means that $P_k^n \in \Pi_n^d$. We now introduce the vector notation that is fundamental in the development below:

$$\mathbb{P}_n(\mathbf{x}) = \left[P_1^n(\mathbf{x}), P_2^n(\mathbf{x}), \ldots, P_{r_n^d}^n(\mathbf{x})\right]^T. \tag{2.2}$$

Using this notation, the orthonormality property of $\{P_k^n\}$ can be described as

$$\mathcal{L}(\mathbb{P}_n\mathbb{P}_m^T) = \delta_{m,n} I_{r_n^d}, \tag{2.3}$$

where $I_n$ denotes the $n \times n$ identity matrix. For convenience, we sometimes call $\mathbb{P}_n$ orthonormal polynomials with respect to $\mathcal{L}$. When $\mathcal{L}$ has an integral representation (2.1) with $d\mu = W(\mathbf{x})d\mathbf{x}$, we call $\mathbb{P}_n$ orthonormal polynomials with respect to $W$ instead of to $\mathcal{L}$. In terms of monomial vectors $\mathbf{x}^n$, we can write $\mathbb{P}_n$ as

$$\mathbb{P}_n = G_n\mathbf{x}^n + G_{n,n-1}\mathbf{x}^{n-1} + G_{n,n-2}\mathbf{x}^{n-2} + \dots \tag{2.4}$$

where $G_{n,i} : r_n^d \times r_{n-i}^d$ are matrices. We call $G_n = G_{n,n}$ the leading coefficient of $\mathbb{P}_n$, which can be seen to be invertible since $\mathcal{L}$ is square positive.

For each $k \ge 0$, let $V_k^d \subset \Pi_k^d$ be the set of polynomials spanned by $\mathbb{P}_k$, that is, spanned by the components of $\mathbb{P}_k$, together with zero. Then $V_k^d$ is a vector space of dimension $r_k^d$ which is orthogonal to all polynomials in $\Pi_{k-1}^d$. Clearly

$$\Pi_n^d = \bigoplus_{k=0}^{n} V_k^d \quad \text{and} \quad \Pi^d = \bigoplus_{k=0}^{\infty} V_k^d \ ,$$

and $V_k^d$ are mutually orthogonal. As we mentioned before, the sequence of orthonormal polynomial is not unique. Actually, it is easy to see that each orthogonal matrix $Q$ of order $r_k^d$ gives rise to an orthonormal basis $Q\mathbb{P}_k$ of $V_k^d$ and every orthonormal basis of $V_k^d$ is of the form $Q\mathbb{P}_n$. One can also work with other bases of $V_k$ that are not necessarily orthonormal. In particular, one basis consists of polynomials $\tilde{P}_\alpha^k$ of the form

$$\tilde{P}_\alpha^k = \mathbf{x}^\alpha + R_\alpha^{k-1}, \qquad |\alpha| = k, \qquad R_\alpha^{k-1} \in \Pi_{k-1}^d.$$

This basis is sometimes called monomial basis; in general, $\mathcal{L}(\tilde{P}_\alpha^k \tilde{P}_\beta^k) \ne 0$, although $\tilde{P}_\alpha^k$ are orthogonal to all polynomials of lower degrees. It is easy to see that the matrix $H_n = \mathcal{L}(\tilde{\mathbb{P}}_n\tilde{\mathbb{P}}_n^T)$ is positive definite and $\tilde{\mathbb{P}}_n = H_n^{-1/2}\mathbb{P}_n$. Because of the relation, most of the results below can be stated in terms of the monomial basis. We should mention that one can define orthogonal polynomials with respect to linear functionals that are not necessarily square positive. However, the square positiveness is necessary for obtaining orthonormal polynomials.

For a square positive linear functional expressible by (2.1), we can consider the orthogonal expansion of a function $f \in L_{d\mu}$. Using the vector notation $\mathbb{P}_n$, the Fourier orthogonal expansion of $f$ with respect to a sequence of orthonormal polynomials $\{P_k^n\}$ is given by

$$f \sim \sum_{n=0}^{\infty} \mathbf{a}_n^T(f)\mathbb{P}_n, \qquad \mathbf{a}_n(f) = \int f(\mathbf{x})\mathbb{P}_n(\mathbf{x})d\mu \tag{2.5}$$

where $\mathbf{a}_n$ is a vector of $\mathbb{R}^{r_n^d}$. If we replace $\mathbf{a}_n^T\mathbb{P}_n$ by $\mathrm{proj}_{V_n} f$, then the expansion can be viewed as in terms of the $V_n^d$ and is independent of the particular basis of $V_n^d$. In fact, the $n$-th reproducing kernel of the orthonormal polynomials, defined by

$$\mathbf{K}_n(\mathbf{x},\mathbf{y}) = \sum_{k=0}^{n-1}\sum_{j=0}^{r_k^d} P_j^k(\mathbf{x})P_j^k(\mathbf{y}) = \sum_{k=0}^{n} \mathbb{P}_n^T(\mathbf{x})\mathbb{P}_n(\mathbf{y}), \tag{2.6}$$

is easily seen to depend on $V_k^d$ rather than a particular basis of $V_k^d$. The $n$-th partial sum $S_n f$ of the expansion can be written in terms of $\mathbf{K}_n(\cdot,\cdot)$ as

$$S_n f = \sum_{k=0}^{n-1} \mathbf{a}_k^T(f)\mathbb{P}_k = \int \mathbf{K}_n(\cdot,\mathbf{y}) f(\mathbf{y}) d\mu. \tag{2.7}$$

From time to time, we may use results for orthogonal polynomials in one variable for motivation or comparison. We follow the standard notation in one variable (cf. Szegő [1975] and Chihara [1978]). The orthonormal polynomial of degree $n$ on $\mathbb{R}$ is denoted by $p_n$. The three-term relation satisfied by $p_n$ is denoted by

$$x p_n = a_n p_{n+1} + b_n p_n + a_{n-1} p_{n-1}, \qquad x \in \mathbb{R}, \tag{2.8}$$

where $a_n$ and $b_n$ are called the coefficients of the three-term relation.

## 3 Three-term relation

Our development of a general theory of orthogonal polynomials in several variables starts from a three-term relation in a vector-matrix notation very much like in the one variable theory.

*Three-term relation. For $n \ge 0$, there exist matrices $A_{n,i} : r_n^d \times r_{n+1}^d$ and $B_{n,i} : r_n^d \times r_n^d$, such that*

$$x_i \mathbb{P}_n = A_{n,i}\mathbb{P}_{n+1} + B_{n,i}\mathbb{P}_n + A_{n-1,i}^T \mathbb{P}_{n-1}, \qquad 1 \le i \le d, \tag{3.1}$$

*where we define $\mathbb{P}_{-1} = 0$ and $A_{-1,i} = 0$.*

In fact, since the components of $x_i\mathbb{P}_n$ are polynomials of degree $n+1$, they can be written in terms of linear combinations of $\mathbb{P}_{n+1}, \ldots, \mathbb{P}_0$; the orthonormal property of $\mathbb{P}_n$ implies that only the coefficients of $\mathbb{P}_{n+1}$, $\mathbb{P}_n$ and $\mathbb{P}_{n-1}$ are nonzero. The result is the relation (3.1). Moreover, the matrices in the three-term relation are expressible as

$$A_{n,i} = \mathcal{L}(x_i \mathbb{P}_n \mathbb{P}_{n+1}^T) \quad \text{and} \quad B_{n,i} = \mathcal{L}(x_i \mathbb{P}_n \mathbb{P}_n^T). \tag{3.2}$$

As a consequence, the matrices $B_{n,i}$ are symmetric. If we are dealing with orthogonal polynomials, $\tilde{\mathbb{P}}_n$, which are not necessarily orthonormal, then the three-term relation takes the form

$$x_i \tilde{\mathbb{P}}_n = A_{n,i}\tilde{\mathbb{P}}_{n+1} + B_{n,i}\tilde{\mathbb{P}}_n + C_{n,i}^T \tilde{\mathbb{P}}_{n-1}, \qquad 1 \le i \le d, \tag{3.3}$$

where $C_{n,i} : r_n^d \times r_{n-1}^d$ is related to $A_{n,i}$ by

$$A_{n,i} H_{n+1} = H_n C_{n+1,i}, \qquad \text{where} \qquad H_n = \mathcal{L}(\tilde{\mathbb{P}}_n \tilde{\mathbb{P}}_n^T).$$

Moreover, comparing the highest coefficient matrices at both sides of (3.1), it follows that

$$A_{n,i} G_{n+1} = G_n L_{n,i}, \qquad 1 \le i \le d, \tag{3.4}$$

where $L_{n,i}$ are matrices of size $r_n^d \times r_{n+1}^d$ which are defined by

$$L_{n,i}\mathbf{x}^{n+1} = x_i \mathbf{x}^n, \qquad 1 \le i \le d.$$

Clearly, $\operatorname{rank} L_{n,i} = r_n^d$, and $\operatorname{rank} L_n = r_{n+1}^d$, where $L_n = (L_{n,1}^T | \dots | L_{n,d}^T)^T$. For example, for $d = 2$ we have

$$L_{n,1} = \begin{bmatrix} 1 & & \bigcirc & 0 \\ & \ddots & & \vdots \\ \bigcirc & & 1 & 0 \end{bmatrix} \quad \text{and} \quad L_{n,2} = \begin{bmatrix} 0 & 1 & & \bigcirc \\ \vdots & & \ddots & \\ 0 & \bigcirc & & 1 \end{bmatrix}.$$

From the relation (3.4) and the fact that $G_n$ is invertible, it follows readily that the matrices $A_{n,i}$ satisfy

*Rank conditions.* For $n \ge 0$, $\operatorname{rank} A_{n,i} = r_n^d$ for $1 \le i \le d$, and

$$\operatorname{rank} A_n = r_{n+1}^d, \qquad A_n = (A_{n,1}^T, \dots, A_{n,d}^T)^T. \tag{3.5}$$

The importance of the three-term relation is readily seen in the following analog of Favard's theorem of one variable. We extend the notation (2.2) to an arbitrary sequence of polynomials $\{P_k^n\}_{k=1}^{r_n^d}$.

**Theorem 3.1** (Xu [1993a]). *Let $\{\mathbb{P}_n\}_{n=0}^\infty$, $\mathbb{P}_0 = 1$, be a sequence in $\Pi^d$. Then the following statements are equivalent:*

1. *There exists a linear functional which is square positive and which makes $\{\mathbb{P}_n\}_{n=0}^\infty$ an orthonormal basis in $\Pi^d$;*
2. *there exist matrices $A_{n,i} : r_n^d \times r_{n+1}^d$ and $B_{n,i} : r_n^d \times r_n^d$ such that*
   (1) *the polynomial vectors $\mathbb{P}_n$ satisfy the three-term relation* (3.1),
   (2) *the matrices in the relation satisfies the rank condition.*

An earlier version of this theorem is stated in Kowalski [1982b] with respect to the three-term relation (3.3) using the notation $\mathbf{x}\mathbb{P}_n = [x_1\mathbb{P}_n^T | \dots | x_d\mathbb{P}_n^T]^T$, where it is stated under an additional condition in part 2:

(3) *for an arbitrary sequence $D_0, D_1, \dots, D_n$ of matrices satisfying $D_k A_k = I$, the recursion*

$$J_0 = [1], J_{k+1} = D_k(C_{k+1,1} J_k^T | \dots | C_{k+1,1} J_k^T)^T, \qquad k = 0, 1, \dots,$$

*produces nonsingular symmetric matrices $J_k$.*

In Xu [1993a], the theorem is stated in the present form with respect to the three-term relation (3.3), where it is proved that the condition (3) is equivalent to that $\operatorname{rank} C_n = r_n^d$, here $C_n = (C_{k,1} | \dots | C_{k,d})$. The theorem in its present form for orthonormal polynomials is stated in Xu [1994a] for the first time.

The theorem is an analog of Favard's theorem in one variable, but it is not as strong as the classical Favard's theorem. It does not state, for example, when the linear functional $\mathcal{L}$ in the theorem will have an integral representation. We will address this question in the following section. For now, we concentrate on the three-term relation (3.1). It is an analog of the three-term relation in one variable; the fact that its coefficients are matrices reflect the complexity of the structure for $d \ge 2$.

In one variable, the sequence of orthogonal polynomials $\{p_n\}$ can also be viewed as a solution of the difference equation

$$y_{n+1} = a_n^{-1}(x y_n - b_n y_n - a_{n-1} y_{n-1}), \tag{3.6}$$

with initial values $p_0 = 1$ and $p_1 = a_1^{-1}(x - b_0)$. In particular, the three-term relation is also called *recurrence* relation, since $p_n$ can be computed recursively through such an relation. It follows trivially that all solutions of (3.6) satisfy the three-term relation (2.8).

To derive an analog of (3.6) in several variables, we first note that the rank condition (3.5) implies that there exists a matrix $D_n^T : r_{n+1}^d \times d\, r_n^d$, which we write as $D_n^T = (D_{n,1}^T | \dots | D_{n,d}^T)$ where $D_{n,i}$ are of size $r_n^d \times r_{n+1}^d$, such that

$$D_n^T A_n = \sum_{i=1}^d D_{n,i}^T A_{n,i} = I_{r_{n+1}^d}. \tag{3.7}$$

We note that such a $D_n$ is not unique, as can be seen from the singular value decomposition of $A_n$. Multiplying the three-term relation (3.1) from the left by $D_{n,i}^T$ and summing up the $d$ relations, we derived that $\mathbb{P}_n$ satisfies a recurrence relation

$$\begin{aligned}\mathbb{P}_{n+1}(\mathbf{x}) = \sum_{i=1}^d D_{n,i}^T x_i \mathbb{P}_n(\mathbf{x}) - &\left(\sum_{i=1}^d D_{n,i}^T B_{n,i}\right) \mathbb{P}_n(\mathbf{x}) \\ - &\left(\sum_{i=1}^d D_{n,i}^T A_{n-1,i}^T\right) \mathbb{P}_{n-1}(\mathbf{x})\end{aligned} \tag{3.8}$$

where $\mathbb{P}_{-1} = 0$, $\mathbb{P}_0 = 1$. For any given sequences $A_{n,i}$ and $B_{n,i}$, the relation (3.8) can be used to define a sequence of polynomials. However, in contrast to one variable, polynomials so defined may not satisfy the three-term relation (3.1), since (3.8) cannot be split into $d$ relations (3.1) in general. The question is answered by the following theorem.

**Theorem 3.2** (Xu [1994c]) *Let $\{\mathbb{P}_k\}_{k=0}^\infty$ be defined by* (3.8). *Then $\{\mathbb{P}_k\}_{k=0}^\infty$ satisfies the three-term relation* (3.1) *if, and only if, $B_{k,i}$ are symmetric, $A_{k,i}$ satisfy the rank condition* (3.5), *and together they satisfy the commuting conditions*

$$A_{k,i} A_{k+1,j} = A_{k,j} A_{k+1,i}\ , \tag{3.9a}$$

$$A_{k,i} B_{k+1,j} + B_{k,i} A_{k,j} = B_{k,j} A_{k,i} + A_{k,j} B_{k+1,i}\ , \tag{3.9b}$$

$$\begin{aligned}A_{k-1,i}^T A_{k-1,j} + B_{k,i} B_{k,j} &+ A_{k,i} A_{k,j}^T \\ &= A_{k-1,j}^T A_{k-1,i} + B_{k,j} B_{k,i} + A_{k,j} A_{k,i}^T\end{aligned} \tag{3.9c}$$

*for $i \neq j$, $1 \le i, j \le d$, and $k \ge 0$, where $A_{-1,i} = 0$.*

In one variable, one can start with a sequence of positive numbers $a_n$ and an arbitrary sequence of real numbers $b_n$, use (3.6) to generate a sequence of polynomials, which is then automatically a sequence of orthogonal polynomials by Favard's theorem. For $d \ge 2$, such an approach is no longer feasible according to Theorem 3.2, since it is very difficult to tell when matrices will satisfy the commuting conditions (3.9). The reason why we call the equations in (3.9) commuting conditions will become clear in the next section. The necessity of these conditions follows easily from the fact that there are two ways to compute $\mathcal{L}(x_i x_j \mathbb{P}_n \mathbb{P}_n^T)$, $\mathcal{L}(x_i x_j \mathbb{P}_n \mathbb{P}_{n+1}^T)$ and $\mathcal{L}(x_i x_j \mathbb{P}_n \mathbb{P}_{n+2}^T)$ using the three-term relation. The sufficiency part is long and quite complicated, we refer to Xu [1994c].

In view of (3.6), we can also look at $\mathbb{P}_n$ as the solution of the multiparameter difference equations

$$x_i Y_k = A_{k,i} Y_{k+1} + B_{k,i} Y_k + A_{k-1,i}^T Y_{k-1}\ , \qquad 1 \le i \le d, \quad k \ge 1, \tag{3.10}$$

with the initial values

$$Y_0 = 1, \qquad Y_1 = A_0^{-1}(\mathbf{x} - B_0). \tag{3.11}$$

For $d = 1$, the difference equation (3.6) has two linearly independent solutions. One is $\{p_n\}$, which corresponds to the initial values $y_0 = 1$ and $y_{-1} = 0$, another is usually denoted by $\{q_n\}$, which corresponds to the initial values $y_0 = 0$ and $y_{-1} = 1$. The polynomials $q_n$ are usually called the *associated polynomials*, or polynomials of the second kind. They play an important role in areas such as problem of moments, spectral theory of the Jacobi matrices, and continuous fractions (cf. Akheizer [1965]). For the multiparameter difference equations, one may easily guess that there will be $d+1$ linearly independent solutions. However, the following theorem says that $\mathbb{P}_n$ is essentially the only possible solution.

**Theorem 3.3** (Xu [1994e]) *If the multiparameter difference equation* (3.10) *has a solution* $\mathbb{P} = \{\mathbb{P}_k\}_{k=0}^{\infty}$ *for the particular initial value* (3.11), *then all other solutions of* (3.10) *are multiples of* $\mathbb{P}$ *with the possible exception of the first component. More precisely, if* $Y = \{Y_n\}_{n=0}^{\infty}$ *is a solution of* (3.11), *then* $Y_n = h\mathbb{P}_n$ *for all* $n \ge 1$, *where* $h$ *is a function independent of* $n$.

The theorem comes out somewhat surprising, especially in view of the important role played by the associated polynomials in one variable. It indicates that it is much harder to extract information from the three-term relation (3.1) for $d \ge 2$.

Among other consequences of the three-term relation, we mention an extension of the Christoffel-Darboux formula of one variable. Let $\mathbf{K}_n(\cdot,\cdot)$ be the reproducing kernel defined in (2.6). Then we have (Xu [1993a])

*Christoffel-Darboux formula: For* $n \ge 1$, $1 \le i \le d$,

$$\mathbf{K}_n(\mathbf{x},\mathbf{y}) = \frac{[A_{n-1,i}\mathbb{P}_n(\mathbf{x})]^T \mathbb{P}_{n-1}(\mathbf{y}) - \mathbb{P}_{n-1}^T(\mathbf{x})[A_{n-1,i}\mathbb{P}_n(\mathbf{y})]}{x_i - y_i}. \tag{3.12}$$

It is interesting to note that although the right hand side of the formula seems to depend on $i$, the left hand side shows that it does not. Because of (2.6), the Christoffel-Darboux formula is important in the study of Fourier orthogonal series. It also plays an important role in the study of the common zeros of orthogonal polynomials.

## 4 Block Jacobi Matrices

For $d = 1$, the linear functional in Favard's theorem is known to be given by an integral with respect to $d\phi$, where $\phi$ is a non-decreasing function with infinite support. There are several ways to establish this result, one of them uses the spectral theorem of self-adjoint operators. With the three-term relation (2.8) of one variable we associate a tridiagonal matrix $J$, customarily called the *Jacobi matrix*, with $a_n$ on the main diagonal and $b_n$ on the two subdiagonals, that acts as an operator on $\ell^2$ (cf. Stone [1932]). This matrix has been studied extensively in operator theory and it plays a very important role in the study of orthogonal polynomials in one variable.

For $d \geq 1$, the coefficients of the three-term relation (3.1) can be used to define a family of block tridiagonal matrices (Xu [1994a]),

$$J_i = \begin{bmatrix} B_{0,i} & A_{0,i} & & \bigcirc \\ A_{0,i}^T & B_{1,i} & A_{1,i} & \\ & A_{1,i}^T & B_{2,i} & \ddots \\ \bigcirc & & \ddots & \ddots \end{bmatrix}, \qquad 1 \leq i \leq d. \tag{4.1}$$

We call them *block Jacobi matrices*. These matrices can be viewed as a family of linear operators which act via matrix multiplication on $\ell^2$. The domain of $J_i$ consists of all sequences in $\ell^2$ for which matrix multiplication yields sequences in $\ell^2$. Under proper conditions, the matrices $J_i$ form a family of commuting self-adjoint operators, which will yield via spectral theorem that $\mathcal{L}$ in Theorem 3.1 has an integral representation. In order to proceed, we need some notion from the spectral theory of self-adjoint operators in a Hilbert space (cf. Riesz and Nagy [1955]).

Let $\mathcal{H}$ be a separable Hilbert space. Each self-adjoint operator $T$ in $\mathcal{H}$ associates with a spectral measure $E$ on $\mathbb{R}$ such that $T = \int x dE(x)$. $E$ is a projection valued measure defined for the Borel sets of $\mathbb{R}$ such that $E(\mathbb{R})$ is the identity operator in $\mathcal{H}$ and $E(B \cap C) = E(B) \cap E(C)$ for Borel sets $B, C \subseteq \mathbb{R}$. For any $f \in \mathcal{H}$ the mapping $B \to \langle E(B)f, f\rangle$ is an ordinary measure defined for the Borel sets $B \subseteq \mathbb{R}$ and denoted $\langle Ef, f\rangle$. A family of operators $\{T_1, \ldots, T_d\}$ in $\mathcal{H}$ *commutes*, by definition, if their spectral measures commute, i.e. $E_i(B)E_j(C) = E_i(C)E_j(B)$ for any $i, j = 1, \ldots, d$ and any two Borel sets $B, C \subseteq \mathbb{R}$. If $T_1, \ldots, T_d$ commute, then $E = E_1 \otimes \cdots \otimes E_d$ is a spectral measure on $\mathbb{R}^d$ with values that are self-adjoint projections in $\mathcal{H}$. In particular, $E$ is the unique measure such that

$$E(B_1 \times \cdots \times B_d) = E_1(B_1) \cdots E_d(B_d)$$

for any Borel sets $B_1, \ldots, B_d \subseteq \mathbb{R}$. The measure $E$ is called the spectral measure of the commuting family $T_1, \ldots, T_d$. A vector $\Phi_0 \in \mathcal{H}$ is a *cyclic vector* in $\mathcal{H}$ with respect to the commuting family of self-adjoint operators $T_1, \ldots, T_d$ in $\mathcal{H}$ if the linear manifold $\{P(T_1, \ldots, T_d)\Phi_0,\ P \in \Pi^d\}$ is dense in $\mathcal{H}$. The spectral theorem for $T_1, \ldots, T_d$ states

*If $T_1, \ldots, T_d$ are a commuting family of self-adjoint operators with a cyclic vector $\Phi_0$, then $T_1, \ldots, T_d$ are unitarily equivalent to the multiplication operators $X_1, \ldots, X_d$ ,*

$$(X_i f)(\mathbf{x}) = x_i f(\mathbf{x}) \ , \qquad 1 \leq i \leq d \ ,$$

*defined on $L^2(\mathbb{R}^d, \mu)$, where the measure $\mu$ is defined by $\mu(B) = \langle E(B)\Phi_0, \Phi_0\rangle$ for the Borel set $B \subset \mathbb{R}^d$.*

The unitary equivalence means that there exists a unitary mapping $U : \mathcal{H} \to L^2(\mathbb{R}^d, \mu)$ such that $UT_iU^{-1} = X_i$, $1 \leq i \leq d$.

We apply the spectral theorem on the operators $J_1, \ldots, J_d$ on $\ell^2$ defined by the block Jacobi matrices. The connection to orthogonal polynomials in several variables is as follows. Let $\{\psi_n\}_{n=0}^\infty$ be the canonical orthonormal basis for $\ell^2$. We rewrite this basis as $\{\psi_n\}_{n=0}^\infty = \{\phi_j^k\}_{j=1}^{r_k^d}{}_{k=0}^{\infty}$ according to the lexicographical order, and introduce the formal vector notation $\Phi_k = [\phi_1^k, \ldots, \phi_{r_k^d}^k]^T$, $k \in \mathbb{N}_0$. Then, if the spectral theorem can be applied, the unitary mapping $U$ maps $\Phi_n$ to $\mathbb{P}_n$ and the relation

$$\int \mathbb{P}_n(\mathbf{x})\mathbb{P}_m^T(\mathbf{x})\, d\mu(\mathbf{x}) = \langle \Phi_n, \Phi_m^T \rangle \ , \quad \text{where} \quad \mu(B) = \langle E(B)\Phi_0, \Phi_0 \rangle$$

establishes the integral representation of $\mathcal{L}$. The main task is therefore to show that $J_1, \ldots, J_d$ are commuting and self-adjoint. In Xu [1994a] this is done under the assumption that $J_1, \ldots, J_d$ are bounded operators, which also implies that the support set of the spectral measure is bounded. Moreover, it is shown that $J_1, \ldots, J_d$ are bounded if, and only if,

$$\sup_{k \ge 0} \|A_{k,i}\|_2 < +\infty \quad , \quad \sup_{k \ge 0} \|B_{k,i}\|_2 < +\infty \ , \qquad 1 \le i \le d \ , \tag{4.2}$$

where $\|\cdot\|_2$ is the matrix norm induced by the Euclidean norm for vectors. Thus, one can strengthen Favard's theorem as follows.

**Theorem 4.1** (Xu [1994a]) *Let $\{\mathbb{P}_n\}_{n=0}^\infty$, $\mathbb{P}_0 = 1$, be a sequence in $\Pi^d$. Then the following statements are equivalent:*

1. *There exists a determinate measure $\mu \in \mathcal{M}$ with compact support in $\mathbb{R}^d$ such that $\{\mathbb{P}_n\}_{n=0}^\infty$ is orthonormal with respect to $\mu$.*
2. *The statement 2) in Theorem* 3.2 *holds together with* (4.2).

For bounded self-adjoint operators $T_1, \ldots, T_d$, the commuting of the spectral measures is equivalent to the formal commuting $T_iT_j = T_jT_i$. For block Jacobi matrices, the formal commuting can be easily verified as equivalent to the conditions in (3.9), which is why we call (3.9) commuting conditions. However, there are examples (cf. Nelson [1959]) of unbounded self-adjoint operators with a common dense domain such that they formally commute but their spectral measures do not commute. Fortunately, a sufficient condition for the commuting of operators in Nelson [1959] can be applied to $J_1, \ldots, J_d$, which leads to

**Theorem 4.2** (Xu [1993b]) *Let $\{\mathbb{P}_n\}_{n=0}^\infty$, $\mathbb{P}_0 = 1$, be a sequence in $\Pi^d$ that satisfies the three-term relation* (3.1) *and the rank condition. If*

$$\sum_{k=0}^\infty \frac{1}{\|A_{n,i}\|_2} = \infty, \qquad 1 \le i \le d, \tag{4.3}$$

*then there exists a determinate measure $\mu \in \mathcal{M}$ such that $\{\mathbb{P}_n\}$ is orthonormal with respect to $\mu$.*

It is worthwhile to point out that there is a close resemblance between the above set-up and the operator theoretic approach for moment problems, see Berg [1987] and Fuglede [1983]. For $d = 1$, the condition (4.3) is well-known; it implies the classical result of Carleman on the determinate of moment problem Akheizer [1965, p. 86]. For discussion of (4.3) in several variables we refer to Xu [1993b]. For further results concerning the block Jacobi matrices see Xu [1994b]; in particular, Dombrowski's formula for orthogonal polynomials of one variable (cf. Dombrowski [1990]) is extended to several variables in Xu [1994b]. The operator approach to the orthogonal polynomials in several variables also appears in Gekhtman and Kalyuzhny [1994].

For $d = 1$, the perturbation theory of self-adjoint operators can be applied to study the spectrum of Jacobi matrices, from which one can derive important results about the spectral measure. This is made possible by Weyl's theorem, which states that the essential spectrum of a self-adjoint operator does not change when a compact operator is added (cf. Riesz and Nagy [1955, p. 367]), and by the fact that the difference of two Jacobi matrices with respect to a large class of measures is a compact operator. The latter fact follows from the following limit relations on the coefficients of the three-term relation (2.8),

$$\lim_{n\to\infty} a_n = 1/2, \qquad \lim_{n\to\infty} b_n = 0, \tag{4.4}$$

which is satisfied by a wide class of measures, including all $\alpha \in \mathcal{M}$ such that $\alpha' > 0$ *a.e.* and $\operatorname{supp}\alpha = [-1,1]$. Moreover, the perturbation theorem shows that the relation (4.4) is a characteristic property of the class of measures in view of the spectrum. For the importance of the limit relation, we refer to Nevai [1979] and [1986] and the references given there. However, there seems to be no analog that plays the role of (4.4) in several variables. In fact, since the sizes of the coefficient matrices in the three-term relation (3.1) tend to infinity as $n$ goes to infinity, one can only deal with the limit of $\|A_{n,i}\|$ or the limit of some other unitarily invariant characters, such as $\det G_n$ or $\|G\|$, which are related to $A_{n,i}$. It's highly unlikely that a limit relation as such will determine the spectrum of a family of block Jacobi matrices, which can be any geometric region in $\mathbb{R}^d$. In this respect, one interesting question is to understand the interrelation among the block Jacobi matrices of the same commuting family, since whether the essential spectrum is, say, a simplex or a ball has to be reflected as some interrelations.

So far, there is little work done in this direction. This is partly due to the problem within the operator theory. In fact, there are several definitions of joint spectrum for a family of commuting operators. Moreover, it is not clear whether there is a proper perturbation theory. From the discussion above, it is no surprise that the difference of two block Jacobi matrices from different measures is not a compact operator in general, which can be easily seen by examining the example of product measures. It is likely that the non-uniqueness of orthogonal polynomials plays a role here. For each given measure, the block Jacobi matrices defined above are unique up to a unitary transformation, which can be ignored when we deal with the spectral measure. However, when we deal with the difference of two families of block Jacobi matrices from different measures, the unitary transformation may play an important role.

Since the block Jacobi matrices can be viewed as a prototype of a family of commuting operators, the study of them may yield new results in operator theory.

## 5 Common zeros of orthogonal polynomials

Zeros of a polynomial in $d$-variables are algebraic varieties of dimension $d-1$ or less; they are difficult to deal with for $d \ge 2$. Moreover, in dealing with zeros of orthogonal polynomials, we are mainly interested in the zeros that are real. In one variable, zeros of orthogonal polynomial are all real and distinct. For orthogonal polynomials in several variables, however, the right notion seems to be the *common zeros* of a family of polynomials. The simplest case is to consider the common zeros of $\mathbb{P}_n$, or equivalently, zeros of $V_n^d$, which we will deal with in this section.

For $d = 1$, it is well-known that the zeros of orthogonal polynomials are the eigenvalues of a truncated Jacobi matrix (cf. Chihara [1978]). It is remarkable that the same is true in several variables. Let $J_i$ be the block Jacobi matrices in the previous section. For $n \in \mathbb{N}_0$, the $n$-th truncated block Jacobi matrices, denoted by $J_{n,i}$, are given as follows

$$J_{n,i} = \begin{bmatrix} B_{0,i} & A_{0,i} & & & \bigcirc \\ A_{0,i}^T & B_{1,i} & A_{1,i} & & \\ & \ddots & \ddots & \ddots & \\ & & A_{n-3,i}^T & B_{n-2,i} & A_{n-2,i} \\ \bigcirc & & & A_{n-2,i}^T & B_{n-1,i} \end{bmatrix}, \qquad 1 \le i \le d .$$

We note that $J_{n,i}$ is a square matrix of order $N = \dim \Pi_{n-1}^d$. We say that $\Lambda = (\lambda_1, \dots, \lambda_d)^T \in \mathbb{R}^d$ is a *joint eigenvalue* of $J_{n,1}, \dots, J_{n,d}$, if there is an $\xi \neq 0$, $\xi \in \mathbb{R}^N$, such that $J_{n,i}\xi = \lambda_i \xi$ for $i = 1, \dots, d$; the vector $\xi$ is called a *joint eigenvector* associated to $\Lambda$. By a common zero $\mathbf{x} \in \mathbb{R}^d$ of $\mathbb{P}_n$ we mean a zero for every component of $\mathbb{P}_n$, *i.e.* $P_j^n(\mathbf{x}) = 0$ for all $1 \le j \le r_n^d$. The common zeros of $\mathbb{P}_n$ are characterized in the following theorem, which is proved using the three-term relation (3.1) and the recurrence relation (3.8).

**Theorem 5.1** (Xu [1994b]) *A point $\Lambda = (\lambda_1, \dots, \lambda_d)^T \in \mathbb{R}^d$ is a common zero of $\mathbb{P}_n$ if, and only if, it is a joint eigenvalue of $J_{n,1}, \dots, J_{n,d}$; moreover, a joint eigenvector of $\Lambda$ is $(\mathbb{P}_0^T(\Lambda), \dots, \mathbb{P}_{n-1}^T(\Lambda))^T$.*

Since all $J_{n,i}$ are real symmetric matrices, the theorem immediately implies that common zeros of $\mathbb{P}_n$ are all real. Setting both variables in the Christoffel-Darboux formula as one common zero of $\mathbb{P}_n$, where a standard limiting process is necessary, it follows readily that one of the partial derivatives of $\mathbb{P}_n$ is not identically zero, which implies that all zeros of $\mathbb{P}_n$ are simple. Due to the size of $J_{n,i}$, there can be at most $N$ common zeros of $\mathbb{P}_n$. Moreover, it follows that

**Theorem 5.2** (Xu [1994b]) *The orthogonal polynomial $\mathbb{P}_n$ has $N = \dim \Pi_{n-1}^d$ distinct real common zeros if, and only if,*

$$A_{n-1,i} A_{n-1,j}^T = A_{n-1,j} A_{n-1,i}^T, \quad 1 \le i, j \le d. \tag{5.1}$$

For $d = 2$, the condition (5.1) is equivalent to a set of conditions derived much earlier by Mysovskikh [1976] using an entirely different method, where he used monic orthogonal polynomials and did not use the matrix notation. The theorem has important applications in numerical cubature formulae.

Common zeros of orthogonal polynomials in several variables are first studied in connection with cubature formulae, where the word *cubature* stands for the higher dimensional quadrature. For a square positive linear functional $\mathcal{L}$, a cubature formula of degree $2n-1$ is a linear functional

$$\mathcal{I}_n(f) = \sum_{k=1}^{N} \lambda_k f(\mathbf{x}_k) , \quad \lambda_k > 0, \quad \mathbf{x}_k \in \mathbb{R}^d, \tag{5.2}$$

defined on $\Pi^d$, such that $\mathcal{L}(f) = \mathcal{I}_n(f)$ whenever $f \in \Pi_{2n-1}^d$, and $\mathcal{L}(f^*) \neq \mathcal{I}_n(f^*)$ for at least one $f^* \in \Pi_{2n}$. The points $\mathbf{x}_1, \dots, \mathbf{x}_N$ are called *nodes* and the numbers $\lambda_1, \dots, \lambda_N$ are called *weights*. Such a formula is called minimal, if $N$, the number of nodes, is minimal among all cubature formulae of degree $2n-1$. It is easily seen (cf. Stroud [1971]) that

$$N \ge \dim \Pi_{n-1}^d \tag{5.3}$$

in general. Indeed, if the number of nodes $N$ were less than $\dim \Pi_{n-1}^d$, then there would exist a polynomial $P \in \Pi_{n-1}^d$ vanishing on all nodes, which would then imply that $\mathcal{L}(P^2) = 0$. We call formulae attaining the lower bound (5.3) *Gaussian*, since for $d = 1$ these are the Gaussian quadrature formulae.

The problem of the existence of Gaussian cubature formulae was raised by Radon [1948]. The important contributions are made by, among others, Mysovskikh and his school (cf. Mysovskikh [1981]) and Möller ([1976] and [1979]). Let $\mathbb{P}_n$ be the vector of orthonormal polynomials with respect to $\mathcal{L}$. The following important theorem is due to Mysovskikh.

**Theorem 5.3** (Mysovskikh [1976]) *A Gaussian cubature formula exists if and only if* $\mathbb{P}_n$ *has* $N = \dim \Pi_{n-1}^d$ *common zeros.*

Since matrix multiplication is not commutative, it is evident from Theorem 5.2 that $\mathbb{P}_n$ does not have $N = \dim \Pi_{n-1}^d$ common zeros in general, which means that Gaussian cubature formulae do not exist in general. In his ground-breaking work on minimal cubature formulae, Möller [1976] showed that there is an improved lower bound for the number of nodes. For $d = 2$ this bound states, in the formulation of Xu [1994h], that

$$N \geq \dim \Pi_{n-1}^2 + \frac{1}{2} \operatorname{rank}(A_{n-1,1} A_{n-1,2}^T - A_{n-1,2} A_{n-1,1}^T), \tag{5.4}$$

which agrees with the lower bound (5.3) if and only if the Gaussian cubature formula exists. Moreover, Möller proved a surprising result which, for $d = 2$, states that

**Theorem 5.4** (Möller [1976]) *For a centrally symmetric linear functional* $\mathcal{L}$ *the lower bound* (5.4) *holds with*

$$\operatorname{rank}(A_{n-1,1} A_{n-1,2}^T - A_{n-1,2} A_{n-1,1}^T) = [n/2]. \tag{5.5}$$

A linear functional $\mathcal{L}$ on $\Pi^d$ is called centrally symmetric if it satisfies

$$\mathcal{L}(\mathbf{x}^\alpha) = 0, \quad \alpha \in \mathbb{N}^d, \quad |\alpha| = \text{odd integer}.$$

In the case that $\mathcal{L}$ is given by (2.1) with a weight function $W$, the central symmetry of $\mathcal{L}$ means that

$$\mathbf{x} \in \Omega \Rightarrow -\mathbf{x} \in \Omega \quad \text{and} \quad W(\mathbf{x}) = W(-\mathbf{x}),$$

where $\Omega \subset \mathbb{R}^d$ is the support set of $W$. Examples of centrally symmetric weight functions include the product weight functions $W(x,y) = w(x)w(y)$, where $w$ is a symmetric function on a symmetric interval $[-a, a]$, and the radial weight functions $W(\mathbf{x}) = w(\|\mathbf{x}\|)$ on Euclidean balls centered at origin. In terms of the three-term relation (3.1), it is shown in Xu [1994h] that

> *a square positive linear functional* $\mathcal{L}$ *is centrally symmetric if, and only if,* $B_{n,i} = 0$ *for all* $n \in \mathbb{N}_0$ *and* $1 \leq i \leq d$.

For $d > 2$, analogs of the lower bound (5.4) and the rank condition (5.5) are also given by Möller [1976] and [1979], see also Xu [1994h] where the bound is proved under the condition $B_{n,i}B_{n,j} = B_{n,j}B_{n,i}$. Moreover, the condition (5.5) has also been shown to be true for several families of non-centrally symmetric functions (Berens et al. [1995a] and Xu [1994g]). Whether the bound in (5.4) can be attained will be discussed in the following section. For now, we give a positive example in which the lower bound in (5.3) is attained.

From the result of Mysovskikh and Möller, it becomes clear that Gaussian cubature must be rare and there is no Gaussian cubature for most of the classical weight functions. In Mysovskikh and Chernitsina [1971], a Gaussian cubature formula of degree 5 is constructed numerically, which answers the original question of Radon. One natural question is whether there exists a weight function for which Gaussian cubature formulae exist for all $n$. The question is answered only recently in Berens et al. [1995b] and Schmid and Xu [1994], where two families of weight functions that admit Gaussian cubature formulae are given for $d = 2$ in Schmid and Xu [1994] and $d \geq 2$ in Berens et al. [1995b]. We describe the results below

We need to recall the definition of symmetric polynomials. A polynomial $P \in \Pi^d$ is called symmetric, if $P$ is invariant under any permutation of its variables. In particular, the degree of $P$, considered as a polynomial of variable $x_i$, $1 \leq i \leq d$, remains unchanged, we denote it by $\tau(P)$. The elementary symmetric polynomials in $\Pi^d$ are defined by

$$u_k := u_k(x_1, \ldots x_d) = \sum_{1 \leq i_1 < \ldots < i_k \leq d} x_{i_1} \cdots x_{i_k}, \qquad k = 1, 2, \ldots, d,$$

and any symmetric polynomial $P$ can be uniquely represented as

$$\sum_{\alpha_1 + \alpha_2 + \cdots + \alpha_d \leq \tau(P)} c_{\alpha_1, \ldots, \alpha_n} \, u_1^{\alpha_1} \cdots u_d^{\alpha_d};$$

where $\tau(P)$ is the degree of $P$. Let $\mathbf{x} = (x_1, x_2, \ldots, x_d)^T$ and $\mathbf{u} = (u_1, u_2, \ldots, u_d)^T$, where $u_k$ be the $k$-th symmetric polynomials in $x_1, \ldots, x_d$. We consider the mapping $\mathbf{x} \mapsto \mathbf{u}$. The Jacobian of $\mathbf{u} = \mathbf{u}(\mathbf{x})$ can be expressed as

$$J(\mathbf{x}) := \det \frac{\partial \mathbf{u}}{\partial \mathbf{x}} = \prod_{1 \leq i < j \leq d} (x_i - x_j).$$

Since $J^2$ is a symmetric polynomial, we shall further use the notation $\Delta(\mathbf{u}) := J^2(\mathbf{x})$. Let $D = \{\mathbf{x} \in \mathbb{R}^d : x_1 < x_2 < \ldots < x_d\}$. We define $R$ to be the image of $D$ under the mapping $\mathbf{x} \mapsto \mathbf{u}$; *i.e.*, $R = \mathbf{u}(D)$. Let $\mu$ be a nonnegative measure on $\mathbb{R}$ with finite moments and infinite support on $\mathbb{R}$. We define a measure $d\nu(\mathbf{u})$ on $R \subset \mathbb{R}^d$ by $d\nu(\mathbf{u}) = d\mu(\mathbf{x}) := d\mu(x_1) \cdots d\mu(x_d)$; *i.e.* $\nu$ is the image of the product measure under the mapping $\mathbf{x} \to \mathbf{u}$.

**Theorem 5.5** (Berens et al. [1995b]) *Let $\nu$ be defined as above. The orthogonal polynomials $\mathbb{P}_n$ with respect to the measure $[\Delta(\mathbf{u})]^{\frac{1}{2}} d\nu(\mathbf{u})$ or $[\Delta(\mathbf{u})]^{-\frac{1}{2}} d\nu(\mathbf{u})$ have $N = \dim \Pi_{n-1}^d$ many common zeros. Consequently, there exist Gaussian cubature formulae for these measures.*

In fact, the above construction uses the structure of the tensor product of orthogonal polynomials and quadrature formulae. The tensor product is usually not suitable for generating cubature formulae, since the degree of the polynomials tends to be too high; however, the symmetric mapping $\mathbf{x} \mapsto \mathbf{u}$ maps the tensor product into a system of polynomials with proper degree. For example, let $\{p_n\}_{n=0}^\infty$ be the orthonormal polynomials with respect to $d\mu$. Then for $\alpha = (\alpha_1, \ldots, \alpha_d)$, $0 \leq \alpha_1 \leq \ldots \leq \alpha_d = n$, the polynomials $P_\alpha^n$ defined by,

$$P_\alpha^n(\mathbf{u}) = \sum_{\beta \in \{1,2,\ldots,d\}} p_{\alpha_1}(x_{\beta_1}) \cdots p_{\alpha_d}(x_{\beta_d}),$$

where $\Sigma$ is over all permutation $\beta$, are orthogonal with respect to $[\Delta(\mathbf{u})]^{-\frac{1}{2}}d\nu(\mathbf{u})$. For $\alpha = (\alpha_1, \ldots, \alpha_d)$, $0 \le \alpha_1 < \ldots < \alpha_d = n + d - 1$ and $n \in \mathbb{N}_0$,

$$P_\alpha^n(\mathbf{u}) = \frac{V_\alpha^n(\mathbf{x})}{J(\mathbf{x})}, \qquad V_\alpha^n(\mathbf{x}) = \det(p_{\alpha_i}(x_j))_{i,j=1}^d.$$

are orthogonal with respect to $[\Delta(\mathbf{u})]^{\frac{1}{2}}d\nu(\mathbf{u})$. From this point of view, the symmetric orthogonal polynomials so defined are the most natural extensions of orthogonal polynomials in one variable.

Questions on cubature formulae of degree $2n-1$ can also be asked for cubature formulae of degree $2n-2$. Although there are interesting examples that distinguish the cubature formulae of even degree from those of odd degree, the theorems are of the similar nature. We will not go into details, but refer to the results in Schmid [1978], Xu [1994d] and [1994h].

## 6 Common zeros and cubature formulae

After the discussion in the previous section, it is natural to ask how many common zeros $\mathbb{P}_n$ can have if it does not have $\dim \Pi_{n-1}^d$ many. The answer, however, is a little surprising; the following result is proved in Xu [1994h]:

> *Let $\mathcal{L}$ be centrally symmetric. If $n$ is an even integer, then $\mathbb{P}_n$ has no common zero; if $n$ is an odd integer, then $\mathbb{P}_n$ has the origin as the only common zero.*

This result indicates that it is often too much to ask for common zeros of *all* polynomials of $\mathbb{P}_n$, or *all* polynomials in $V_n^d$. Instead, one might look for common zeros of a part of $V_n^d$, such as common zeros of $\mathcal{U}_n^d$ spanned by

$$U^T\mathbb{P}_n, \qquad \text{where} \qquad U : r_n^d \times (r_n^d - \sigma)$$

where the matrix $U$ has full rank $r_n^d - \sigma$. Such a consideration is important in the study of cubature formula. Möller [1976] characterized the minimal cubature formulae that attain his lower bound (5.4). The nodes are shown to be common zeros of $U^T\mathbb{P}_n$ for some $U$ with $\sigma = [n/2]$. His characterization is derived under the assumption that the span of $\{xU^T\mathbb{P}_n, yU^T\mathbb{P}_n\}$ contains $r_{n+1}^2$ many linearly independent polynomials of degree $n+1$; in other words, that $U^T\mathbb{P}_n$ generates the polynomial ideal that consists of all polynomials vanishing on the nodes of the cubature formula. Roughly speaking, such an assumption requires that $\mathcal{U}_n$ contains at least half of the polynomials in $V_n^d$; and to attain the lower bound (5.4), $\mathcal{U}_n$ has to have almost exactly half of the polynomials in $V_n^d$.

Clearly, such an assumption may not be satisfied in general. Indeed, the lower bound (5.4) is known to be attained only in a few cases (cf. Möller [1976], Morrow and Patterson [1978] and Cools and Schmid [1989]), and it has been shown to be not attainable for various classical weight functions (cf. Morrow and Patterson [1978], Verlinden and Cools [1992] and Cools and Schmid [1993]). When a cubature formula has more nodes than the lower bound (5.4), $\mathcal{U}_n$ will have fewer polynomials in $V_n^d$ and the nodes of the cubature formula will be determined as common zeros of $U^T\mathbb{P}_n$ and some polynomials of degree greater than $n$; together they span the ideal determined by the nodes. Among these polynomials, those of degree $n+1$ are necessarily orthogonal to polynomials of degree $n-2$, since the cubature formula is of order $2n-1$. They are what we would call quasi-orthogonal polynomials of order 2. In general, polynomials that determine the nodes of a cubature formula may involve quasi-orthogonal polynomials of higher order and several consecutive degrees higher than $n$. The above consideration motivates the study of common zeros of a set of quasi-orthogonal polynomials in Xu [1994h].

For fixed $n$, we denote the vector of quasi-orthogonal polynomials of degree $n+s$, which are orthogonal to polynomials in $\Pi^d_{n-1-s}$, as

$$\mathbb{Q}_{n+s} = \mathbb{P}_{n+s} + \Gamma_{1,s}\mathbb{P}_{n+s-1} + \ldots + \Gamma_{2s,s}\mathbb{P}_{n-s}.$$

In Xu [1994h], the common zeros of the polynomials in the set

$$\mathcal{Q}_r = \{U_0^T\mathbb{P}_n, U_1^T\mathbb{Q}_{n+1}, \ldots, U_r^T\mathbb{Q}_{n+r}, \mathbb{Q}_{n+r+1}\}, \tag{6.1}$$

are studied, where $U_k$, $0 \le k \le r$, are matrices such that

$$U_k : r^d_{n+k} \times r^d_{n+k} - \sigma_k, \qquad \operatorname{rank} U_k = r^d_{n+k} - \sigma_k, \quad \sigma_k \ge 0.$$

We note that if $U_k$ is the identity matrix, then the set of common zeros of $\mathcal{Q}_r$ will be the same as that of $\mathcal{Q}_{r-1}$. If a cubature formula is based on the zeros of $\mathcal{Q}_r$, then we say that the cubature formula is generated by the zeros of $\mathcal{Q}_r$. Moreover, we call $\mathcal{Q}_r$ *maximal*, if for $1 \le k \le r$ every polynomial in $\Pi^d_{n+k}$ vanishing on all the common zeros of $\mathcal{Q}_r$ belongs to the linear space spanned by $U_0^T\mathbb{P}_n, U_1^T\mathbb{Q}_{n+1}, \ldots, U_{n+k}^T\mathbb{Q}_{n+k}$. The main result in Xu [1994h] characterizes the common zeros of $\mathcal{Q}_r$ that generates cubature formulas in terms of the nonlinear matrix equations satisfied by the matrices $\Gamma_i$ and $V_k$, where $V_k$ are orthogonal compliments of $U_k$, defined by

$$U_k^T V_k = 0, \qquad V_k : r^d_n \times \sigma_k.$$

The statement in the general case is rather complicated, we shall restrict ourselves to the case $r = 0$, where we deal with

$$\mathcal{Q}_0 = \{U^T\mathbb{P}_n, \mathbb{Q}_{n+1}\}, \qquad \mathbb{Q}_{n+1} = \mathbb{P}_{n+1} + \Gamma_1\mathbb{P}_n + \Gamma_2\mathbb{P}_{n-1}, \tag{6.2}$$

and we write $U$ and $V$ instead of $U_0$ and $V_0$ in this case. we have

**Theorem 6.1** (Xu [1994h]) *The set $\mathcal{Q}_0 = \{U^T\mathbb{P}_n, \mathbb{Q}_{n+1}\}$ is maximal and it has $\dim \Pi^d_{n-1} + \sigma$ many pairwise distinct real zeros that generate a cubature formula of degree $2n-1$ if, and only if, there exists $V$ such that $U^TV = 0$,*

$$\Gamma_2 = \sum_{i=1}^d D^T_{n,i}(I - VV^T)A^T_{n-1,i}, \tag{6.3}$$

*and $\Gamma_1$ and $V$ satisfy the following conditions:*

$$A_{n-1,i}(VV^T - I)A^T_{n-1,j} = A_{n-1,j}(VV^T - I)A^T_{n-1,i}, \quad 1 \le i,j \le d, \tag{6.4}$$

$$(B_{n,i} - A_{n,i}\Gamma_1)VV^T = VV^T(B_{n,i} - \Gamma_1^T A^T_{n,i}), \qquad 1 \le i \le d, \tag{6.5}$$

*and*

$$\begin{aligned} &VV^TA^T_{n-1,i}A_{n-1,j}VV^T + (B_{n,i} - A_{n,i}\Gamma_1)VV^T(B_{n,j} - \Gamma_1^TA^T_{n,j}) \\ &= VV^TA^T_{n-1,j}A_{n-1,i}VV^T + (B_{n,j} - A_{n,j}\Gamma_1)VV^T(B_{n,i} - \Gamma_1^TA^T_{n,i}) \end{aligned} \tag{6.6}$$

*for $1 \le i,j \le d$.*

There is an analogous theorem for $\mathcal{Q}_r$ with many more matrix equations that involve $\Gamma_{i,k}$ and $V_k$. The lower bound (5.4) is attained if and only if $\Gamma_1 = 0$ and $\sigma = [n/2]$ in the above theorem; this special case has been treated by Möller. However, in the present general form, the theorem does not deal with just minimal cubature formulae; in fact, it characterizes *every* (positive) cubature formula generated by $\mathcal{Q}_0$. With the additional $\Gamma_1$ and the flexibility on the value of $\sigma$, it is likely that the equations (6.4)–(6.6) are solvable even if the special case that (5.4) is attained is not solvable. Still, there is no warrant that the system of equations is always solvable; the question depends on the coefficient matrices of the three-term relation. When the system is not solvable, one needs to consider $\mathcal{Q}_r$ for $r \ge 1$.

In Xu [1994h], the sufficiency part of (6.3)–(6.6) in the theorem is contained in two theorems, Theorem 4.1.4, which characterizes the common zeros of $\mathcal{Q}_0$, and Theorem 4.4.3, which established the existence of the cubature formula, necessarily positive, generated by the common zeros. The first theorem also contains a statement of the necessity part which, however, has an error; namely, the clause *that generates a cubature formula of degree* $2n-1$ is missing. The same clause should also be added to the statement of Theorem 7.1.4 in Xu [1994h] which deals with $\mathcal{Q}_r$ in general. The necessity part is the easier half; we give a complete proof below.

**Proof of the necessity part of Theorem 6.1.** Since $\mathbf{x}_k$ generates a cubature formula, the linear functional $\mathcal{I}_n(f)$ in (5.2) is square positive. Moreover, that $\mathcal{Q}_0$ is maximal implies that the matrix $\mathcal{I}_n(\mathbb{P}_n\mathbb{P}_n^T)$ is of rank $\sigma$. Therefore, there exists a matrix $V : r_n^d \times \sigma$ of rank $\sigma$ such that $\mathcal{I}_n(\mathbb{P}_n\mathbb{P}_n^T) = VV^T$. Since $U^T\mathbb{P}_n(\mathbf{x}_k) = 0$ and $V$ is of full rank, it follows easily that $U^TV = 0$. For this $V$ we verify the validity of (6.3)–(6.6) by computing the matrices $\mathcal{I}_n(x_i\mathbb{P}_{n-1}\mathbb{P}_n^T)$ and $\mathcal{I}_n(x_ix_j\mathbb{P}_{n-1}\mathbb{P}_{n-1}^T)$, $\mathcal{I}_n(x_i\mathbb{P}_n\mathbb{P}_n^T)$ and $\mathcal{I}_n(x_ix_j\mathbb{P}_n\mathbb{P}_n^T)$, respectively. Since the cubature formula is of degree $2n-1$, it follows readily that $\mathcal{I}_n(\mathbb{P}_{n-1}\mathbb{P}_n^T) = \mathcal{L}(\mathbb{P}_{n-1}\mathbb{P}_n^T) = 0$. Moreover, since $\mathbb{Q}_{n+1}$ vanishes on $\mathbf{x}_k$, it follows that $\mathcal{I}_n(\mathbb{Q}_{n+1}\mathbb{P}_n^T) = 0$. Therefore, using the three-term relation and the definition of $\mathbb{Q}_{n+1}$, we have, for example,

$$\begin{aligned}\mathcal{I}_n(x_i\mathbb{P}_n\mathbb{P}_n^T) =& A_{n,i}\mathcal{I}_n(\mathbb{P}_{n+1}\mathbb{P}_n^T) + B_{n,i}\mathcal{I}_n(\mathbb{P}_n\mathbb{P}_n^T)\\ =& A_{n,i}\mathcal{I}_n[(\mathbb{Q}_{n+1} - \Gamma_1\mathbb{P}_n - \Gamma_2\mathbb{P}_{n-1})\mathbb{P}_n^T] + B_{n,i}\mathcal{I}_n(\mathbb{P}_n\mathbb{P}_n^T)\\ =& (B_{n,i} - A_{n,i}\Gamma_1)VV^T\end{aligned}$$

from which (6.5) follows since the matrix on the left hand side is clearly symmetric. Other matrices are computed similarly, but the computation involves two ways of using the three-term relation; the desired equations are simply the consequence that the two ways should yield the same result. □

In the proof of this direction in Xu [1994h], we assumed that the matrices $S_{n,i}$ defined below are symmetric, which can be justified, when the additional clause is assumed, by computing $\mathcal{I}_n(x_i\mathbb{P}_{n-1}\mathbb{P}_n^T)$ and $\mathcal{I}_n(x_i\mathbb{P}_n\mathbb{P}_n^T)$ as in the above proof.

The proof of the other direction of the theorem is long. In Xu [1994h] it is proved following a approach which is very different from the method introduced for the special case in Möller [1976] and followed by others (cf. Morrow and Patterson [1978] and Schmid [1978]). It is interesting to mention that the approach in Xu [1994h] resembles the approach in one variable very closely (cf. Davis and Rabinowitz [1975] and Xu [1994f]). Assume that (6.4)–(6.6) are satisfied by $V_0$ and $\Gamma_1$, then the approach consists of three steps. As the first step, the common zeros are shown to be exist as the joint eigenvalues of the matrices $S_{n,i}$, defined by

$$S_{n,i} = \begin{bmatrix} B_{0,i} & A_{0,i} & & & \bigcirc \\ A_{0,i}^T & B_{1,i} & A_{1,i} & & \\ & \ddots & \ddots & \ddots & \\ & & & B_{n-1,i} & A_{n-1,i}V \\ \bigcirc & & & V^+A_{n-1,i}^{*T} & V^+B_{n,i}^*V \end{bmatrix}, \qquad 1 \le i \le d,$$

where $B_{n,i}^* = B_{n,i} - A_{n,i}\Gamma_1$ and $A_{n-1,i}^* = A_{n-1,i}^T - A_{n,i}\Gamma_1$. All common zeros are real, since $S_{n,i}$ are symmetric according to (6.4) and (6.5), and they are pairwise distinct, by a modified Christoffel-Darboux formula for

$$\mathbf{K}_n^{(0)}(\mathbf{x},\mathbf{y}) = \mathbf{K}_n(\mathbf{x},\mathbf{y}) + [V^+\mathbb{P}_n(\mathbf{x})]^T V^+\mathbb{P}_n(\mathbf{y}),$$

where $V^+ = (V^TV)^{-1}V^T$ is the generalized inverse of $V$. It is worthwhile to point out that the theorem for the special case in Möller [1976] is stated under the restriction that the zeros are all real, which is removed later in Schmid [1978] as a consequence of a theorem from real algebraic ideal theory; it is not clear whether the theorem from the ideal theory can possibly be applied to the present general setting. Back to the proof of the theorem, the second step shows that the Lagrange interpolation based on the common zeros of $\mathcal{Q}_0$ exists and is unique in the polynomial subspace $\Pi_{n-1}^d \cup \operatorname{span}\{V^+\mathbb{P}_n\}$; moreover, if the common zeros are denoted by $\mathbf{x}_1, \ldots, \mathbf{x}_N$, then the interpolation polynomial, denoted by $L_nf$, can be written down explicitly as

$$L_n(f,\mathbf{x}) = \sum_{k=1}^{N} f(\mathbf{x}_k)\frac{\mathbf{K}_n^{(0)}(\mathbf{x},\mathbf{x}_k)}{\mathbf{K}_n^{(0)}(\mathbf{x}_k,\mathbf{x}_k)}, \tag{6.7}$$

where the interpolation $L_n(f,\mathbf{x}_k) = f(\mathbf{x}_k)$, $1 \le k \le N$, is verified via the modified Christoffel-Darboux formula. In the third step, the cubature formula generated by the common zeros is shown to exist by applying the linear functional $\mathcal{L}$ on $L_nf$. As a by product, it follows readily that the cubature weights are given by the formula

$$\lambda_k = [\mathbf{K}_n^{(0)}(\mathbf{x}_k,\mathbf{x}_k)]^{-1},$$

which are clearly positive and the formula is shown for the first time in Xu [1994h]. It is worthwhile to mention that the Lagrange interpolation polynomials defined in (6.7) are of interests, since interpolation in several variables is very difficult in general and compact formulae are rare. For the product Chebyshev polynomials of the first kind, there exists a well-defined sequence of polynomials $L_nf$ which is proved in Xu [199xa] to converge to $f$ in weighted $L^p$ norm, $0 < p < \infty$, for every continuous function $f$.

The theorem gives a complete characterization for the common zeros of $\mathcal{Q}_0$ that generates a cubature formula. A central question is whether $\mathcal{Q}_0$ always has enough common zeros for some $U$ and $\Gamma_1$, which is equivalent to ask whether (6.4)–(6.6) is always solvable. So far, they are solved only for the product Chebyshev weight functions, $(1-x^2)^{1/2}(1-y^2)^{1/2}$ and $(1-x^2)^{-1/2}(1-y^2)^{-1/2}$, and some simple weight functions for small $n$; moreover, the previous efforts are mostly on the special case that (5.4) is attained, which implies that $\Gamma_1 = 0$, and there are a number of negative results for the case. There is reason to be optimistic about solving (6.4)–(6.6) with some nonzero $\Gamma_1$, although it is likely that the resulting cubature formulae may not be the minimal ones. It should be mentioned that the system (6.4)–(6.6) in its full generality has not yet been explored.

## 7 Fourier orthogonal series and asymptotics

As we mentioned in Section 5, there is no analog of the limit relations (4.4) in several variable. These relations play an important role in proving the asymptotic relations related to orthogonal polynomials of one variable. Let $\mathbf{K}_n(\cdot,\cdot)$ be the reproducing kernel function. The function

$$\Lambda_n(\mathbf{x}) = [\mathbf{K}_n(\mathbf{x},\mathbf{x})]^{-1},$$

is called the *Christoffel function*; it has the following fundamental property.

*If $\mathcal{L}$ is square positive, then*

$$[\mathbf{K}_n(\mathbf{x},\mathbf{x})]^{-1} = \min\{\mathcal{L}(P^2) : P \in \Pi^d_{n-1}, \quad P(\mathbf{x}) = 1\}$$

For $d = 1$, this function plays a significant role in the study of orthogonal polynomials of one variable, which we refer to the extensive survey Nevai [1986]. The limit relation (4.4) is used in proving the following important property of $\lambda_n$, the usual notation for $\Lambda_n$ in $d = 1$,

$$\lim_{n\to\infty} n\lambda_n(x) = \pi\alpha'(x)(1-x^2)^{1/2} \qquad a.e.\ \ x \in [-1,1]$$

for measures $\alpha$ belonging to Szegő's class, i.e., $\log\alpha'(\cos t) \in L^1[0,2\pi]$ (Máté et al. [1991]). Such a relation, like many other limit relations involving orthogonal polynomials on the real line, is proved by first establishing an analogous result for orthogonal polynomials on the unit circle, then coming back to the real line by using the so-called $*$-transform in Szegő's theory. Although part of Szegő's theory undoubtly can be extended to the setting in several complex variables, a moment reflection shows that $*$-transform as a bridge between the real and the complex does not extend to several variables. So far, only very limited effort has been put into proving the limit relation in several variables; namely, certain product weight functions in Xu [1995] and the radial weight function $(1-|\mathbf{x}|^2)^\alpha$ on the unit ball in Bos [1994]. In general, for $\Lambda_n(\cdot)$ associated with orthogonal polynomials with respect to $W$, the limit relation should take the form

$$\lim_{n\to\infty} \binom{n+d}{d} \Lambda_n(\mathbf{x}) = C_d(\mathbf{x})W(\mathbf{x})$$

where $C_d$ is independent of $W$. For the product weight functions in Xu [1995], we have $C_d(\mathbf{x}) = 1/W_0(\mathbf{x})$ where $W_0$ is the product Chebyshev weight function of the first kind,

$$W_0(\mathbf{x}) = \frac{1}{\pi^d}\frac{1}{\sqrt{1-x_1^2}}\cdots\frac{1}{\sqrt{1-x_d^2}}, \qquad \mathbf{x} = (x_1,\ldots,x_d) \in [-1,1]^d;$$

for the radial weight function $(1-|\mathbf{x}|^2)^\alpha$ on the unit ball in $\mathbb{R}^d$, we have that $C_d(\mathbf{x}) = c_d/(1-|\mathbf{x}|^2)^{-1/2}$ where $c_d$ is a constant chosen so that the integral of $C_d$ is 1. One interesting remark is that for the weight function $(1-|\mathbf{x}|^2)^\alpha$ the limit

$$\lim_{n\to\infty} \frac{1}{\binom{n+d-1}{d-1}} \mathbb{P}_n^T(\mathbf{x})\mathbb{P}_n(\mathbf{x}) = \frac{1}{C_d(\mathbf{x})W(\mathbf{x})}$$

exists, which is not true for $d = 1$, where the asymptotic of $p_n$ is much more complicated. Of course, with $\mathbb{P}_n^T \mathbb{P}_n$ we are dealing with a sum instead of an individual term.

On the other hand, the order of $\Lambda_n$ as $n$ goes to infinity is enough to give a sufficient condition for the convergence of orthogonal series. For many weight functions with compact support, it is shown in Xu [1995] that $\Lambda_n(\mathbf{x}) = \mathcal{O}(n^{-d})$. The sufficient condition for the convergence of $S_n(f)$ is derived using the $L^2$ theory as in the case of $d = 1$ (cf. Nevai [1986]). As an example, we give one of the consequences as follows.

**Theorem 7.1** (Xu [1995]) *Let $\mu \in \mathcal{M}(\mathbb{R}^d)$, $\operatorname{supp}\mu = [-1,1]^d$, and suppose that $[n^d \Lambda_n(\mathbf{x})]^{-1}$ is bounded uniformly on a subset $\Delta$ of $[-1,1]^d$. Suppose $f \in C^{[d/2]}([-1,1]^d)$ and each of its $[d/2]$-th derivatives satisfies*

$$|D^{[d/2]} f(\mathbf{x}) - D^{[d/2]} f(\mathbf{y})| \le c h^\beta, \quad |\mathbf{x} - \mathbf{y}| \le h,$$

*where for odd $d$, $\beta > 1/2$, and for even $d$, $\beta > 0$. Then $S_n(f)$ converges uniformly and absolutely to $f$ on $\Delta$.*

Because of the dependence of the asymptotic order of $\Lambda_n$ on the dimension, the requirement on the class of functions in the theorem becomes stronger for higher dimension. The use of $L^2$ theory also implies similar results for the first Cesáro means of the Fourier orthogonal series. These results provide sufficient conditions for the convergence with little restriction on the class of measures; being so, they are usually not sharp. To fully understand the expansion of Fourier orthogonal series, one needs to go back to the reproducing kernel function $\mathbf{K}_n(\cdot,\cdot)$.

For $d = 1$, the Christoffel-Darboux formula (3.12) provides a compact formula of $\mathbf{K}_n(\cdot,\cdot)$ that contains only one term. For $d \ge 2$, however, the nominator in the formula (3.12) is a sum of $r_{n-1}^d = \mathcal{O}(n^{d-1})$ many terms, while the number of terms in $\mathbf{K}_n(\cdot,\cdot)$ is $\dim \Pi_{n-1}^d = \mathcal{O}(n^d)$. In order to understand the behavior of $S_n(f)$, it is necessary to have a compact formula for $\mathbf{K}_n(\cdot,\cdot)$. The first such compact formula is given in Xu [1995] for the product Chebyshev weight function of the first kind. Let $\mathbf{x} = (\cos\theta_1, \dots, \cos\theta_d)$ and $\mathbf{y} = (\cos\phi_1, \dots, \cos\phi_d)$. Then we have

*for the product Chebyshev weight function $W_0$,*

$$\mathbf{K}_n(\mathbf{x},\mathbf{y}) = \sum_{(\varepsilon_1,\dots,\varepsilon_d)\in\{-1,1\}^d} D_{n,d}(\theta_1 + \varepsilon_1\phi_1, \dots, \theta_d + \varepsilon_d \phi_d), \tag{7.1}$$

*where the function $D_{n,d}$ is a divided difference*

$$D_{n,d}(\theta_1, \dots, \theta_d) = [\cos\theta_1, \dots, \cos\theta_d]\, G_{n,d},$$

*and*

$$G_{n,d}(\cos t) = (-1)^{[\frac{d-1}{2}]} 2 \cos\frac{t}{2} (\sin t)^{d-2} \begin{cases} \cos(n-\frac{1}{2})t, & \text{for } d \text{ even} \\ \sin(n-\frac{1}{2})t, & \text{for } d \text{ odd.} \end{cases}$$

In fact, the function $D_{n,d}$ is the Dirichlet kernel of the $\ell$-1 summability of the multiple Fourier series on $\mathbb{T}^d$;

$$D_{n,d}(\Theta) = \sum_{|\alpha|_1 \le n-1} e^{i\alpha\cdot\Theta}, \qquad \Theta \in \mathbb{T}^d,$$

where $|\alpha|_1 = |\alpha_1| + \ldots + |\alpha_d|$ for $\alpha \in \mathbb{Z}^d$. It turns out that the $\ell$-1 summability is very different from the usual summability of the spherical means of multiple Fourier series. For example, there is no critical index for the Riesz means in the $\ell$-1 summability. We refer to Berens and Xu [1995], [1996] and the references there for the results. Here we just point out that overall the $\ell$-1 summability is very much similar to the summability of Fourier series in one variable, which seems to indicate that we can expect a summability theory for orthogonal series in several variable similar to that of one variable. We quote one result here.

**Theorem 7.2** (Berens and Xu [1996]) *For $W_0$, the Cesàro $(C, 2d-1)$ means of $S_n(f)$ define a positive operator; the order of summability is best possible in the sense that the $(C, \delta)$ means are not positive for $0 < \delta < 2d - 1$.*

For $d = 1$ this is Fejér's theorem which states that the arithmetic means of the Fourier partial sum is positive. The proof of the theorem uses an inequality of Askey-Gasper on Jacobi polynomials (cf. Askey [1975]).

For the product Gegenbauer weight functions $(1-x_1^2)^\alpha \ldots (1-x_d^2)^\alpha$ on $[-1,1]^d$, the compact formula similar to that of (7.1) holds for $2\alpha \in \mathbb{N}_0$; the formula for other values of $\alpha$ has yet to be found. In fact, there is little results known for the summability of Fourier orthogonal expansion with respect to these weight functions. For the weight function

$$W_\mu(\mathbf{x}) = w_\mu(1-|\mathbf{x}|^2)^{\mu-1/2}, \qquad \mathbf{x} \in B^d = \{\mathbf{x} \in \mathbb{R}^d : |\mathbf{x}|^2 = x_1^2 + \ldots + x_d^2 \le 1\},$$

where $w_\mu$ is the normalization constant so that the integral of $W_\mu$ over $B^d$ is 1, there is an interesting compact formula of $\mathbf{K}_n(\cdot,\cdot)$ which has been discovered by the author recently in [199xb]. In fact, a compact formula holds already for $\mathbb{P}_n^T(\mathbf{x})\mathbb{P}_n(\mathbf{y})$.

Let $C_n^\lambda$ denotes the usual Gegenbauer polynomials. Then the compact formula for $\mathbb{P}_n^T(\mathbf{x})\mathbb{P}_n(\mathbf{y})$ is as follows.

$$\begin{aligned}\mathbb{P}_n^T(\mathbf{x})\mathbb{P}_n(\mathbf{y}) =& \frac{n+\mu+\frac{d-1}{2}}{\mu+\frac{d-1}{2}} \int_{-1}^1 C_n^{\mu+\frac{d+1}{2}}(\mathbf{x}\cdot\mathbf{y} + \sqrt{1-|\mathbf{x}|^2}\sqrt{1-|\mathbf{y}|^2}\,t) \\ & \times (1-t^2)^{\mu-1}dt \Big/ \int_{-1}^1 (1-t^2)^{\mu-1}dt, \qquad \mu > 0 \qquad \mathbf{x},\mathbf{y} \in B^d.\end{aligned}$$

A simpler formula holds for the limiting case $\mu = 0$. It's worthwhile to mention that for $d = 1$ this formula coincides with the product formula of the ultraspherical polynomials. Moreover, by setting $|\mathbf{y}| = 1$, the integral in the formula can be removed; we see the formula resembles the addition formula for the spherical harmonics (cf. Erdélyi et al. [1953, Vol. II, p. 244, (2)]). From the formula, a theorem similar to that of Theorem 7.2 follows. More importantly, it allows us to give a complete answer for the Cesáro summability of the Fourier orthogonal series.

**Theorem 7.3** (Xu [199xb]) Let $f$ be continuous on the closed ball $B^d$. The expansion of $f$ in the Fourier orthogonal series with respect to $W_\mu$ is uniformly $(C, \delta)$ summable on $B^d$ if, and only if, $\delta > \mu + \frac{d-1}{2}$.

Other properties and results have been derived from the compact formula in an ongoing work of the author.

Comparing to our understanding of the classical orthogonal polynomials in one variable, we know little about their extensions in several variables. The orthogonal polynomials with respect to the classical weight functions, starting from those in

Koornwinder [1974], should be studied in detail. The knowledge we obtain from studying them not only will enhance our understanding of the general theory, it may also lead to new directions in developing a general theory of orthogonal polynomials in several variables. In this respect, we should at least mention the theory of orthogonal polynomials on spheres with respect to measures invariant under a reflection group. This important theory has been developed by Dunkl recently (cf. Dunkl [1989] and [1991]). Among other things, it opens a way to study orthogonal polynomials on spheres and balls with respect to a large class of weight functions.

## References

Akheizer, N. I. [1965] *The classical moment problem and some related questions in Analysis*, Hafner, New York.

Askey, R. [1975] *Orthogonal polynomials and special functions*, SIAM, Philadelphia.

Berg, C. [1987] *The multidimensional moment problem and semigroups*, Moments in Mathematics, Proc. Sympos. Appl. Math., vol. 37,, Amer. Math. Soc., Providence, RI, pp. 110–124..

Berg, C. and M. Thill [1991] *Rotation invariant moment problems*, Acta Math **167**, 207–227.

Berens, H. and Y. Xu [1995] *$\ell$-1 summability for multivariate Fourier integral*, Approximation Theory VIII (C. K. Chui and L. Schumaker, eds.), World Sci. Publ. Co., pp. 55–62.

Berens, H. and Y. Xu [1996] *Fejér means for multivariate Fourier series*, Math. Z. **221**, 449–465.

Berens, H., H. Schmid and Y. Xu [1995a] *On twodimensional definite orthogonal systems and on a lower bound for the number of nodes of associated cubature formulae*, SIAM J. Math. Anal. **26**, 468–487.

Berens, H., H. Schmid and Y. Xu [1995b] *Multivariate Gaussian cubature formulae*, Arch. Math **64**, 26–32.

Bos, L. [1994] *Asymptotics for the Christoffel function for Jacobi like weights on a ball in $\mathbb{R}^m$*, New Zealand J. Math. **23**, 99–109.

Chihara, T. S. [1978] *An introduction to orthogonal polynomials*, Gordon and Breach, New York.

Cools, R. and H. J. Schmid [1989] *Minimal cubature formulae of degree $2k-1$ for two classical functions*, Computing **43**, 141–157.

Cools, R. and H. J. Schmid [1993] *A new bound for the number of nodes in cubature formulae of degree $4n+1$ for some circularly symmetric integrals*, Numerical Integration IV (H. Braßand G. Hämmerlin, eds.), Birkhüser, Basel, pp. 57–66.

Davis, P. J. and P. Rabinowitz [1975] *Methods of numerical integration*, Academic Press, New York.

Dombrowski, J. [1990] *Orthogonal polynomials and functional analysis*, Orthogonal Polynomials (P. Nevai, ed.), Kluwer Acad. Pub., Netherlands, pp. 147–161.

Dunkl, C [1989] *Differential-difference operators associated to reflection groups*, Trans. Amer. Math. Soc. **311**, 167-183.

Dunkl, C [1991] *Integral kernels with reflection group invariance*, Can. J. Math. **43**, 1213-1227.

Erdélyi, A., W. Magnus, F. Oberhettinger and F. G. Tricomi [1953] *Higher transcendental functions*, vol. 2, McGraw-Hill, New York.

Engles, H. [1980] *Numerical quadrature and cubature*, Academic Press, New York.

Fuglede, B. [1983] *The multidimensional moment problem*, Expo. Math. **1**, 47–65.

Gekhtman, M. I. and A.A. Kalyuzhny [1994] *On the orthogonal polynomials in several variables*, Integr. Equat. Oper. Theory **19**, 404–418.

Koornwinder, T. [1975] *Two-variable analogues of the classical orthogonal polynomials*, Theory and Application of Special Functions (R. A. Askey, ed.), Academic Press, New York, pp. 435–495.

Kowalski, M. A. [1982a] *The recursion formulas for orthogonal polynomials in n variables*, SIAM J. Math. Anal. **13**, 309–315.

Kowalski, M. A. [1982b] *Orthogonality and recursion formulas for polynomials in n variables*, SIAM J. Math. Anal. **13**, 316–323.

Krall, H. L. and I. M. Sheffer [1967] *Orthogonal polynomials in two variables*, Ann. Mat. Pura. Appl. **76**, no. 4, 325–376.

Máté, A, P. Nevai and V. Totik [1991] *Szegő's extremum problem on the unit circle*, Ann. of Math. **134**, 433–453.

Möller, H. M. [1976] *Kubaturformeln mit minimaler Knotenzahl*, Numer. Math. **35**, 185–200.

Möller, H. M. [1979] *Lower bounds for the number of nodes in cubature formulae*, Numerical Integration, ISNM vol. 45 (G. Hämmerlin, ed.), Birkhäuser, Basel.

Morrow, C. R. and T. N. L. Patterson [1978] *Construction of algebraic cubature rules using polynomial ideal theory*, SIAM J. Numer. Anal. **15**, 953–976.

Mysovskikh, I. P. [1976] *Numerical characteristics of orthogonal polynomials in two variables*, Vestnik Leningrad Univ. Math. **3**, 323–332.

Mysovskikh, I. P. [1981] *Interpolatory cubature formulas*, "Nauka", Moscow. (Russian)

Mysovskikh, I. P. and V. Ya. Chernitsina [1971] *Answer to a question of Radon*, Dokl. Akad. Nauk SSSR **198**, 537–539.

Nevai, P [1979] *Orthogonal Polynomials*, Memoirs of Amer. Math. Soc. **213**, Amer. Math. Soc., Providence, RI.

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

Nelson, E. [1959] *Analytic vectors*, Ann. Math. **70**, 572–615.

Riesz, F. and B. Sz.-Nagy [1955] *Functional analysis*, Ungar, New York.

Radon, J. [1948] *Zur mechanischen Kubatur*, Monatsh. Math. **52**, 286–300.

Schmid, H. J. [1978] *On cubature formula with a minimal number of knots*, Numer. Math. **31**, 282–297.

Schmid, H. J. [1983] *Interpolatorische Kubaturformeln*, Diss. Math. **CCXX**, 1-122.

Schmid, H. J. and Y. Xu [1994] *On Bivariate Gaussian cubature formula*, Proc. Amer. Math. Soc. **122**, 833–842.

Stone, M. [1932] *Linear transformations in Hilbert space and their applications to analysis*, reprinted 1983, Amer. Math. Soc., Colloq. Publ. vol.15, Providence, RI.

Stroud, A. [1971] *Approximate calculation of multiple integrals*, Prentice Hall, Englewood Cliffs, NJ.

Suetin, P. K. [1988] *Orthogonal polynomials in two variables*, "Nauka", Moscow. (Russian)

Szegő, H. [1975] *Orthogonal polynomials*, 4th ed., Amer. Math. Soc. Colloq. Publ. vol.23, Providence, RI.

Verlinden, P. and R. Cools [1992] *On cubature formulae of degree $4k+1$ attaining Möller's lower bound for integrals with circular symmetry*, Numer. Math **61**, 395–407.

Xu, Y. [1992] *Gaussian cubature and bivariate polynomial interpolation*, Math. Comp. **59**, 547–555.

Xu, Y. [1993a] *On multivariate orthogonal polynomials*, SIAM J. Math. Anal. **24**, 783–794.

Xu, Y. [1993b] *Unbounded commuting operators and multivariate orthogonal polynomials*, Proc. Amer. Math. Soc. **119**, 1223-1231.

Xu, Y. [1994a] *Multivariate orthogonal polynomials and operator theory*, Trans. Amer. Math. Soc. **343**, 193-202.

Xu, Y. [1994b] *Block Jacobi matrices and zeros of multivariate orthogonal polynomials*, Trans. Amer. Math. Soc. **342**, 855–866.

Xu, Y. [1994c] *Recurrence formulas for multivariate orthogonal polynomials*, Math. Comp. **62**, 687–702.

Xu, Y. [1994d] *On zeros of multivariate quasi-orthogonal polynomials and Gaussian cubature formulae*, SIAM J. Math. Anal. **25**, 991–1001.

Xu, Y. [1994e] *Solutions of three-term relations in several variables*, Proc. Amer. Math. Soc. **122**, 151–155.

Xu, Y. [1994f] *A characterization of positive quadrature formulae*, Math. Comp. **62**, 703–718.

Xu, Y. [1994g] *On a class of bivariate orthogonal polynomials and cubature formula*, Numer. Math. **69**, 233–241.

Xu, Y. [1994h] *Common zeros of polynomials in several variables and higher dimensional quadrature*, Pitman Research Notes in Mathematics Series, Longman, Essex.

Xu, Y. [1995] *Christoffel functions and Fourier series for multivariate orthogonal polynomials*, J. Approx. Theory **82**, 205-239.

Xu, Y. [199xa] *Lagrange interpolation on Chebyshev points of two variables*, J. Approx. Theory (to appear).

Xu, Y. [199xb] *Summability of Fourier orthogonal series for Jacobi weight on a ball in $\mathbb{R}^d$* (to appear).