Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Advances in Applied Mathematics 33 (2004) 615–632

ADVANCES IN Applied Mathematics

www.elsevier.com/locate/yaama

On discrete orthogonal polynomials of several variables ☆

Yuan Xu

Department of Mathematics, University of Oregon, Eugene, OR 97403-1222, USA

Received 7 January 2004; accepted 1 March 2004

Available online 24 June 2004

Abstract

Let V be a set of isolated points in $\mathbb{R}^d$. Define a linear functional $\mathcal{L}$ on the space of real polynomials restricted on V, $\mathcal{L}f = \sum_{x\in V} f(x)\rho(x)$, where ρ is a nonzero function on V. Polynomial subspaces that contain discrete orthogonal polynomials with respect to the bilinear form $\langle f, g\rangle = \mathcal{L}(fg)$ are identified. One result shows that the discrete orthogonal polynomials still satisfy a three-term relation and Favard's theorem holds in this general setting.

Keywords: Discrete orthogonal polynomials; Several variables; Three-term relation; Favard's theorem

1. Introduction

Discrete orthogonal polynomials appear naturally in combinatorics, genetics, statistics and various areas in applied mathematics (see, for example, [4,6]). In one variable they have been studied extensively. Let V be a set of isolated points on the real line, its cardinality $|V|$ is either finite or countable. Let w be a real positive function on V. With respect to the bilinear form $\langle f, g\rangle = \sum_{x\in V} f(x)g(x)w(x)$, there is a sequence of orthogonal polynomials $\{p_n\colon 0 \leqslant n \leqslant |V|\}$ on V with $\langle p_n, p_m\rangle = 0$ for $n \neq m$ (for example, using Gram–Schmidt process). These are the discrete orthogonal polynomials.

☆ Work partially supported by the National Science Foundation under Grant DMS-0201669.
E-mail address: yuan@math.uoregon.edu.

0196-8858/$ – see front matter
doi:10.1016/j.aam.2004.03.002

Their structure is similar to that of the usual continuous orthogonal polynomials. For example, every sequence of discrete orthogonal polynomials satisfies a three-term relation,

$$xp_n = a_n p_{n+1} + b_n p_n + c_n p_{n-1}, \quad 0 \leqslant n \leqslant |V| - 1, \tag{1.1}$$

where a_n, b_n and c_n are real numbers. Furthermore, according to Favard's theorem, the three-term relation essentially characterizes the orthogonality of polynomials.

Discrete orthogonal polynomials of several variables are far less studied. Their orthogonal structure is much more complicated than that of one variable. Even some basic problems have not been addressed. Let us first fix some notation. We use the standard multiindex notation: for $x \in \mathbb{R}^d$ and $\alpha \in \mathbb{N}_0^d$, write $x^\alpha = x_1^{\alpha_1} \cdots x_d^{\alpha_d}$. This is a monomial of (total) degree $|\alpha| := \alpha_1 + \cdots + \alpha_d$. Let $\Pi^d = \mathbb{R}[x_1, \ldots, x_d]$ and let Π_n^d be the subspace of polynomials of degree at most n. Denote by $\mathcal{P}_n^d$ the space of homogeneous polynomials of degree n. It is well known that

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

Let $\mathcal{L}$ be a linear functional defined on Π^d, such that a basis of orthogonal polynomials $\{P_\alpha\colon |\alpha| = n, \alpha \in \mathbb{N}_0^d, n \geqslant 0\}$, where $P_\alpha \in \Pi_n^d$, exists with respect to the bilinear form $\langle f, g\rangle = \mathcal{L}(fg)$, in the sense that $\langle P_\alpha, P_\beta\rangle = 0$ if $|\alpha| \neq |\beta|$. Let $\mathcal{V}_n^d = \operatorname{span}\{P_\alpha\colon |\alpha| = n\}$ be the space of orthogonal polynomials of total degree n. Then $\dim \mathcal{V}_n^d = \dim \mathcal{P}_n^d$. In this case, if we adopt the point of view that the orthogonality holds in terms of the subspaces $\mathcal{V}_n^d$, not in terms of particular bases of $\mathcal{V}_n^d$, then we can have an analog of a three-term relation. Let $\mathbb{P}_n = \{P_\alpha\colon |\alpha| = n\}$; we also use $\mathbb{P}_n$ to denote a column vector, in which the elements are ordered according to a fixed monomial order. Then the following three-term relation holds,

$$x_i \mathbb{P}_n = A_{n,i}\mathbb{P}_{n+1} + B_{n,i}\mathbb{P}_n + C_{n,i}\mathbb{P}_{n-1}, \quad 1 \leqslant i \leqslant d,\ n \geqslant 0, \tag{1.2}$$

where $A_{n,i}$, $B_{n,i}$ and $C_{n,i}$ are matrices of appropriate dimensions, and $\mathbb{P}_{-1} := 0$. Furthermore, there is an analogue of Favard's theorem [11]. For the general theory of orthogonal polynomials of several variables, we refer to Chapter 3 of [2].

Let V be a set of isolated points in $\mathbb{R}^d$. Again we denote by $|V|$ the cardinality of V, which can be finite or countable. The orthogonal polynomials on V depend on the structure of the polynomial ideal $I(V)$ that has V as its variety,

$$I(V) = \bigl\{p \in \mathbb{R}[x_1, \ldots, x_d]\colon\ p(x) = 0,\ \forall x \in V\bigr\}.$$

The discrete orthogonal polynomials on V can only consist of polynomials that do not belong to $I(V)$. Let $\mathcal{L}$ be defined by $\mathcal{L}f = \sum_{x \in V} f(x)W(x)$, where W is a real function on V, $W(x) \neq 0$ for all $x \in V$, and $\sum_{x \in V} |x^\alpha||W(x)| < \infty$ for all $\alpha \in \mathbb{N}_0^d$ in the case where V is a countable set. Only if $I(V) = \{0\}$, can a complete basis $\{P_\alpha\colon \alpha \in \mathbb{N}_0^d\}$ of Π^d with respect to $\mathcal{L}$ exist and the discussion in the previous paragraph applies. If the ideal $I(V)$ is nontrivial, for example, when $|V|$ is finite, then we need to understand the subspace $\mathbb{R}[V] \cong \Pi^d / I(V)$ in order to define orthogonal polynomials. In the case

$\mathbb{R}[V] = \Pi_N^d$, for example, little extra work is needed; the three-term relation in the form (1.2) holds for $0 \leqslant n \leqslant N$. This is the case of straightforward extension of discrete orthogonal polynomials in one variable. However, even in the case where V is a product of two point sets X and Y in one variable, $V = X \times Y$ with $|V| = N$, $\mathbb{R}[V]$ consists of only a subspace of Π_N^d. In general, the space $\mathbb{R}[V]$ can be rather complicated and care is needed for the definition of orthogonal polynomials on a discrete set V.

The purpose of the present study is to define discrete orthogonal polynomials in this general setting. The polynomial subspaces for which the discrete orthogonal polynomials exist are identified and the three-term relation and Favard's theorem are established. In the following section we discuss the structure of polynomial subspaces on V. Discrete orthogonal polynomials on V are studied in Section 3. Various examples will be given in the paper, some are given in terms of the classical discrete orthogonal polynomials, such as Hahn polynomials, in Section 4. It is our hope that this study can help to clarify some of the basic questions in the theory of discrete orthogonal polynomials.

2. Polynomial spaces on *V*

First we need to understand the structure of the quotient ideal Π^d/I, where $I := I(V)$ and V is a set of isolated points in $\mathbb{R}^d$, finite or countable. Most of the results below also hold if V has finitely many accumulation points. We review some results about ideals and varieties, our basic reference is [1].

For $f, g \in \Pi^d$, we say that f is congruent to g modulo I, written as $f = g \bmod I$, if and only if $f - g \in I$. If $|V|$ is finite, then it is known that the codimension of I is equal to $|V|$; that is, $\dim \Pi^d/I(V) = |V|$. Let $\mathbb{R}[V]$ denote the collection of polynomial functions $\phi : V \mapsto \mathbb{R}$. This is a commutative ring and it is isomorphic to the quotient ring $\Pi^d/I(V)$. We can identify $\mathbb{R}[V]$ with Π^d/I as there is an one-to-one correspondence between $\phi \in \mathbb{R}[V]$ and $[\phi] = \{g \in \Pi^d\colon\ g = \phi \bmod I\}$. It is possible to say more about this space. For a fixed monomial order, we denote by $\mathsf{LT}(f)$ the leading monomial term for any polynomial $f \in \Pi^d$; that is, if $f = \sum c_\alpha x^\alpha$, then $\mathsf{LT}(f) = c_\beta x^\beta$, where x^β is the leading monomial among all monomials appearing in f and $c_\beta \neq 0$. For a polynomial ideal I other than [1], we denote by $\mathsf{LT}(I)$ the leading terms of I, that is,

$$\mathsf{LT}(I) = \bigl\{c\mathbf{x}^\alpha \mid \text{ there exists } f \in I \text{ with } \mathsf{LT}(f) = c\mathbf{x}^\alpha\bigr\}.$$

We further denote by $\langle \mathsf{LT}(I)\rangle$ the ideal generated by the leading terms of $\mathsf{LT}(f)$ for all $f \in I \setminus \{0\}$. According to the Hilbert basis theorem, every polynomial ideal has a finite basis. A set $\{g_1, \ldots, g_t\}$ is called a Gröbner basis of I if

$$\bigl\langle \mathsf{LT}(g_1), \ldots, \mathsf{LT}(g_t)\bigr\rangle = \bigl\langle \mathsf{LT}(I)\bigr\rangle.$$

It is known that every polynomial ideal has a Gröbner basis. Such a basis enjoys many interesting properties that have important applications. For example, it is used to prove the following result [1, Chapter 5].

Proposition 2.1. *Fix a monomial ordering on Π^d and let $I \subset \Pi^d$ be an ideal. Then there is an isomorphism between Π^d/I and the space*

$$\mathcal{S}_I := \operatorname{span}\bigl\{\mathbf{x}^\alpha \mid \mathbf{x}^\alpha \notin \bigl\langle \mathsf{LT}(I)\bigr\rangle\bigr\}.$$

More precisely, every $f \in \Pi^d$ is congruent modulo I to a unique polynomial $r \in \mathcal{S}_I$.

In fact, the polynomial r is the remainder of f on division by I. For $I = \langle f_1, \ldots, f_M\rangle$ and a fixed monomial order, the division algorithm states that for every $f \in \Pi^d$, there exist p_i and r in Π^d such that $f = \sum p_i f_i + r$, where $r \in \mathcal{S}_I$ and no term of r is divisible by any of $\mathsf{LT}(f_1), \ldots, \mathsf{LT}(f_M)$. The remainder polynomial r is unique if the basis $f_1, \ldots, f_M$ is a Gröbner basis.

Proposition 2.2. *Let V be a set of isolated points in $\mathbb{R}^d$ and $I = I(V)$. Let $\Lambda := \Lambda(V)$ be the index set $\Lambda = \{\alpha\colon\ x^\alpha \notin \mathsf{LT}(I)\}$. Then every polynomial $P \in \mathbb{R}[V]$ can be written uniquely as*

$$P(x) = \sum_{\alpha \in \Lambda} c_\alpha x^\alpha \quad \bmod I(V), \quad c_\alpha \in \mathbb{R},$$

and the set Λ satisfies the following property

$$\alpha \in \Lambda \quad \textit{implies} \quad \alpha - \beta \in \Lambda, \quad \textit{whenever} \quad \alpha - \beta \in \mathbb{N}_0^d \quad \textit{and} \quad \beta \in \mathbb{N}_0^d. \tag{2.1}$$

Proof. For $I = I(V)$, we can take $\mathcal{S}_I$ in Proposition 2.1 as $\mathbb{R}[V]$, modulus I if needed. The definition shows that we can write $\mathcal{S}_I = \operatorname{span}\{x^\alpha\colon\ \alpha \in \Lambda\}$. Hence, every polynomial P in $\mathbb{R}[V]$ has the stated representation. Since the ideal $\langle \mathsf{LT}(I)\rangle$ is a monomial ideal, $x^\alpha \in \langle \mathsf{LT}(I)\rangle$ implies $x^{\alpha+\beta} \in \langle \mathsf{LT}(I)\rangle$ for any $\beta \in \mathbb{N}_0^d$. Consequently, it follows that the set Λ satisfies the property (2.1). □

In the following, we shall drop modulus I and use $\mathbb{R}[V]$ to denote the space

$$\mathbb{R}[V] = \operatorname{span}\bigl\{x^\alpha\colon\ \alpha \in \Lambda(V)\bigr\}. \tag{2.2}$$

This abuse of notation should not cause problems.

We should point out that the set Λ is not unique, since all equations actually hold under congruence modulo I. In fact, Gröbner bases are not unique, since the choice of monomial orders matters. There are in fact many different representations of elements in $\mathbb{R}[V]$. What is of interest is the property (2.1) satisfied by Λ.

Example 2.1. Consider the set $V = \{(0,0), (0,1), (1,2), (2,3)\}$. It is easy to see that $I(V) = \langle g_1, g_2, g_3\rangle$, where

$$g_1(x,y) = x(x-1)(x-2), \qquad g_2(x,y) = x(x+1-y), \qquad g_3(x,y) = y(x+1-y).$$

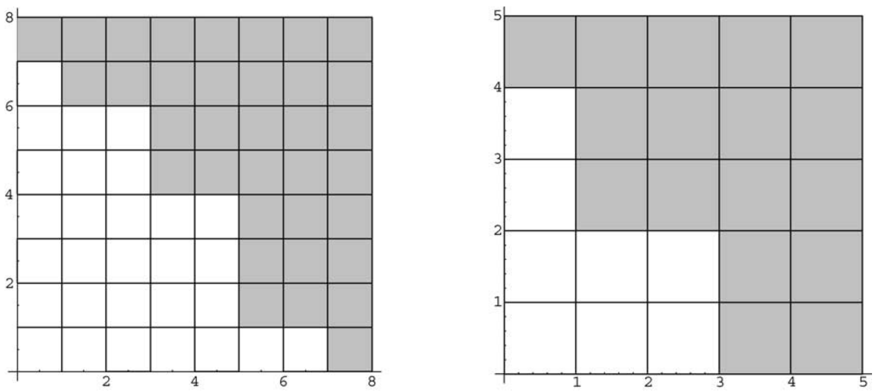

Fig. 1.

If we use the graded reverse lexicographical order $(n-k-1, k+1) \succ (n-k, k)$, then $\langle \mathsf{LT}(I) \rangle = \langle y^2, xy, x^3 \rangle$ and $\mathbb{R}[V] = \operatorname{span}\{1, x, y, x^2\}$. If we use the graded lexicographical order $(n-k, k) \succ (n-k-1, k+1)$, then $\langle \mathsf{LT}(I) \rangle = \langle x^2, xy, y^3 \rangle$ since we also have $I(V) = \langle g_2, g_3, g_4 \rangle$ where $g_4(x, y) = y^3 - 6y^2 + 5y + 6x$, and $\mathbb{R}[V] = \operatorname{span}\{1, x, y, y^2\}$.

For $d = 2$, the property (2.1) of Λ shows that the set Λ must be of a stair shape as the lattice points in the unshaded area depicted in Fig. 1.

That is, in the case of $d = 2$, for each set Λ there is a sequence of positive integers n_i, which satisfies $n_m \leqslant n_{m-1} \leqslant \cdots \leqslant n_0$ (some of the n_i can be positive infinity), such that

$$\Lambda = \{(k, l):\ 0 \leqslant l \leqslant m,\ 0 \leqslant k \leqslant n_l\}. \tag{2.3}$$

Example 2.2. Let Λ be the lattice set in the two figures in Fig. 1. For the left one, $m = 6$ and $\Lambda = \{(i, j):\ 0 \leqslant i \leqslant n_j, 0 \leqslant j \leqslant 6\}$ with $(n_0, \ldots, n_6) = (6, 4, 4, 4, 2, 2, 0)$. For the right one, $m = 3$ and

$$\Lambda = \{(i, 0):\ 0 \leqslant i \leqslant 2\} \cup \{(i, 1):\ 0 \leqslant i \leqslant 2\} \cup \{(0, 2)\} \cup \{(0, 3)\}$$

with $(n_0, n_1, n_2, n_3) = (2, 2, 0, 0)$.

Proposition 2.3. *There exists a point set V for which $\mathbb{R}[V]$ in* (2.2) *is given by the index set Λ in* (2.3)*.*

Proof. Let $x_0, x_1, \ldots, x_{n_0}$ and $y_0, y_1, \ldots, y_m$ be isolated real numbers. We define the point set V as follows:

$$\begin{aligned} V = \{&(x_0, y_0),\ \ (x_1, y_0),\ \ \ldots,\ \ (x_{n_0}, y_0),\\ &(x_0, y_1),\ \ (x_1, y_1),\ \ \ldots,\ \ (x_{n_1}, y_1)\\ &\ \ \vdots \qquad\qquad \vdots \qquad\quad \ldots \qquad \vdots\\ &(x_0, y_m),\ (x_1, y_m),\ \ldots,\ (x_{n_m}, y_m)\}. \end{aligned}$$

If n_0 is finite, then $|V| = n_0 + n_1 + \cdots + n_m + m + 1$. We let $n_{m+1} = -1$ and adopt the convention that $\prod_{i=0}^{-1} a_i = 1$. Define polynomials

$$g_k(x) = \prod_{i=0}^{n_k}(x - x_i)\prod_{j=0}^{k-1}(y - y_j), \quad 0 \leqslant k \leqslant m + 1,$$

where if $n_k = \infty$, then we take $g_k(x) = 1$. Then it is easy to see that $\langle I \rangle = \langle g_0, g_1, \ldots, g_{m+1} \rangle$, so that

$$\langle \mathsf{LT}(I) \rangle = \langle x^{n_0+1}, x^{n_1+1}y, \ldots, x^{n_m+1}y^m, y^{m+1} \rangle,$$

from which it follows that $\mathbb{R}[V] = \{x^k y^l\colon (k,l) \in \Lambda\}$ with Λ given in (2.3). $\square$

Example 2.3. Consider the "triangle" point set

$$\begin{aligned} V = \bigl\{(x_0, y_0),\ (x_0, y_1),\ \ldots,\ &(x_0, y_m), \\ (x_1, y_1),\ \ldots,\ &(x_1, y_m), \\ \ddots\quad &\ \vdots \\ &(x_m, y_m)\bigr\}, \end{aligned}$$

where all points are isolated and $|V| = (m+1)(m+2)/2$. In this case, $n_l = m - l$ for $0 \leqslant l \leqslant m$ (making a transpose of the array as in matrix transpose), so that $\Lambda = \{(k,l)\colon 0 \leqslant k + l \leqslant m\}$ and $\mathbb{R}[V] = \Pi_m^2$.

Example 2.4. Consider the "product" point set

$$V = X \times Y = \bigl\{(x_i, y_j)\colon 0 \leqslant i \leqslant n,\ 0 \leqslant j \leqslant m\bigr\},$$

for which $|V| = (n+1)(m+1)$. In this case $n_0 = n_1 = \cdots = n_m = n$ and $\mathbb{R}[V] = \{x^k y^l\colon 0 \leqslant k \leqslant n, 0 \leqslant l \leqslant m\} = \Pi_n^1 \times \Pi_m^1$.

The above discussion can be extended to $\mathbb{R}[x_1, \ldots, x_d]$. We stick to the case $d = 2$ to keep the notation simple.

A couple of remarks are in order. First, if $m = \infty$ in Example 2.2, then $\mathbb{R}[V] = \mathbb{R}[x, y]$. This is also the case for Example 2.3 when both n and m are infinity. If, however, n is infinite and m is finite (or other way round), then $\Lambda = \{(k,l)\colon 0 \leqslant l \leqslant m, k \geqslant 0\}$ is an infinite set but not all $\mathbb{N}_0^2$ and $\mathbb{R}[V] = \Pi^d \times \Pi_m$. Furthermore, the space of polynomials of one variable appears as a special case.

Example 2.5. If V is a point set on the x coordinate line, $V = \{(x_0, 0), \ldots, (x_n, 0)\}$ (n can be infinity), then $\mathbb{R}[V] = \Pi_n^1$.

Second, we should point out that the point set V in the proposition, and the above examples, are the simplest examples for which $\mathbb{R}[V]$ can be determined. For a generic point set V, the problem of determining $\mathbb{R}[V]$ is highly nontrivial. One possible algorithm, at least for V is finite and moderate in size, is to check the rank of the matrices whose rows are the vectors $X_\alpha := \{x^\alpha\colon x \in V\}$.

The algorithm goes as follows: fix a graded monomial order, starting with $\alpha \in \mathbb{N}_m := \{\beta\colon |\beta| \leqslant m\}$ for a small m, so that the resulting matrix $(X_\alpha)_{\alpha\in\mathbb{N}_m}$ has full rank. Then add new rows X_α according to the order in $|\alpha| = m+1$ to the matrix. For each new row added, check the rank of the new matrix; if it has full rank, add the next X_α and proceed; if it does not have full rank, remove this row and add the next X_α and continue. When the matrix becomes a square nonsingular matrix, the corresponding set of x^α will be a basis for $\mathbb{R}[V]$.

3. Discrete orthogonal polynomials

Let V be a set of isolated points in $\mathbb{R}^d$. Let W be a real function on V and $W(x) \neq 0$ for any $x \in V$. Assume that

$$\sum_{x\in V} \left|x^\alpha\right|\left|W(x)\right| < \infty \quad \text{for all } \alpha \in \mathbb{N}_0^d$$

in the case where V is an infinite set. We define a bilinear form $\langle\cdot,\cdot\rangle$ on $\Pi^d \times \Pi^d$ by

$$\langle f, g\rangle = \mathcal{L}(fg), \quad \text{where } \mathcal{L}(f) := \sum_{x\in V} f(x)W(x).$$

If $\langle f, g\rangle = 0$, we say that f and g are orthogonal to each other with respect to W on the discrete set V. The notation $\mathcal{L}(fg)$ is more convenient for the matrix operations below.

Fix a graded monomial order. Let $\mathbb{R}[V]$ and $\Lambda = \Lambda(V)$ be defined as in the previous section (see (2.2)). Let $\mathbf{n} = \max\{|\alpha|\colon \alpha \in \Lambda(V)\}$ and $\Lambda_k(V) = \{\alpha \in \Lambda(V)\colon |\alpha| = k\}$, $0 \leqslant k \leqslant \mathbf{n}$. Note that $\mathbf{n}$ can be infinity. Define $r_k = |\Lambda_k(V)|$. We denote by $\mathbf{x}^\Lambda$, $\mathbf{x}_k^\Lambda$ and $\mathbf{x}^k$, $0 \leqslant k \leqslant \mathbf{n}$, the sets

$$\mathbf{x}^\Lambda = \left\{x^\alpha\colon \alpha \in \Lambda(V)\right\}, \qquad \mathbf{x}_k^\Lambda = \left\{x^\alpha \in \mathbf{x}^\Lambda\colon |\alpha| \leqslant k\right\}, \qquad \mathbf{x}^k = \left\{x^\alpha\colon \alpha \in \Lambda_k(V)\right\},$$

respectively. We also regard them as column vectors in which the elements are ordered according to the fixed graded monomial order.

The set of orthogonal polynomials on V will be denoted by $\{P_\alpha\colon \alpha \in \Lambda(V)\}$, where P_α has degree $|\alpha|$. We introduce the following notion. If $\{P_\alpha\colon \alpha \in \Lambda(V)\}$ is a sequence of polynomials in $\mathbb{R}[V]$, then set $\mathbb{P}_k := \{P_\alpha\colon \alpha \in \Lambda_k(V)\}$. Just as in the case of $\mathbf{x}^k$, we also regard $\mathbb{P}_k$ as a column vector.

Definition 3.1. Let V be the set of isolated points and W be a nonzero real function on V as above. A sequence of polynomials $\{P_\alpha \in \Pi^d_{|\alpha|}\colon \alpha \in \Lambda(V)\}$ is orthogonal with respect to W if

$$\mathcal{L}\big(\mathbf{x}^l \mathbb{P}_k^T\big) = \big\langle \mathbf{x}^l, \mathbb{P}_k \big\rangle = 0, \quad k > l, \qquad \text{and} \qquad \mathcal{L}\big(\mathbf{x}^k \mathbb{P}_k^T\big) = \big\langle \mathbf{x}^k, \mathbb{P}_k \big\rangle = S_k,$$

for $0 \leqslant k \leqslant \mathbf{n}$, where S_k is an invertible matrix of size $r_k \times r_k$. The sequence is orthonormal with respect to W on V if $\langle \mathbb{P}_k, \mathbb{P}_k \rangle = I_{r_k}$, the identity matrix, for $0 \leqslant k \leqslant \mathbf{n}$.

The notation $\mathcal{L}(\mathbf{x}^l \mathbb{P}_k^T)$ is more convenient than $\langle \mathbf{x}^k, \mathbb{P}_k \rangle$, since it shows clearly that this is a matrix of size $r_k \times r_k$. The orthogonality of P_α is defined as orthogonal to lower degree polynomials, as in the continuous case. The polynomials of the same degree may not be pairwise orthogonal. By definition, we can write

$$\mathbb{P}_k = G_k \mathbf{x}^k + G_{k-1,k} \mathbf{x}^{k-1} + \cdots + G_{1,k} \mathbf{x}^0,$$

where G_k is a $r_k \times r_k$ matrix, called the leading coefficient of $\mathbb{P}_k$. Assume a sequence of orthogonal polynomials P_α exists on V. Then we can follow the proof in [2, Section 3.1] to show that $\{\mathbb{P}_0, \ldots, \mathbb{P}_k\}$ is a basis for $\mathbb{R}[V] \cap \Pi^d_k$, $H_k := \mathcal{L}(\mathbb{P}_k \mathbb{P}_k^T)$ and G_k are both invertible. Furthermore, the following theorem still holds.

Theorem 3.2. *A sequence of orthogonal polynomials $\{P_\alpha\colon \alpha \in \Lambda(V)\}$ with respect to W on V exists if and only if the matrices $M_k := \langle \mathbf{x}_k^\Lambda, \mathbf{x}_k^\Lambda \rangle$ are nonsingular for $0 \leqslant k \leqslant \mathbf{n}$.*

The proof is based on linear algebra and follows exactly as in the continuous case; see [2, Theorem 3.1.6].

If W is positive on V, $W(x) > 0$ for all $x \in V$, then $\langle f, f \rangle = 0$ implies $f \equiv 0$ on V, so that the bilinear form $\langle \cdot, \cdot \rangle = 0$ becomes an inner product on $\mathbb{R}[V]$. For such a W, orthogonal polynomials on V exist. Furthermore, in this case, we can have orthonormal bases.

Theorem 3.3. *If W is a positive function on V, then a sequence of orthonormal polynomials $\{P_\alpha\colon \alpha \in \Lambda(V)\}$ with respect to W on V exists.*

Proof. In this case, the matrix M_k is positive definite since for any nonzero column vector $\mathbf{c}$, $\mathbf{c}^T M_k \mathbf{c} = \langle \mathbf{c}^T \mathbf{x}^k, \mathbf{c}^T \mathbf{x}^k \rangle > 0$. In particular, $M_\Lambda := \langle \mathbf{x}^\Lambda, \mathbf{x}^\Lambda \rangle$ is a symmetric and positive definite matrix. It follows that it can be factored as $M_\Lambda = SDS^T$ where S is a nonsingular lower triangular matrix and $D = \mathrm{diag}\{d_1, d_2, \ldots, d_{|V|}\}$ with all $d_i > 0$. Let $D^{-1/2} = \mathrm{diag}\{d_1^{-1/2}, \ldots, d_{|V|}^{-1/2}\}$ and $R = D^{-1/2} S^{-1}$. Then $R M_\Lambda R^T = \langle R\mathbf{x}^\Lambda, R\mathbf{x}^\Lambda \rangle = I_{|V|}$. Since S is lower triangular and $\mathbf{x}^\Lambda = \{\mathbf{x}^0, \mathbf{x}^1, \ldots, \mathbf{x}^{\mathbf{n}}\}$ as a column vector, we can write the components of $R\mathbf{x}^\Lambda$ as $\mathbb{P}_0, \ldots, \mathbb{P}_{\mathbf{n}}$ where $\mathbb{P}_k$ consists of polynomials of degree k. These are the orthonormal polynomials. $\square$

The proof of the theorem provides an algorithm that can be used to construct discrete polynomials in several variables. For $|V|$ is finite and moderate in size, it is rather effective; an example is given in the following section (Example 4.3).

Let $\mathcal{V}_k(W)$ denote the space of orthogonal polynomials of degree k; that is, $\mathcal{V}_k(W) = \operatorname{span} \mathbb{P}_k$. Evidently $\dim \mathcal{V}_k(W) = r_k$. Comparing to the orthogonal polynomials in the continuous case, the numbers r_k depend on V and there is no closed formula for them. Furthermore, $r_k \leqslant \dim \mathcal{P}_k^d$ and the equality often does not hold. In fact, r_k may no longer be an increasing sequence.

Example 3.1. Let $V = X \times Y$ be the product point set in Example 2.3, where $X = \{x_0, \dots, x_n\}$ and $Y = \{y_0, \dots, y_m\}$. Here $\mathbf{n} = n + m$. Let $W(x, y) = w_1(x)w_2(y)$, where w_1 is positive on X and w_2 is positive on Y. Then the orthogonal polynomials $\mathbb{P}_k$, $0 \leqslant k \leqslant \mathbf{n}$, exist and can be constructed as follows: let $\{p_0, p_1, \dots, p_n\}$ and $\{q_0, q_1, \dots, q_n\}$ denote the discrete orthogonal polynomials with respect to w_1 on X and w_2 on Y, respectively. Then the orthogonal polynomials on V are given by $P_{k,l}(x, y) = p_k(x)q_l(y)$. We assume m is finite and $n > m$. Then in our vector notation,

$$\mathbb{P}_k(x, y) = \bigl\{p_0(x)q_k(y), p_1(x)q_{k-1}(y), \dots, p_k(x)q_0(y)\bigr\} \quad \text{for } 0 \leqslant k \leqslant m,$$

so that $r_k = k + 1$ for $0 \leqslant k \leqslant m$,

$$\mathbb{P}_k(x, y) = \bigl\{p_{k-m}(x)q_m(y), \dots, p_k(x)q_0(y)\bigr\} \quad \text{for } m + 1 \leqslant k \leqslant n,$$

so that $r_k = m + 1$ for $m + 1 \leqslant k \leqslant n$, and

$$\mathbb{P}_k(x, y) = \bigl\{p_{k-m}(x)q_m(y), \dots, p_n(x)q_{k-n}(y)\bigr\} \quad \text{for } n + 1 \leqslant k \leqslant n + m,$$

so that $r_k = n + m - k + 1$ for $n + 1 \leqslant k \leqslant n + m - 1$.

Next we consider the three-term relations satisfied by the orthogonal polynomials. If P_α is an orthogonal polynomial, then $P_\alpha \in \mathbb{R}[V]$ so that it is a linear combination of x^α for $\alpha \in \Lambda(V)$. Clearly, multiplying by a coordinate x_i gives a polynomial $x_i P_\alpha$ of degree $|\alpha| + 1$. However, unlike the continuous case, $x_i x^\alpha$ may not belong to $\mathbb{R}[V]$ for some $\alpha \in \Lambda(V)$. Nevertheless, it is congruent modulus $I(V)$ to a unique polynomial in $\mathbb{R}[V]$. Recall that $\mathbf{n} = \max\{|\alpha|: \alpha \in \Lambda(V)\}$.

Proposition 3.4. *Let $I := I(V)$. For $0 \leqslant k \leqslant \mathbf{n} - 1$, there exist matrices $A_{k,i}$: $r_k \times r_{k+1}$, $B_{k,i}$: $r_k \times r_k$, and $C_{k,i}$: $r_k \times r_{k-1}$, such that for $1 \leqslant i \leqslant d$,*

$$x_i \mathbb{P}_k(x) = A_{k,i} \mathbb{P}_{k+1}(x) + B_{k,i} \mathbb{P}_k(x) + C_{k,i} \mathbb{P}_{k-1}(x) \mod I, \tag{3.1}$$

where $0 \leqslant k \leqslant \mathbf{n} - 1$ and we define $\mathbb{P}_{-1} = 0$, $A_{k,i} = 0$ and $C_{-1,i} = 0$; moreover,

$$A_{k,i} H_{k+1} = \mathcal{L}\bigl(x_i \mathbb{P}_k \mathbb{P}_{k+1}^T\bigr) = H_k C_{k+1,i}^T, \qquad B_{k,i} H_k = \mathcal{L}\bigl(x_i \mathbb{P}_k \mathbb{P}_k^T\bigr). \tag{3.2}$$

Proof. If all components of $x_i \mathbb{P}_k$ are in $\mathbb{R}[V]$, this is proved as in the usual case, by writing $x_i \mathbb{P}_k$ in terms of orthogonal polynomials $\mathbb{P}_0, \mathbb{P}_1, \dots, \mathbb{P}_{k+1}$ and then use orthogonality. If $P_\alpha \in \mathbb{R}[V]$ but $x_i P_\alpha \notin \mathbb{R}[V]$ for some $|\alpha| = k$, then there exist a Q in $I(V)$ and an $R_\alpha \in \mathbb{R}[V]$ such that $x_i P_\alpha(x) = Q(x) + R_\alpha(x)$, and the degree of R_α is at most $k + 1$. We can write $R_\alpha = \mathbf{a}_{k+1}\mathbb{P}_{k+1} + \mathbf{a}_k\mathbb{P}_k + \mathbf{a}_{k-1}\mathbb{P}_{k-1} + \cdots$, where $\mathbf{a}_j$ are some row vectors of appropriate size. By the orthogonality, we get

$$\mathbf{a}_{k+1}\mathcal{L}\big(\mathbb{P}_{k+1}\mathbb{P}_{k+1}^T\big) = \mathcal{L}\big(R_\alpha\mathbb{P}_{k+1}^T\big) = \mathcal{L}\big(P_\alpha\mathbb{P}_{k+1}^T\big),$$

since Q vanishes on V and $\langle Q, P\rangle = 0$ for any P. Similarly, we get $\mathbf{a}_k H_k = \mathcal{L}(P_\alpha\mathbb{P}_k^T)$, $\mathbf{a}_{k-1}H_{k-1} = \mathcal{L}(P_\alpha\mathbb{P}_{k-1}^T)$, and all other $\mathbf{a}_j$ are equal to zero. In vector and matrix notation, this is the three-term relation. The presence of Q means that the equality holds under modulus $I(V)$ in general. □

Corollary 3.5. *Let $I := I(V)$. If $\{\mathbb{P}_k\}$ are orthonormal polynomials, then for $1 \leqslant i \leqslant d$,*

$$x_i\mathbb{P}_k(x) = A_{k,i}\mathbb{P}_{k+1}(x) + B_{k,i}\mathbb{P}_k(x) + A_{k-1,i}^T\mathbb{P}_{k-1}(x) \quad \mod I, \tag{3.3}$$

where $0 \leqslant k \leqslant \mathbf{n} - 1$, $\mathbb{P}_{-1} = 0$, $A_{k,i} = 0$ and $C_{-1,i} = 0$; moreover, $B_{n,i}$ are symmetric.

In the case of continuous orthogonal polynomials, the matrix $A_{n,i}$ has more columns than rows and it has full rank. This is no longer true in the discrete case. Since r_k may no longer be an increasing sequence, the matrix $A_{n,i}$ can have more rows than columns; moreover, it may not have full rank.

Example 3.1*. We continue the example $V = X \times Y$ in Example 3.1. Assume that $\{p_0, p_1, \ldots, p_n\}$ satisfies the three-term relation

$$xp_k(x) = a_n p_{k+1}(x) + b_k p_k(x) + c_k p_{k-1}(x)$$

and $\{q_0, q_1, \ldots, q_m\}$ satisfies the three-term relation

$$yq_k(y) = a_n' q_{k+1}(y) + b_k' q_k(y) + c_k' q_{k-1}(y),$$

respectively. With $x = x_1$ and $y = x_2$, the matrices $A_{k,1}$ and $A_{k,2}$ take the form

$$A_{k,1} = \begin{bmatrix} 0 & a_0 & & \bigcirc \\ \vdots & & \ddots & \\ 0 & \bigcirc & & a_k \end{bmatrix} \quad \text{and} \quad A_{k,2} = \begin{bmatrix} a_k' & & \bigcirc & 0 \\ & \ddots & & \vdots \\ \bigcirc & & a_0' & 0 \end{bmatrix}, \quad 0 \leqslant k < m,$$

of size $(k+1) \times (k+2)$;

$$A_{k,1} = \begin{bmatrix} a_{k-m} & & \bigcirc \\ & \ddots & \\ \bigcirc & & a_k \end{bmatrix} \quad \text{and} \quad A_{k,2} = \begin{bmatrix} 0 & 0 & \ldots & 0 \\ a_{m-1}' & 0 & & \bigcirc \\ & \ddots & \ddots & \\ \bigcirc & \ldots & a_0' & 0 \end{bmatrix}, \quad m \leqslant k < n,$$

of size $(m+1) \times (m+1)$; and

$$A_{k,1} = \begin{bmatrix} a_{k-m} & & \bigcirc \\ & \ddots & \\ \bigcirc & & a_k \\ 0 & \dots & 0 \end{bmatrix} \quad \text{and} \quad A_{k,2} = \begin{bmatrix} 0 & \dots & 0 \\ a'_{m-1} & & \bigcirc \\ & \ddots & \\ \bigcirc & & a'_{k-n} \end{bmatrix}, \quad n \leqslant k < n+m,$$

of size $(n+m-k+1) \times (n+m-k)$. Note that $A_{k,2}$ in the case $m \leqslant k < n$ does not have full rank. For $n \leqslant k < n+m$, $A_{k,i}$ has more rows than columns.

One consequence of the three-term relation is the Christoffel–Darboux formula,

$$\sum_{j=0}^{k} \mathbb{P}_k^T(x)\mathbb{P}_k(y) = \frac{\mathbb{P}_{k+1}(x)A_{k,i}^T\mathbb{P}_k(y) - \mathbb{P}_k(x)A_{k,i}\mathbb{P}_{k+1}(y)}{x_i - y_i}, \quad \text{mod } I(V),$$

where $1 \leqslant i \leqslant d$ and $0 \leqslant k < \mathbf{n}$. The proof follows as in the continuous case.

The composite matrix $A_k = (A_{k,1}^T, \dots, A_{k,d}^T)^T$ plays an important role in Favard's theorem of several variables. This matrix is of size $dr_k \times r_{k+1}$.

Proposition 3.6. *For $0 \leqslant k \leqslant \mathbf{n}-1$, $dr_k \geqslant r_{k+1}$; the composite matrix A_k of $A_{k,1}, \dots, A_{k,d}$ and the composite matrix C_{k+1} of $C_{k+1,1}, \dots, C_{k+1,d}$ both have full rank,*

$$\operatorname{rank} A_k = \operatorname{rank} C_{k+1}^T = r_{k+1}. \tag{3.4}$$

Proof. Recall that G_k denotes the leading coefficient of $\mathbb{P}_k$ and it is an invertible matrix. Let $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ be the ith element of the standard Euclidean basis. By (2.1), $\alpha \in \Lambda(V)$ implies that $\alpha - e_i \in \Lambda(V)$. If $\alpha \in \Lambda_k(V)$ but $\alpha + e_i \notin \Lambda_{k+1}(V)$, then $x_i x^\alpha = x^{\alpha+e_i} \in \langle \mathsf{LT}(I) \rangle$. We then define the matrix $L_{k,i}$ by

$$x_i \mathbf{x}^k = L_{k,i} \mathbf{x}^{k+1} \quad \text{mod } \langle \mathsf{LT}(I) \rangle, \quad 1 \leqslant i \leqslant d.$$

The matrix $L_{k,i}$ is of the size $r_k \times r_{k+1}$ and it is uniquely determined. Comparing the coefficients of $\mathbf{x}^{k+1}$ in both sides of the three-term relation, we see that

$$G_k L_{k,i} = A_{k,i} G_{k+1}, \quad 1 \leqslant i \leqslant d. \tag{3.5}$$

The entries of the matrix $L_{k,i}$ are mostly 0 with at most one 1 in each row. However, $L_{k,i}$ may not have full rank.

Since $\alpha \in \Lambda_{k+1}(V)$ implies that $\alpha - e_i \in \Lambda_k$ whenever $\alpha_i - 1 \geqslant 0$, it follows that $dr_k \geqslant r_{k+1}$ since $r_k = |\Lambda_k(V)|$. Moreover, since the column vector $(x_1\mathbf{x}^k, \dots, x_d\mathbf{x}^k)$ is equal to $L_k\mathbf{x}^{k+1}$, where L_k is the composite matrix of $L_{k,1}, \dots, L_{k,d}$, and clearly $\{x_i\mathbf{x}^k\colon 1 \leqslant i \leqslant d\}$ includes $\mathbf{x}^{k+1}$ as a subset, it follows that L_k has full rank r_{k+1}.

Equation (3.5) implies that $A_k G_{k+1} = \operatorname{diag}\{G_k, \dots, G_k\} L_k$. Since G_k invertible implies $\operatorname{diag}\{G_k, \dots, G_k\}$ invertible, it follows that $\operatorname{rank} A_{k+1} = r_{k+1}$. Furthermore, (3.2) implies that $A_k H_{k+1} = \mathrm{diag}\{H_k, \ldots, H_k\} C_{k+1}^T$ and H_k is invertible; hence, $\mathrm{rank}\, C_{k+1} = \mathrm{rank}\, A_k$. $\square$

Since the matrix A_k has full rank and $dr_k \geqslant r_{k+1}$, it has a generalized inverse, D_k^T, which is of the size $r_{k+1} \times dr_k$ and can be assumed to be of the form $D_k^T = (D_{k,1}^T, \ldots, D_{k,d}^T)$, where $D_{k,i}^T$ are of the size $r_{k+1} \times r_k$. Then

$$D_k^T A_k = D_{k,1}^T A_{k,1} + \cdots + D_{k,d}^T A_{k,d} = I_{r_{k+1}}.$$

We note that the generalized inverse is in general not unique. Using D_n^T, we get from the three-term relation a recursive formula

$$\mathbb{P}_{k+1} = \sum_{i=1}^{d} x_i A_{k,i} \mathbb{P}_k - \sum_{i=1}^{d} B_{k,i} \mathbb{P}_k - \sum_{i=1}^{d} C_{k,i} \mathbb{P}_{k-1}, \tag{3.6}$$

which allows us to compute $\mathbb{P}_{k+1}$ using $\mathbb{P}_k$ and $\mathbb{P}_{k-1}$. This formula is useful in the proof of the analog of Favard's theorem.

According to Propositions 3.4 and 3.6, orthogonal polynomials on a set V of isolated points satisfy a three-term relation whose coefficient satisfies a rank condition. We want to establish that the converse is also true, that is, an analog of Favard's theorem. To this end, we start with a sequence of polynomials that satisfies the three-term relation (3.1) and the rank condition (3.4), and show that there exist a set V of isolated points and a weight function W on V with respect to which the polynomials are orthogonal.

For this purpose let us start with an ideal $I \subset \Pi^d$ and let $\Lambda := \{\alpha \in \mathbb{N}_0^d\colon x^\alpha \notin \langle \mathsf{LT}(I)\rangle\}$. The proof of Proposition 2.2 shows that Λ satisfies the property (2.1). Assume that there is a sequence of polynomials $P_\alpha \in \mathcal{S}_I := \mathrm{span}\{x^\alpha\colon x^\alpha \notin \langle \mathsf{LT}(I)\rangle\}$, where P_α is indexed by $\alpha \in \Lambda$ such that $P_0 = 1$ and $P_\alpha \in \Pi_{|\alpha|}^d$. Set $\mathbf{n} = \max\{|\alpha|\colon \alpha \in \Lambda\}$, which can be infinity. Let $\Lambda_k = \{\alpha \in \Lambda\colon |\alpha| = k\}$ and let $\mathbb{P}_k = \{P_\alpha\colon \alpha \in \Lambda_k\}$ for $0 \leqslant k \leqslant \mathbf{n}$, and regard $\mathbb{P}_k$ as column vectors according to a fixed graded monomial order.

Theorem 3.7. *Let I be an ideal of Π^d and let Λ and $\mathbb{P}_k$ be as above. Assume that $\mathbb{P}_k$ satisfies the three-term relation* (3.1) *whose coefficient matrices satisfy the rank condition* (3.4).

(i) *There is a linear functional $\mathcal{L}$ on $\mathcal{S}_I$ for which P_α are orthogonal polynomials with respect to the bilinear form $\mathcal{L}(fg) = \langle f, g\rangle$.*
(ii) *If $\mathbf{n}$ is finite then there exist a set V of isolated points and a real function W on V, $W(x) \neq 0$ for all $x \in V$, such that $\{P_\alpha\colon \alpha \in \Lambda\}$ is a sequence of orthogonal polynomials with respect to W on V.*

Proof. (i) The proof follows along the line of the proof of Favard's theorem for continuous orthogonal polynomials of several variables. We shall be brief whenever proofs in the two cases are essentially the same. Using induction, it follows from the three-term relation and the rank condition that the leading coefficients G_k of $\mathbb{P}_k$ are invertible. The linear functional, $\mathcal{L}$, defined by

$$\mathcal{L}1 = 1 \qquad \text{and} \qquad \mathcal{L}(\mathbb{P}_k) = 0, \quad 1 \leqslant k \leqslant \mathbf{n},$$

is well defined for $\mathcal{S}_I$, since every polynomial in $\mathcal{S}_I$ takes the form $\sum_{\alpha \in \Lambda} c_\alpha x^\alpha$ and Λ satisfies (2.1). Furthermore, using (3.6), one can show by induction that $\mathcal{L}$ satisfies $\mathcal{L}(\mathbb{P}_k \mathbb{P}_j^T) = 0$ for $k \neq j$, and the matrix $H_k = \mathcal{L}(\mathbb{P}_k \mathbb{P}_k^T)$ is invertible. Consequently, $\mathbb{P}_k$ are orthogonal polynomials with respect to the bilinear form $\mathcal{L}(fg) = \langle f, g \rangle$.

(ii) We only need to show that the linear functional $\mathcal{L}$ can be represented by a sum over a set of isolated points; that is, $\mathcal{L}$ can be written as

$$\mathcal{L}f = \lambda_1 f(\mathbf{x}_1) + \cdots + \lambda_N f(\mathbf{x}_N) \tag{3.7}$$

for some $\mathbf{x}_i$ with $\lambda_i \neq 0$ for $1 \leqslant i \leqslant N$. Assume that $\mathcal{L}$ has such an expression. If $\mathbf{x}_i$ are known then $\mathcal{L}P_\alpha = \delta_{\alpha,0}$, $\alpha \in \Lambda$, becomes a system of equations on λ_i. The coefficient matrix of this system is $[P_\alpha(\mathbf{x}_i)]$, where $\alpha \in \Lambda$ and $1 \leqslant i \leqslant N$. In particular, if $N = |\Lambda|$ then the matrix is a square matrix. Its determinant is a polynomial in variables $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$ and defines a hypersurface in $\mathbb{R}^{dN}$. Hence, for almost all choices of the values of $\mathbf{x}_1, \ldots, \mathbf{x}_N$, the determinant is nonzero. Furthermore, by Cremer's rule, it is possible to choose a set $V = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ such that $\lambda_i \neq 0$ for $1 \leqslant i \leqslant N$. Hence, P_α are orthogonal with respect to W on V, where W is defined by $W(\mathbf{x}_i) = \lambda_i$, $1 \leqslant i \leqslant N$. □

A couple of remarks are in order. First of all, if $\mathbf{n}$ is infinity, we do not know if $\mathcal{L}$ can be written as a sum of point evaluations, $\mathcal{L}f = \sum_{x \in V} f(x)$, on a countable set V. In fact, if I is the trivial ideal $\langle 1 \rangle$, then $\Lambda = \mathbb{N}_0^d$; the three-term relation takes the same form as that for continuous orthogonal polynomials and the rank condition remains the same. Hence, the three-term relation and the rank condition are not enough to give further information on the linear functional.

The same phenomenon will happen to the case that $\mathcal{S}_I$ is a product polynomial space, say, $\Pi_m^d \times \Pi^d$, for which the orthogonal polynomials are $p_j(x) q_k(y)$, $0 \leqslant j \leqslant m$, and $k \geqslant 0$. No matter if q_k are orthogonal with respect to a linear functional defined by an integral or to a linear functional defined by an infinite sum, the three-term relation will take the same form and the rank condition will also remain the same.

Secondly, the proof of the theorem only shows that W is nonzero at every point of V. This is enough if we only deal with orthogonality but not orthonormality. The theorem simply states that the rank condition (3.4) and the three-term relation (3.1) are enough to ensure orthogonality. The following corollary is about the case of orthonormality, where we do get positive weight. The difference is in the three-term relations (3.3) vs (3.1).

Corollary 3.8. *Let I be an ideal of Π^d and let Λ and $\mathbb{P}_k$ be as in the theorem. Assume that $\mathbf{n}$ is finite. If $\mathbb{P}_k$ satisfies the three-term relation* (3.3) *whose coefficient matrices satisfy the rank condition* (3.4)*, then there is a set V of isolated points and a positive function W on V such that $\{P_\alpha\colon \alpha \in \Lambda\}$ is a sequence of orthonormal polynomials with respect to W on V.*

Proof. According to the theorem, there is a set V so that P_α are orthogonal with respect to the bilinear form defined by the linear functional $\mathcal{L}$ of the form (3.7). We need to show that $H_k = \mathcal{L}(\mathbb{P}_k\mathbb{P}_k^T)$ is an identity matrix for $0 \leqslant k \leqslant n$. This can be established by induction. Since $\mathbb{P}_0 = 1$ and $\mathcal{L}1 = 1$, we have $H_0 = 1$. By (3.2) with $C_{k+1}^T = A_k$, $A_k H_{k+1} = \mathrm{diag}\{H_k, \ldots, H_k\}A_k$. Assume $H_k = I_{r_k}$. Then $\mathrm{diag}\{H_k, \ldots, H_k\}$ is an identity matrix, so that H_{k+1} is an identity matrix by the rank condition. From the fact that $\mathcal{L}(\mathbb{P}_k\mathbb{P}_k^T) = I$, it follows easily that $\mathcal{L}$ is a positive definite linear functional, which shows in particular that $W(x_i) = \lambda_i > 0$ for all $x_i \in V$. □

The coefficient matrices of the three-term relation (3.3) can be used to define the analog of Jacobi matrices,

$$J_i = \begin{bmatrix} B_{0,i} & A_{0,i} & & & \bigcirc \\ A_{0,i}^T & B_{1,i} & A_{1,i} & & \\ & \ddots & \ddots & \ddots & \\ & & A_{\mathbf{n}-3,i}^T & B_{\mathbf{n}-2,i} & A_{\mathbf{n}-2,i} \\ \bigcirc & & & A_{\mathbf{n}-2,i}^T & B_{\mathbf{n}-1,i} \end{bmatrix}, \quad 1 \leqslant i \leqslant d.$$

According to the three-term relation, J_i is the matrix representation of the operator $x_i : P \mapsto x_i P$. Since $x_i x_j \mathbb{P}_k = x_j x_i \mathbb{P}_k \bmod I(V)$, $1 \leqslant k \leqslant \mathbf{n}$, it follows that these matrices commute. In other words, we have the following proposition.

Proposition 3.9. *If* $\mathbf{n}$ *is finite, then the Jacobi matrices commute,* $J_i J_j = J_j J_i$, $1 \leqslant i, j \leqslant d$.

If $\mathbf{n}$ is infinity, then the matrices become infinite and we need to consider them as operators. Still, the Jacobi matrices formally commute. In the case that $\Lambda = \mathbb{N}_0^d$, see [12].

4. Examples

Some examples have been given in the previous sections. To make them more concrete, let us mention two classical discrete polynomials. See, for example, [3,5].

The Hahn polynomials, $Q(x; a, b, N)$, are discrete orthogonal polynomials defined on the set $V = \{0, 1, \ldots, N\}$ and are orthogonal with respect to the hypergeometric distribution $(a+1)_x(b+1)_{N-x}/(x!(N-x)!)$,

$$\sum_{x=0}^{N} \binom{x+a}{x}\binom{N-y+b}{N-y} Q_n(x; a, b, N) Q_m(x; a, b, N)$$
$$= \frac{(-1)^n n!(b+1)_n(n+a+b+1)_{N+1}}{N!(2n+a+b+1)(-N)_n(a+1)_n}\delta_{n,m}, \quad n, m \leqslant N.$$

Their explicit formulas are given in terms of ${}_3F_2$ series,

$$Q_n(x; a, b, N) := {}_3\widetilde{F}_2\left(\begin{matrix} -n, n+a+b+1, -x \\ a+1, -N \end{matrix}; 1\right), \quad n = 0, 1, \ldots, N,$$

where ${}_3\widetilde{F}_2$ is defined as the usual ${}_3F_2$ with the summation terminating at N. The Hahn polynomials, $Q_n(x) = Q_n(x; a, b, N)$, satisfy the three-term relation

$$-xQ_n(x) = A_nQ_{n+1}(x) - (A_n + C_n)Q_n(x) + C_nQ_{n-1}(x),$$

where

$$A_n = \frac{(n+a+b+1)(n+a+1)(N-n)}{(2n+a+b)(2n+a+b+1)}, \qquad C_n = \frac{n(n+b)(n+a+b+N+1)}{(2n+a+b)(2n+a+b+1)}.$$

The Meixner polynomials, $M_n(x; b, c)$, are discrete orthogonal polynomials defined on the set $V = \mathbb{N}_0$ and are orthogonal with respect to the negative binomial distribution $(b)_x c^x/x!$,

$$\sum_{x=0}^{\infty} \frac{(b)_x}{x!} c^x M_m(x; b, c) M_n(x; b, c) = \frac{c^{-n} n!}{(b)_n (1-c)^b} \delta_{m,n},$$

where $(a)_m$ denote the Pochhammer symbol $(a)_m = a(a+1)\ldots(a+m-1)$. Their explicit formula is given in terms of ${}_2F_1$ series,

$$M_n(x) := M_n(x; b, c) = {}_2\widetilde{F}_1\big(-n, -x; b; 1 - c^{-1}\big), \quad n = 0, 1, 2, \ldots,$$

which satisfies the three-term relation

$$(c-1)xM_n(x) = c(n+b)M_{n+1} - \big(n + (n+b)c\big)M_n(x) + nM_{n-1}(x).$$

Example 4.1. As a special case of Example 3.1, we have the product Hahn polynomials of two variables, $Q_n(x; a_1, b_1, N)Q_m(x; a_2, b_2, M)$, $0 \leqslant n \leqslant N$, $0 \leqslant m \leqslant M$, which are orthogonal on the set $V = \{0, 1, \ldots, N\} \times \{0, 1, \ldots, M\}$ with respect to the weight function

$$W(x, y) = \binom{x + a_1}{x}\binom{N - x + a_1}{N - x}\binom{y + b}{y}\binom{N - y + b}{N - y},$$

and the product of Hahn and Meixner polynomials, $Q_n(x; a_1, a_2, N)M_m(x; b, c)$, $0 \leqslant n \leqslant N$, $m \geqslant 0$, which are orthogonal on the set $V = \{0, 1, \ldots, N\} \times \mathbb{N}_0$ with respect to the weight function

$$W(x, y) = \binom{x + a_1}{x}\binom{N - x + a_1}{N - x} \cdot \frac{(b)_y}{y!} c^y.$$

We also have product Meixner polynomials $M_n(x; b_1, c_1)M_m(x; b_2, c_2)$, $n \geqslant 0$, $m \geqslant 0$, which are orthogonal on the set $V = \mathbb{N}_0^2$ with respect to the weight function $(b_1)_x(b_2)_y c^{x+y}/(x!y!)$. In the last case, we have $r_k = k + 1$ for all $k \geqslant 0$, as for the usual continuous orthogonal polynomials.

For the product orthogonal polynomials, the coefficient matrices $A_{k,i}$ in the three-term relation are given in Example 3.1*. Note that it is easy to get orthonormal Hahn and Meixner polynomials (multiply by square root of the normalization constant), and the product of the orthonormal polynomials gives orthonormal polynomials in two variables.

Example 4.2. Let V be the "triangle" point set in Example 2.2. Then the orthogonal polynomials $\mathbb{P}_k$, $0 \leqslant k \leqslant m$, exist with $r_k = k + 1$. This is the case that works exactly as in the case of the continuous orthogonal polynomials. As an example, let us mention the Hahn polynomials of two variables as defined in [4]. They are given by

$$\begin{aligned}\phi_{n,m}(x, y; \sigma, N) = (-1)^{n+m} &\frac{(\sigma_1 + 1)_n(\sigma_2 + 1)_m(-N + x)_m}{(\sigma_3 + 1)_m(\sigma_2 + \sigma_3 + 2n + 1)_n(-N)_m} \\ &\times Q_n(x; \sigma_1, \sigma_2 + \sigma_3 + 2n + 1, N - m)Q_m(y; \sigma_2, \sigma_3, N - x),\end{aligned}$$

and are orthogonal on the set $V = \{(x, y) \in \mathbb{N}_0^2\colon 0 \leqslant x + y \leqslant N\}$ with respect to the weight function

$$W(x, y) = \binom{x + \sigma_1}{x}\binom{y + \sigma_1}{y}\binom{N - x - y + \sigma_3}{N - x - y}.$$

In this case $r_k = k + 1$ and the matrix $A_{k,1}$ and $A_{k,2}$ are of the size $(k + 1) \times (k + 2)$, just as in the usual continuous case.

For extensions of classical discrete orthogonal polynomials to several variables, we refer to [4,9,10]. One can also extend discrete q-orthogonal polynomials to several variables.

In the following we consider an example in which V contains 8 points and Λ is given as in the right figure of Fig. 1.

Example 4.3. We set

$$V = \big\{(-1, -1),\ (0, -1),\ (1, -1),\ (-1, 0),\ (0, 0),\ (1, 0),\ (-1, 1),\ (-1, 2)\big\}.$$

Then $\mathbb{R}[V] = \operatorname{span}\{1, x, y, x^2, xy, y^2, x^2y, y^3\}$ as shown in Example 2.2. We use the method described in Theorem 3.3 to construct orthogonal polynomials on V with respect to the linear functional

$$\mathcal{L}(f) = \frac{1}{8}\sum_{x \in V} f(x).$$

That is, we compute the matrix $M_\Lambda = \langle x^\Lambda, x^\Lambda\rangle$, where $\langle f, g\rangle = \mathcal{L}(fg)$, and factor it as SDS^T. The orthogonal polynomials are given as follows:

$$
\begin{aligned}
&P_0^0(x,y) = 1,\\
&P_0^1(x,y) = 1 + 4x,\\
&P_1^1(x,y) = 3 + 12x + 22y,\\
&P_0^2(x,y) = -26 + x + 35x^2 - 4y,\\
&P_1^2(x,y) = 3 + 3x + x^2 + 6y + 8xy,\\
&P_2^2(x,y) = -20 + 31x - x^2 + 11y + 60xy + 51y^2,\\
&P_0^3(x,y) = -20 + 3x + 27x^2 - 45y + 4xy + 56x^2y - 5y^2,\\
&P_1^3(x,y) = -9x + 9x^2 - 50y - 12xy + 12x^2y - 30y^2 + 20y^3,
\end{aligned}
$$

where P_i^k is a polynomial of degree k. By construction, these polynomials are mutually orthogonal and become orthonormal upon multiplying by proper constants. The corresponding dimensions of $\mathbb{P}_0, \mathbb{P}_1, \mathbb{P}_2, \mathbb{P}_3$ are $1, 2, 3, 2$.

In the case that V is a set of lattice point and V satisfies (2.1), we can take $\Lambda = V$. Examples 4.1 and 4.2 are examples of such a case. If we take $V = \Lambda$ in Example 4.3, we get orthogonal polynomials $Q_j^k(x,y) = P_j^k(x+1, y+1)$, where P_j^k are those in the example.

Acknowledgments

The author thanks Paul Terwilliger for the discussion on the Leonard pairs and tridiagonal pairs (see [7,8] and the reference there), during the 7th International Symposium on Orthogonal Polynomials and Special Functions, held in Copenhagen in August, 2003, which suggests this study. The author thanks one referee for his careful review and helpful suggestions.

References

[1] D. Cox, J. Little, D. O'Shea, Ideals, Varieties, and Algorithms, 2nd ed., Springer, Berlin, 1997.
[2] C.F. Dunkl, Y. Xu, Orthogonal Polynomials of Several Variables, Cambridge Univ. Press, 2001.
[3] S. Karlin, J. McGregor, The Hahn polynomials, formulas and an application, Scripta Math. 45 (1974) 176–198.
[4] S. Karlin, J. McGregor, Linear growth models with many types and multidimensional Hahn polynomials, in: R.A. Askey (Ed.), Theory and Applications of Special Functions, Academic Press, New York, 1975, pp. 261–288.
[5] R. Koekoek, R. Swarttouw, The Askey-scheme of hypergeometric orthogonal polynomials and its q-analog, in: Rep. Fac. Tech. Math. Inform., vol. 98-17, Delft, the Netherlands, 1998, available at http://aw.twi.tudelft.nl/~koekoek/research_en.html.
[6] A.F. Nikiforov, S.K. Suslov, V.B. Uvarov, Classical Orthogonal Polynomials of a Discrete Variable, Springer-Verlag, Berlin, 1991.
[7] T. Ito, P. Terwilliger, Tridiagonal pairs and the quantum affine algebra $U_q(\hat{sl}_2)$, http://front.math.ucdavis.edu/math.QA/0310042.

[8] P. Terwilliger, Introduction to Leonard pairs, OPSFA Rome 2001, J. Comput. Appl. Math. 153 (2) (2003) 463–475.
[9] M.V. Tratnik, Multivariable Meixner, Krawtchouk, and Meixner–Pollaczek polynomials, J. Math. Phys. 30 (1989) 2740–2749.
[10] M.V. Tratnik, Some multivariable orthogonal polynomials of the Askey tableau—discrete families, J. Math. Phys. 32 (1991) 2337–2342.
[11] Y. Xu, On multivariable orthogonal polynomials, SIAM J. Math. Anal. 24 (1993) 783–794.
[12] Y. Xu, Block Jacobi matrices and zeros of multivariable orthogonal polynomials, Trans. Amer. Math. Soc. 342 (1994) 855–866.