Lecture notes on combinatorics of orthogonal polynomials

Jang Soo Kim

September 25, 2023

Contents

Chapter 1

Introduction

Orthogonal polynomials are classical objects arising from the study of continued fractions. Due to the long history of orthogonal polynomials, they have now become important objects of study in many areas: classical analysis and PDE, mathematical physics, probability, random matrix theory, and combinatorics.

The combinatorial study of orthogonal polynomials was pioneered by Flajolet and Viennot in 1980s. In these lecture notes we will learn fascinating combinatorial properties of orthogonal polynomials.

We will first study basic properties of orthogonal polynomials based on Chihara's book, Chapter 1 [1]. We will then focus on the combinatorial approach of orthogonal polynomials, which will be based on Viennot's lecture notes [3]. We will also cover more recent developments in the combinatorics of orthogonal polynomials such as their connections with ASEP, staircase tableaux, lecture hall partitions, and orthogonal polynomials of type R_1.

In Chapter 2 we study elementary and classical results of orthogonal polynomials. In Chapter 3 we review basics of enumerative combinatorics. Starting from Chapter 4 we focus on the combinatorics of orthogonal polynomials.

Chapter 2

Basics of orthogonal polynomials

In this chapter we will cover the first chapter of Chihara's book [1].

2.1 Introduction

Since

$$2\cos m\theta \cos n\theta = \cos(m+n)\theta + \cos(m-n)\theta,$$

for nonnegative integers m and n, we have

$$\int_0^\pi \cos m\theta \cos n\theta d\theta = 0, \qquad m \neq n. \tag{2.1.1}$$

In this situation we say that $\cos m\theta$ and $\cos n\theta$ are orthogonal over the interval $(0, \pi)$.

Note that $\cos n\theta$ is a polynomial in $\cos\theta$ of degree n. So we can write $\cos n\theta = T_n(\cos\theta)$ for a polynomial $T_n(x)$ of degree x.

By the change of variable $x = \cos\theta$, (2.1.1) can be rewritten as

$$\int_{-1}^1 T_m(x)T_n(x)(1-x^2)^{-1/2}dx = 0, \qquad m \neq n.$$

The polynomials $T_n(x)$, $n \geq 0$, are called the **Tchebyshev polynomials of the first kind**. The first few polynomials are:

$$\begin{aligned}
T_0(x) &= 1,\\
T_1(x) &= \cos\theta = x,\\
T_2(x) &= \cos 2\theta = 2\cos^2\theta - 1 = 2x^2 - 1,\\
T_3(x) &= 4x^3 - 3x.
\end{aligned}$$

Recall that in an inner product space V with inner product $\langle\cdot,\cdot\rangle$, a set of vectors $v_1, \dots, v_n$ are said to be orthogonal if $\langle v_i, v_j\rangle = 0$ for all $i \neq j$. In this sense the Tchebyshev polynomials $T_n(x)$ are orthogonal, where $V = \mathbb{R}[x]$ is the space of polynomials with real coefficients with the inner product given by

$$\langle f(x), g(x)\rangle = \int_{-1}^1 f(x)g(x)(1-x^2)^{-1/2}dx.$$

We say that $T_n(x)$ are **orthogonal polynomials** with respect to the **weight function** $(1-x^2)^{-1/2}$ on the interval $(-1, 1)$.

Definition 2.1.1. Suppose that $w(x)$ is a nonnegative and integrable function on (a, b) with $\int_a^b w(x)dx > 0$ and $\int_a^b x^n dx < \infty$ for all $n \geq 0$. A sequence of polynomials $\{P_n(x)\}_{n\geq 0}$ is called an **orthogonal polynomial sequence (OPS)** with respect to the **weight function** (or **measure**) $w(x)$ on (a, b) if the following conditions hold:

(1) $\deg P_n(x) = n$, for $n \ge 0$,

(2) $\int_a^b P_m(x)P_n(x)w(x)dx = 0$ for $m \ne n$.

There is another way to define orthogonal polynomials without using the weight function. For a polynomial $f(x)$, if we define

$$\mathcal{L}(f(x)) = \int_a^b f(x)w(x)dx,$$

then $\mathcal{L}(f(x))$ is completely determined by the **moments** $\mu_n = \int_a^b x^n w(x)dx$. So, if we are only interested in polynomials, then we can define a linear functional $\mathcal{L}$ using a moment sequence $\mu_0, \mu_1, \dots$. Not every sequence $\mu_0, \mu_1, \dots$ gives rise to an OPS, though. We will see later a criterion for a sequence to be a moment sequence.

Definition 2.1.2. Let $\mathcal{L}$ be a linear functional defined on the space of polynomials in x. A sequence of polynomials $\{P_n(x)\}_{n\ge 0}$ is called an **orthogonal polynomial sequence (OPS)** with respect to $\mathcal{L}$ if the following conditions hold:

(1) $\deg P_n(x) = n$, $n \ge 0$,

(2) $\mathcal{L}(P_m(x)^2) \ne 0$ for $m \ge 0$,

(3) $\mathcal{L}(P_m(x)P_n(x)) = 0$ for $m \ne n$.

Note that the second condition above was not necessary in Definition 2.1.1 because it follows from the facts that $w(x)$ is nonnegative and $\int_a^b w(x)dx > 0$.

Remark 2.1.3. The moments of the Tchebyshev polynomials are

$$\mu_{2n} = \int_{-1}^1 x^{2n}(1-x^2)^{-1/2}dx = \frac{\pi}{2^{2n}}\binom{2n}{n}, \qquad \mu_{2n+1} = 0.$$

This suggests that there could be some interesting combinatorics behind the scene. We will later find a combinatorial way to understand this situation.

Example 2.1.4 (Charlier polynomials)**.** The **Charlier polynomials** $P_n(x)$ are defined by

$$P_n(x) = \sum_{k=0}^n \binom{x}{k}\frac{(-a)^{n-k}}{(n-k)!},$$

where $\binom{x}{k} = x(x-1)\cdots(x-k+1)/k!$. We will find a different type of orthogonality for $P_n(x)$.

The generating function for $P_n(x)$ is

$$G(x,w) = \sum_{n\ge 0} P_n(x)w^n = \sum_{n\ge 0}\left(\sum_{k=0}^n \binom{x}{k}\frac{(-a)^{n-k}}{(n-k)!}\right)w^n = \sum_{n\ge 0}\binom{x}{n}w^n \sum_{n\ge 0}\frac{(-a)^m}{m!}w^m,$$

which means

$$G(x,w) = e^{-aw}(1+w)^x.$$

Thus

$$a^x G(x,v)G(x,w) = e^{-a(v+w)}\left(a(1+v)(1+w)\right)^x.$$

We have

$$\sum_{k\ge 0}\frac{a^k G(k,v)G(k,w)}{k!} = \sum_{k\ge 0}\frac{e^{-a(v+w)}\left(a(1+v)(1+w)\right)^k}{k!} = e^{-a(v+w)}e^{a(1+v)(1+w)} = e^a e^{avw}.$$

Thus

$$\sum_{k\geq 0} \frac{a^k G(k,v)G(k,w)}{k!} = \sum_{n\geq 0} \frac{e^a (avw)^n}{n!}. \tag{2.1.2}$$

On the other hand

$$\begin{aligned}\sum_{k\geq 0} \frac{a^k G(k,v)G(k,w)}{k!} &= \sum_{k\geq 0} \frac{a^k}{k!} \sum_{m,n\geq 0} P_m(k)P_n(k)v^m w^n \\ &= \sum_{m,n\geq 0} \left(\sum_{k\geq 0} P_m(k)P_n(k)\frac{a^k}{k!} \right) v^m w^n. \end{aligned} \tag{2.1.3}$$

Comparing the coefficients of $v^m w^n$ in (2.1.2) and (2.1.3) we obtain

$$\sum_{k\geq 0} P_m(k)P_n(k)\frac{a^k}{k!} = \frac{e^a a^n}{n!}\delta_{n,m}. \tag{2.1.4}$$

Therefore, if we define a linear functional $\mathcal{L}$ by

$$\mathcal{L}(x^n) = \sum_{k\geq 0} k^n \frac{a^k}{k!},$$

then $P_n(x)$ are orthogonal polynomials with respect to $\mathcal{L}$.

Note that we describe the orthogonality of $P_n(x)$ using only the linear functional $\mathcal{L}$ without referring to any weight function. However, we can also find a weight function in this case. Let $\psi(x)$ be the step function with a jump at $k = 0, 1, 2, \dots$ of magnitude $a^k/k!$. Then the linear functional $\mathcal{L}$ can be written as the following Riemann–Stieltjes integral

$$\mathcal{L}(f(x)) = \int_{-\infty}^{\infty} f(x) d\psi(x).$$

We can also prove (2.1.4) in a combinatorial way, see Appendix A.

Remark 2.1.5. In the theory of orthogonal polynomials, finding an explicit weight function is an important problem. However, in these lecture notes, we will not pursue in this direction and we will be mostly satisfied with Definition 2.1.2.

2.2 The moment functional and orthogonality

We will consider the space $\mathbb{C}[x]$ of polynomials with complex coefficients. A **linear functional** on $\mathbb{C}[x]$ is a map $\mathcal{L} : \mathbb{C}[x] \to \mathbb{C}$ such that $\mathcal{L}(af(x) + bg(x)) = a\mathcal{L}(f(x)) + b\mathcal{L}(g(x))$ for all $f(x), g(x) \in \mathbb{C}[x]$ and $a, b \in \mathbb{C}$.

Definition 2.2.1. Let $\{\mu_n\}_{n\geq 0}$ be a sequence of complex numbers. Let $\mathcal{L}$ be the linear functional on the space of polynomials defined by $\mathcal{L}(x^n) = \mu_n$, $n \geq 0$. In this case we say that $\mathcal{L}$ is the **moment functional** determined by the **moment sequence** $\{\mu_n\}$, and μ_n is called the **nth moment**.

We recall the definition of orthogonal polynomials.

Definition 2.2.2. Let $\mathcal{L}$ be the linear functional defined on the space of polynomials in x. A sequence of polynomials $\{P_n(x)\}_{n\geq 0}$ is called an **orthogonal polynomial sequence (OPS)** with respect to $\mathcal{L}$ if the following conditions hold:

(1) $\deg P_n(x) = n$, $n \geq 0$,

(2) $\mathcal{L}(P_m(x)P_n(x)) = K_n\delta_{m,n}$, for some $K_n \neq 0$.

We say that $P_n(x)$ are **orthonormal** if $\mathcal{L}(P_m(x)P_n(x)) = \delta_{m,n}$.

Theorem 2.2.3. *Let $\{P_n(x)\}$ be a sequence of polynomials and let $\mathcal{L}$ be a linear functional. The following are equivalent:*

(1) $\{P_n(x)\}$ is an OPS with respect to $\mathcal{L}$;

(2) $\mathcal{L}(\pi(x)P_n(x)) = 0$ if $\deg \pi(x) < n$ and $\mathcal{L}(\pi(x)P_n(x)) \neq 0$ if $\deg \pi(x) = n$;

(3) $\mathcal{L}(x^m P_n(x)) = K_n\delta_{m,n}$, $0 \le m \le n$, for some $K_n \neq 0$.

Proof. (1) ⇒ (2): Suppose that $\deg \pi(x) \le n$. Since $\{P_n(x)\}$ is a basis of $\mathbb{C}[x]$, we can write

$$\pi(x) = c_0 + c_1P_1(x) + \cdots + c_nP_n(x).$$

Then

$$\mathcal{L}(\pi(x)P_n(x)) = \sum_{k=0}^{n} \mathcal{L}\left(c_kP_k(x)P_n(x)\right) = c_n\mathcal{L}(P_n(x)^2),$$

which is zero if $\deg \pi(x) < n$ and nonzero if $\deg \pi(x) = n$.

(2) ⇒ (3): Trivial. (2) ⇒ (3): Trivial. □

Theorem 2.2.4. *Suppose that $\{P_n(x)\}_{n\ge 0}$ be an OPS with respect to $\mathcal{L}$. Then for any polynomial $\pi(x)$ of degree n,*

$$\pi(x) = \sum_{k=0}^{n} c_kP_k(x), \qquad c_k = \frac{\mathcal{L}(\pi(x)P_k(x))}{\mathcal{L}(P_k(x)^2)}.$$

Proof. Clearly, we can write

$$\pi(x) = \sum_{k=0}^{n} c_kP_k(x),$$

for some c_k. Multiplying $P_j(x)$ both sides and taking $\mathcal{L}$, we get

$$\mathcal{L}(\pi(x)P_j(x)) = \sum_{k=0}^{n} \mathcal{L}\left(c_kP_k(x)P_j(x)\right) = c_j\mathcal{L}(P_j(x)^2).$$

Dividing both sides by $\mathcal{L}(P_j(x)^2)$, we obtain the theorem. □

Theorem 2.2.5. *Suppose that $\{P_n(x)\}_{n\ge 0}$ be an OPS with respect to $\mathcal{L}$. Then $P_n(x)$ is uniquely determined by $\mathcal{L}$ up to a nonzero factor. More precisely, if $\{Q_n(x)\}_{n\ge 0}$ is an OPS with respect to $\mathcal{L}$, then there are constants $c_n \neq 0$ such that $Q_n(x) = c_nP_n(x)$ for all $n \ge 0$.*

Proof. Let us write $Q_n(x) = \sum_{k=0}^{n} c_kP_k(x)$. Then by Theorem 2.2.4, $c_k = \mathcal{L}(Q_n(x)P_k(x))/\mathcal{L}(P_k(x)^2)$. But by Theorem 2.2.3, $\mathcal{L}(Q_n(x)P_k(x)) = 0$ unless $k = n$. Thus $Q_n(x) = c_nP_n(x)$. □

Note that if $\{P_n(x)\}_{n\ge 0}$ is an OPS for $\mathcal{L}$, then so is $\{c_nP_n(x)\}_{n\ge 0}$ for any $c_n \neq 0$. Therefore there is a unique monic OPS, which is obtained by dividing each $P_n(x)$ by its leading coefficient. Note also that there is a unique orthonormal OPS as well given by $p_n(x) = P_n(x)/\mathcal{L}(P_n(x)^2)^{1/2}$. In summary we have the following corollary.

Corollary 2.2.6. *Suppose that $\mathcal{L}$ is a moment sequence such that there is an OPS for $\mathcal{L}$. Let K_n, $n \ge 0$, be a sequence of nonzero numbers. Then the following hold.*

(1) There is a unique monic OPS $\{P_n(x)\}_{n\ge 0}$ for $\mathcal{L}$.

(2) There is a unique OPS $\{P_n(x)\}_{n\ge 0}$ for $\mathcal{L}$ such that the leading coefficient of $P_n(x)$ is K_n.

(3) There is a unique OPS $\{P_n(x)\}_{n\ge 0}$ for $\mathcal{L}$ such that $\mathcal{L}(x^nP_n(x)) = K_n$.

Clearly, if $\{P_n(x)\}_{n\ge 0}$ is an OPS for $\mathcal{L}$, then it is also an OPS for $\mathcal{L}'$ given by $\mathcal{L}'(f(x)) = c\mathcal{L}(f(x))$ for some $c \neq 0$. Therefore, by dividing the linear functional by the value $\mathcal{L}(1)$, we may assume that $\mathcal{L}(1) = 1$.

2.3 Existence of OPS

The main question in this section is: for what linear functional $\mathcal{L}$ does there exist an OPS? To answer this question we need the following definition.

Definition 2.3.1. The **Hankel determinant** of a moment sequence $\{\mu_n\}$ is defined by

$$\Delta_n = \det(\mu_{i+j})_{i,j=0}^n = \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix}.$$

Theorem 2.3.2. *Let $\mathcal{L}$ be a linear functional with moment sequence $\{\mu_n\}$. Then there is an OPS for $\mathcal{L}$ if and only if $\Delta_n \neq 0$ for all $n \geq 0$.*

Proof. Fix a sequence $\{K_n\}$ of nonzero real numbers K_n. By Corollary 2.2.6, if there is an OPS $\{P_n(x)\}_{n\geq 0}$ for $\mathcal{L}$, it is uniquely determined by the condition $\mathcal{L}(x^n P_n(x)) = K_n$, $n \geq 0$. In other words, using Theorem 2.2.3, there is an OPS for $\mathcal{L}$ if and only if there is a unique sequence $\{P_n(x)\}_{n\geq 0}$ of polynomials such that

$$\mathcal{L}(x^m P_n(x)) = K_n \delta_{m,n}, \qquad 0 \leq m \leq n. \tag{2.3.1}$$

Now let $P_n(x) = \sum_{k=0}^n c_{n,k} x^k$. Multiplying both sides by x^m and taking $\mathcal{L}$, we get

$$\mathcal{L}(x^m P_n(x)) = \sum_{k=0}^n c_{n,k} \mu_{n+k}.$$

Thus (2.3.1) can be written as the matrix equation

$$\begin{pmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{pmatrix} \begin{pmatrix} c_{n,0} \\ c_{n,1} \\ \vdots \\ c_{n,n} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ K_n \end{pmatrix}. \tag{2.3.2}$$

Then the uniqueness of the polynomials $P_n(x)$ satisfying (2.3.1) is equivalent to the uniqueness of the solution of the matrix equation (2.3.2) in $c_{n,0}, c_{n,1}, \ldots, c_{n,n}$. In order for (2.3.2) to have a unique solution, the Hankel determinant Δ_n must be nonzero for all $n \geq 0$. Moreover, by Cramer's rule, $c_{n,n} = K_n \Delta_n / \Delta_{n-1}$ is nonzero iff $\Delta_n \neq 0$. This proves the theorem. □

Applying Cramer's rule to (2.3.2) we can prove the following lemma, which will be used later.

Lemma 2.3.3. *Let $\{P_n(x)\}_{n\geq 0}$ be an OPS for $\mathcal{L}$. Then for a polynomial $\pi(x)$ of degree n we have*

$$\mathcal{L}(\pi(x) P_n(x)) = \frac{ab\Delta_n}{\Delta_{n-1}},$$

where a and b are the leading coefficients of $\pi(x)$ and $P_n(x)$, respectively. In particular, if $\{P_n(x)\}_{n\geq 0}$ is the monic OPS for $\mathcal{L}$, then

$$\mathcal{L}(P_n(x)^2) = \frac{\Delta_n}{\Delta_{n-1}}.$$

Proof. We use the notation in the proof of Theorem 2.3.2. By solving (2.3.2) using Cramer's rule, we obtain that the leading coefficient of $P_n(x)$ is $b = c_{n,n} = K_n \Delta_{n-1}/\Delta_n$. Thus, if we let $\pi(x) = \sum_{k=0}^n a_k x^k$, we have

$$\mathcal{L}(\pi(x) P_n(x)) = \sum_{k=0}^n \mathcal{L}(a_k x^k P_n(x)) = a_n \mathcal{L}(x^n P_n(x)) = a K_n = \frac{ab\Delta_n}{\Delta_{n-1}},$$

as desired. □

Similarly every coefficient $c_{n,i}$ of $P_n(x)$ can be computed using (2.3.2). Thus we have an explicit determinant formula for $P_n(x)$.

Theorem 2.3.4. *Let $\mathcal{L}$ be a linear functional with moment sequence $\{\mu_n\}$ with $\Delta_n \neq 0$ for all $n \geq 0$. Then the monic OPS for $\mathcal{L}$ is given by*

$$P_n(x) = \frac{1}{\Delta_{n-1}} \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-1} \\ 1 & x & \cdots & x^n \end{vmatrix}.$$

Proof. This can be proved using (2.3.2). We can also prove directly that $\{P_n(x)\}_{n\geq 0}$ satisfies the conditions for an OPS. First, the coefficient of x^n in $P_n(x)$ is 1, so $\deg P_n(x) = n$. For $0 \leq k \leq n$, we have

$$\mathcal{L}(x^k P_n(x)) = \frac{1}{\Delta_{n-1}} \mathcal{L}\left(\begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-1} \\ x^k & x^{k+1} & \cdots & x^{n+k} \end{vmatrix} \right) = \frac{1}{\Delta_{n-1}} \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-1} \\ \mu_k & \mu_{k+1} & \cdots & \mu_{n+k} \end{vmatrix}.$$

If $k < n$, then the right-hand side of the above equation has two identical rows, hence zero. If $k = n$, the right-hand side is $\Delta_n/\Delta_{n-1} \neq 0$. This implies that $\{P_n(x)\}_{n\geq 0}$ is an OPS for $\mathcal{L}$. □

In many important cases of orthogonal polynomials there is a nonnegative weight function $w(x)$ representing the moment functional: $\mathcal{L}(x^n) = \int_a^b x^n w(x)dx$. In more general cases, $\mathcal{L}$ can be represented using the Riemann–Stieltjes integral $\mathcal{L}(x^n) = \int_a^b x^n d\psi(x)$, where $\psi(x)$ is a non-decreasing function such that $\{x : \psi(x+\epsilon) - \psi(x-\epsilon) > 0 \text{ for all } \epsilon > 0\}$ is an infinite set. It is known [1, Chapter 2] that there is such an expression if and only if $\mathcal{L}(\pi(x)) > 0$ for all nonzero polynomials $\pi(x)$ such that $\pi(x) \geq 0$ for all $x \in \mathbb{R}$.

A **nonnegative-valued** polynomial is a polynomial $\pi(x)$ such that $\pi(x) \geq 0$ for all $x \in \mathbb{R}$.

Definition 2.3.5. A linear functional $\mathcal{L}$ is **positive-definite** if $\mathcal{L}(\pi(x)) > 0$ for all nonzero nonnegative-valued polynomials $\pi(x)$.

If $\mathcal{L}$ is positive-definite, then it has a real OPS. We will see later that the converse is not true.

Theorem 2.3.6. *Let $\mathcal{L}$ be a positive-definite linear functional. Then $\mathcal{L}$ has real moments and there is a real OPS for $\mathcal{L}$.*

Proof. First, we show that the moments μ_n are real. Since $\mathcal{L}$ is positive-definite, $\mu_{2n} = \mathcal{L}(x^{2n}) > 0$ is real. Since $\mathcal{L}((x+1)^{2n}) = \sum_{k=0}^{2n} \binom{2n}{k}\mu_k$ is real, by induction, we obtain that μ_{2n-1} is also real.

Now, we construct a real OPS $\{P_n(x)\}_{n\geq 0}$ for $\mathcal{L}$. Let $P_0(x) = 1$. Suppose that we have constructed real polynomials $P_0, \dots, P_n$ which are orthogonal with respect to $\mathcal{L}$, i.e., for $0 \leq i, j \leq n$, $\mathcal{L}(P_i(x)P_j(x))$ is zero if $i \neq j$ and nonzero if $i = j$. Now we need to find

$$P_{n+1}(x) = x^{n+1} + \sum_{k=0}^{n} a_k P_k(x) \tag{2.3.3}$$

such that $\mathcal{L}(P_k(x)P_{n+1}(x)) = 0$ for all $0 \leq k \leq n$. Multiplying $P_k(x)$ and taking $\mathcal{L}$ in (2.3.3) we get $\mathcal{L}(P_k(x)P_{n+1}(x)) = \mathcal{L}(x^{n+1}P_k(x)) + a_k\mathcal{L}(P_k(x)^2)$. Thus, if we set

$$a_k = -\frac{\mathcal{L}(x^{n+1}P_k(x))}{\mathcal{L}(P_k(x)^2)},$$

which is real, then $P_{n+1}(x)$ is orthogonal to $P_0(x), \dots, P_n(x)$. In this way we can construct a real OPS $\{P_n(x)\}_{n\geq 0}$ for $\mathcal{L}$. □

Note that if $\mathcal{L}$ is positive-definite, then $\mathcal{L}(P_n(x)^2) > 0$. Thus in this case we can construct a real orthonormal OPS $\{p_n(x)\}_{n\geq 0}$ by rescaling: $p_n(x) = P_n(x)/\sqrt{\mathcal{L}(P_n(x)^2)}$.

Nonnegative-valued polynomials have the following useful property.

Lemma 2.3.7. *Let $\pi(x)$ be a nonnegative-valued polynomial. Then $\pi(x) = p(x)^2 + q(x)^2$ for some real polynomials $p(x)$ and $q(x)$.*

Proof. Since $\pi(x)$ is real for all real x, the coefficients of $\pi(x)$ are real. This can be seen inductively by observing that if $\deg \pi(x) = n$, then the leading coefficient of $\pi(x)$ is equal to

$$\lim_{x\to\infty} \frac{\pi(x)}{x^n}.$$

Since $\pi(x)$ is a real polynomial such that $\pi(x) \geq 0$, every real zero of $\pi(x)$ has even multiplicity and complex roots appear in conjugate pairs. Thus we can write

$$\pi(x) = r(x)^2 \prod_{k=1}^{m} (x - \alpha_k - \beta_k i)(x - \alpha_k + \beta_k i),$$

where $r(x)$ is a real polynomial and $\alpha_k, \beta_k \in \mathbb{R}$. If we write $\prod_{k=1}^{m}(x - \alpha_k - \beta_k i) = A(x) + iB(x)$, then $\prod_{k=1}^{m}(x - \alpha_k + \beta_k i) = A(x) - iB(x)$. Thus $\pi(x) = r(x)^2(A(x)^2 + B(x)^2)$ as desired. $\square$

By Lemma 2.3.7, we have the following criterion for linear functionals.

Corollary 2.3.8. *A linear functional $\mathcal{L}$ is positive-definite if and only if $\mathcal{L}(p(x)^2) > 0$ for every nonzero real polynomial $p(x)$.*

You may wonder why $\mathcal{L}$ is called "positive-definite". To see this recall that a real $n \times n$ matrix A is positive definite if $u^T A u > 0$ for every nonzero vector $u \in \mathbb{R}^n$. Sylvester's criterion says that A is positive definite if and only if every principal minor of A is positive. The following theorem justifies the terminology "positive-definite" for $\mathcal{L}$.

Theorem 2.3.9. *A linear functional $\mathcal{L}$ is positive-definite if and only if every moment μ_n is real and $\Delta_n > 0$ for all $n \geq 0$. In other words, $\mathcal{L}$ is positive-definite if and only if the Hankel matrix $(\mu_{i+j})_{i,j=0}^{n}$ is positive-definite for all $n \geq 0$.*

Proof. ($\Rightarrow$) By Theorem 2.3.6, the moments are real and there is a real OPS $\{P_n(x)\}_{n\geq 0}$ for $\mathcal{L}$. By Lemma 2.3.3, $\Delta_n/\Delta_{n-1} = \mathcal{L}(P_n(x)^2) > 0$ for $n \geq 0$, where $\Delta_{-1} = 1$. Thus by induction we obtain $\Delta_n > 0$ for all $n \geq 0$.

($\Leftarrow$) Since $\Delta_n > 0$, by Theorem 2.3.2, there is an OPS $\{P_n(x)\}_{n\geq 0}$ for $\mathcal{L}$. By Corollary 2.3.8, it suffices to show that $\mathcal{L}(p(x)^2) > 0$ for any nonzero real polynomial $p(x)$. To do this let $p(x) = \sum_{k=0}^{n} a_k P_k(x)$. Then by the orthogonality,

$$\mathcal{L}(p(x)^2) = \sum_{k=0}^{n} a_k^2 \mathcal{L}(P_k(x)^2).$$

Since $\Delta_n > 0$, we have $\mathcal{L}(P_k(x)^2) > 0$ by Lemma 2.3.3. Thus $\mathcal{L}(p(x)^2) > 0$ as desired. $\square$

2.4 The three-term recurrence

One important property of orthogonal polynomials is that they satisfy a 3-term recurrence relation.

Theorem 2.4.1. *Let $\mathcal{L}$ be a linear functional with monic OPS $\{P_n(x)\}_{n\geq 0}$. Then these monic orthogonal polynomials satisfy the following 3-term recurrence relation:*

$$P_{n+1}(x) = (x - b_n)P_n(x) - \lambda_n P_{n-1}(x), \qquad n \geq 0, \tag{2.4.1}$$

with initial conditions $P_{-1}(x) = 0$ and $P_0(x) = 1$ for some sequences $\{b_n\}_{n\geq 0}$ and $\{\lambda_n\}_{n\geq 1}$ such that $\lambda_n \neq 0$.

Proof. Since $P_n(x)$ are monic polynomials, $P_{n+1}(x) - xP_n(x)$ has degree at most n. Thus we can write

$$P_{n+1}(x) - xP_n(x) = \sum_{k=0}^{n} a_k P_k(x).$$

By Theorem 2.2.3, multiplying both sides by $P_j(x)$ for $0 \le j \le n-2$ and taking $\mathcal{L}$ gives

$$0 = \mathcal{L}(P_j(x)P_{n+1}(x) - xP_j(x)P_n(x)) = \sum_{k=0}^{n} a_k \mathcal{L}(P_j(x)P_k(x)) = a_j \mathcal{L}(P_j(x)^2).$$

Since $\mathcal{L}(P_j(x)^2) \neq 0$, we obtain $a_j = 0$ for all $0 \le j \le n-2$. Then we alway have $P_{n+1}(x) - xP_n(x) = a_n P_n(x) + a_{n-1}P_{n-1}(x)$ for some constants a_n and a_{n-1}. This implies that the polynomials $P_n(x)$ satisfy the 3-term recurrence relation (2.4.1).

It remains to show that $\lambda_n \neq 0$. Multiplying x^{n-1} both sides of (2.4.1) and taking $\mathcal{L}$ gives

$$0 = \mathcal{L}(x^{n-1}P_{n+1}(x)) = \mathcal{L}(x^n P_n(x)) - b_n \mathcal{L}(x^{n-1}P_n(x)) - \lambda_n \mathcal{L}(x^{n-1}P_{n-1}(x)). \tag{2.4.2}$$

By Lemma 2.3.3, we have $\mathcal{L}(x^n P_n(x)) = \mathcal{L}(P_n(x)P_n(x))$. Thus (2.4.2) implies

$$\lambda_n = \frac{\mathcal{L}(P_n(x)^2)}{\mathcal{L}(P_{n-1}(x)^2)}. \tag{2.4.3}$$

Since $\mathcal{L}(P_n(x)^2) \neq 0$, we get $\lambda_n \neq 0$. ☐

Theorem 2.4.2. *Following the notation in Theorem 2.4.1, we have*

$$\lambda_n = \frac{\mathcal{L}(P_n(x)^2)}{\mathcal{L}(P_{n-1}(x)^2)} = \frac{\Delta_{n-2}\Delta_n}{\Delta_{n-1}^2}, \tag{2.4.4}$$

$$b_n = \frac{\mathcal{L}(xP_n(x)^2)}{\mathcal{L}(P_n(x)^2)}, \tag{2.4.5}$$

$$\mathcal{L}(P_n(x)^2) = \lambda_1 \cdots \lambda_n \mathcal{L}(1) = \frac{\Delta_n}{\Delta_{n-1}}, \tag{2.4.6}$$

$$\Delta_n = \lambda_1^n \lambda_2^{n-1} \cdots \lambda_n^1 \mathcal{L}(1)^{n+1}. \tag{2.4.7}$$

Proof. By Lemma 2.3.3, we have $\mathcal{L}(P_n(x)^2) = \Delta_n/\Delta_{n-1}$. Thus the first identity (2.4.4) follows from (2.4.3).

Multiplying $P_n(x)$ both sides of (2.4.1) and taking $\mathcal{L}$ gives

$$\begin{aligned} 0 &= \mathcal{L}(P_n(x)P_{n+1}(x)) = \mathcal{L}(xP_n(x)^2) - b_n\mathcal{L}(P_n(x)^2) - \lambda_n \mathcal{L}(P_n P_{n-1}(x)) \\ &= \mathcal{L}(xP_n(x)^2) - b_n \mathcal{L}(P_n(x)^2), \end{aligned}$$

which implies (2.4.5).

The identity (2.4.6) is an immediate consequence of (2.4.4). The identity (2.4.7) follows from (2.4.6). ☐

Corollary 2.4.3. *Following the notation in Theorem 2.4.1, the linear functional $\mathcal{L}$ is positive-definite if and only if $b_n \in \mathbb{R}$ and $\lambda_n > 0$ for all n and $\mathcal{L}(1) > 0$.*

Proof. Suppose that $\mathcal{L}$ is positive-definite. Then by Theorem 2.3.6 the polynomials $P_n(x)$ are real, hence the recurrence coefficients b_n and λ_n are real. By Theorem 2.3.9, we have $\Delta_n > 0$, which together with (2.4.4) implies $\lambda_n > 0$.

Now suppose that $b_n \in \mathbb{R}$ and $\lambda_n > 0$ for all n. By (2.4.4) and (2.4.5), one can easily check by induction that all the moments are real. By (2.4.7), we have $\Delta_n > 0$. Thus by Theorem 2.3.9, $\mathcal{L}$ is positive-definite. ☐

Oftentimes non-monic orthogonal polynomials are used in the literature. We can always make them monic by dividing each polynomial by its leading coefficient. This allows us to convert a 3-term recurrence of monic orthogonal polynomials to that of non-monic orthogonal polynomials and vice versa.

Suppose that $\{P_n(x)\}_{n\geq 0}$ is an OPS for $\mathcal{L}$, which is not monic. If k_n is the leading coefficient of $P_n(x)$, then the monic OPS for $\mathcal{L}$ is given by $\{\hat{p}_n(x)\}_{n\geq 0}$, where $\hat{p}_n(x) = P_n(x)/k_n$. Then, by Theorem 2.4.1, we have

$$\hat{p}_{n+1}(x) = (x - b_n)\hat{p}_n(x) - \lambda_n \hat{p}_{n-1}(x), \quad n \geq 0; \quad \hat{p}_{-1}(x) = 0, \hat{p}_0(x) = 1. \tag{2.4.8}$$

Substituting $\hat{p}_n(x) = P_n(x)/k_n$ in the above formula, we get

$$P_{n+1}(x) = (A_n x - B_n)P_n(x) - C_n P_{n-1}(x), \quad n \geq 0; \quad P_{-1}(x) = 0, P_0(x) = k_0, \tag{2.4.9}$$

where $A_n = k_{n+1}/k_n$, $B_n = b_n k_{n+1}/k_n$, and $C_n = \lambda_n k_{n+1}/k_{n-1}$.

Conversely, from the recurrence (2.4.9), the leading coefficient of $P_n(x)$ is $k_n = A_{n-1}A_{n-2}\cdots A_0 k_0$. Hence

$$\hat{p}_n(x) = (A_{n-1}A_{n-2}\cdots A_0 k_0)^{-1} P_n(x),$$

and we can obtain the recurrence (2.4.8) by dividing (2.4.9) by $A_n A_{n-1}\cdots A_0 k_0$.

Example 2.4.4. Since

$$\cos(n+1)\theta + \cos(n-1)\theta = 2\cos\theta\cos n\theta, \qquad n \geq 1,$$

we have

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \qquad n \geq 1.$$

Since $T_0(x) = 1$ and $T_1(x) = x$, we have

$$T_{n+1}(x) = A_n x T_n(x) - T_{n-1}(x), \qquad n \geq 0, \tag{2.4.10}$$

where $T_{-1}(x) = 0$, $A_0 = 1$ and $A_n = 2$ for $n \geq 1$. Thus the monic Tchebyshev polynomials are given by $\hat{T}_n(x) = 2^{1-n}T_n(x)$ for $n \geq 1$. Dividing (2.4.10) by 2^n gives

$$\hat{T}_{n+1}(x) = x\hat{T}_n(x) - \lambda_n \hat{T}_{n-1}(x), \qquad n \geq 0, \tag{2.4.11}$$

where $\lambda_1 = 1/2$ and $\lambda_n = 1/4$ for $n \geq 2$.

Note that in the recurrence (2.4.11) for the (monic) Tchebyshev polynomials, $b_n = 0$. This, in fact, implies that $T_{2n}(x)$ is an even function and $T_{2n+1}(x)$ is an odd function. It also turns out that the odd moments are all zero.

Definition 2.4.5. A linear functional $\mathcal{L}$ is **symmetric** if all of its odd moments are zero.

Theorem 2.4.6. *Let $\mathcal{L}$ be a linear functional with monic OPS $\{P_n(x)\}_{n\geq 0}$. The following are equivalent:*

(1) $\mathcal{L}$ is symmetric.

(2) $P_n(-x) = (-1)^n P_n(x)$ for $n \geq 0$.

(3) In the 3-term recurrence (2.4.1)*, $b_n = 0$ for $n \geq 0$.*

Proof. (1) $\Rightarrow$ (2): Since $\mathcal{L}$ is symmetric, $\mathcal{L}(\pi(-x)) = \mathcal{L}(\pi(x))$ for all polynomials $\pi(x)$. Thus $\mathcal{L}(P_m(-x)P_n(-x)) = \mathcal{L}(P_m(x)P_n(x)) = K_n\delta_{m,n}$. By the uniqueness of orthogonal polynomials, Theorem 2.2.5, we have $P_n(-x) = c_n P_n(x)$ for some $c_n \neq 0$. Comparing their leading coefficients, we obtain $c_n = (-1)^n$.

(2) $\Rightarrow$ (1): Since $P_{2n+1}(-x) = -P_{2n+1}(x)$, $P_{2n+1}(x)$ is an odd polynomial. Thus $\mathcal{L}(P_{2n+1}(x)) = 0$ is a sum of odd moments. This shows by induction that all odd moments are zero.

(2) $\Leftrightarrow$ (3): Let $Q_n(x) = (-1)^n P_n(-x)$. Then the condition in (2) is the same as $P_n(x) = Q_n(x)$. By Theorem 2.4.1, we have

$$\begin{aligned} P_{n+1}(x) &= (x - b_n)P_n(x) - \lambda_n P_{n-1}(x), \\ Q_{n+1}(x) &= (x + b_n)Q_n(x) - \lambda_n Q_{n-1}(x), \end{aligned}$$

where the second recurrence is obtained from the first by replacing x by $-x$ and multiplying both sides by $(-1)^{n+1}$. Clearly, the condition $P_n(x) = Q_n(x)$ is equivalent to $b_n = 0$, $n \geq 0$. □

Recall Theorem 2.4.1, which states that orthogonal polynomials satisfy a 3-term recurrence. The converse of this theorem is also true.

Theorem 2.4.7 (Favard's theorem)**.** *Let $\{P_n(x)\}_{n\geq 0}$ be a sequence of monic polynomials. Then there is a (unique) linear functional $\mathcal{L}$ with $\mathcal{L}(1) = 1$ for which $\{P_n(x)\}_{n\geq 0}$ is an OPS if and only if*

$$P_{n+1}(x) = (x - b_n)P_n(x) - \lambda_n P_{n-1}(x), \qquad n \geq 0, \tag{2.4.12}$$

for some sequences $\{b_n\}_{n\geq 0}$ and $\{\lambda_n\}_{n\geq 1}$ of complex numbers with $\lambda_n \neq 0$. Moreover, $\mathcal{L}$ is positive-definite if and only if $b_n \in \mathbb{R}$ and $\lambda_n > 0$ for all $n \geq 1$.

Proof. The "only if" part is done in Theorem 2.4.1. To prove the "if" part, we assume $\lambda_n \neq 0$ for all $n \geq 1$. Note that if $\{P_n(x)\}_{n\geq 0}$ is an OPS for $\mathcal{L}$, then we must have $\mathcal{L}(P_n(x)) = 0$ for $n \geq 1$. This together with $\mathcal{L}(1) = 1$ completely determines the moments of $\mathcal{L}$. Thus we define $\mathcal{L}$ to be the unique linear functional such that $\mathcal{L}(1) = 1$ and $\mathcal{L}(P_n(x)) = 0$ for $n \geq 1$. We need to show that $\{P_n(x)\}_{n\geq 0}$ is indeed an OPS for $\mathcal{L}$. By Theorem 2.2.3, it suffices to show that

$$\mathcal{L}(x^k P_n(x)) = \lambda_1 \cdots \lambda_n \delta_{k,n}, \qquad 0 \leq k \leq n. \tag{2.4.13}$$

We will prove this by induction on k. By the constriction of $\mathcal{L}$, (2.4.13) is true when $k = 0$. Let $k \geq 1$ and suppose that (2.4.13) holds for $k - 1$. To prove (2.4.13) for k, consider an integer $n \geq k$. Multiplying x^{k-1} to (2.4.12), we get

$$x^k P_n(x) = x^{k-1} P_{n+1}(x) + b_n x^{k-1} P_n(x) + \lambda_n x^{k-1} P_{n-1}(x).$$

By the induction hypothesis, taking $\mathcal{L}$ in the above formula gives

$$\mathcal{L}(x^k P_n(x)) = \begin{cases} 0 & \text{if } 1 \leq k \leq n-1, \\ \lambda_n \mathcal{L}(x^{n-1} P_{n-1}(x)) & \text{if } k = n. \end{cases}$$

Thus (2.4.13) also holds for k, and the claim is established.

The "moreover" statement follows from Corollary 2.4.3. □

2.5 Christoffel–Darboux identities and zeros of orthogonal polynomials

The Christoffel–Darboux identities are useful identities which have many applications in the theory of orthogonal polynomials. In this section we prove these identities and and their application to the zeros of orthogonal polynomials.

Theorem 2.5.1 (The Christoffel–Darboux identities)**.** *Let $\{P_n(x)\}_{n\geq 0}$ be given by the 3-term recurrence* (2.4.1)*. For $n \geq 0$, we have*

$$\sum_{k=0}^{n} \frac{P_k(x)P_k(y)}{\lambda_1 \cdots \lambda_k} = \frac{P_{n+1}(x)P_n(y) - P_{n+1}(y)P_n(x)}{\lambda_1 \cdots \lambda_n (x - y)}, \tag{2.5.1}$$

$$\sum_{k=0}^{n} \frac{P_k(x)^2}{\lambda_1 \cdots \lambda_k} = \frac{P'_{n+1}(x)P_n(x) - P_{n+1}(x)P'_n(x)}{\lambda_1 \cdots \lambda_n}. \tag{2.5.2}$$

Proof. Multiply $P_n(y)$ to (2.4.1) to get

$$P_{n+1}(x)P_n(y) = (x - b_n)P_n(x)P_n(y) - \lambda_n P_{n-1}(x)P_n(y). \tag{2.5.3}$$

Interchanging x and y in (2.5.3) gives

$$P_{n+1}(y)P_n(x) = (y - b_n)P_n(x)P_n(y) - \lambda_n P_{n-1}(y)P_n(x). \tag{2.5.4}$$

Subtracting (2.5.4) from (2.5.3), we have

$$P_{n+1}(x)P_n(y) - P_{n+1}(y)P_n(x) = (x - y)P_n(x)P_n(y) + \lambda_n(P_n(x)P_{n-1}(y) - P_n(y)P_{n-1}(x)).$$

Let $f_k = P_{k+1}(x)P_k(y) - P_{k+1}(y)P_k(x)$. Then we can rewrite the above equation (with n replaced by k) as

$$(x - y)P_k(x)P_k(y) = f_k - \lambda_k f_{k-1}.$$

Dividing both sides by $\lambda_1 \cdots \lambda_k(x - y)$ gives

$$\frac{P_k(x)P_k(y)}{\lambda_1 \cdots \lambda_k} = \frac{f_k}{\lambda_1 \cdots \lambda_k(x - y)} - \frac{f_{k-1}}{\lambda_1 \cdots \lambda_{k-1}(x - y)}.$$

Summing the equation for $k = 0, \dots, n$, we obtain (2.5.1).

Rewriting (2.5.1) as

$$\sum_{k=0}^{n} \frac{P_k(x)P_k(y)}{\lambda_1 \cdots \lambda_k} = \frac{(P_{n+1}(x) - P_{n+1}(y))P_n(y) - P_{n+1}(y)(P_n(x) - P_n(y))}{\lambda_1 \cdots \lambda_n(x - y)}$$

and taking the limit $y \to x$ gives (2.5.2). □

The Christoffel–Darboux identities have an interesting application on the zeros of orthogonal polynomials. We first show that orthogonal polynomials have distinct real zeros if $\mathcal{L}$ is positive-definite.

Lemma 2.5.2. *Let $\mathcal{L}$ be a positive-definite linear functional with monic OPS $\{P_n(x)\}_{n\ge 0}$. Then $P_n(x)$ has n distinct real roots for all $n \ge 1$.*

Proof. Since $\mathcal{L}(P_n(x)) = 0$, $P_n(x)$ must have a root of odd multiplicity. (Because otherwise $P_n(x) \ge 0$ for all $x \in \mathbb{R}$, which in turn implies $\mathcal{L}(P_n(x)) > 0$ by the assumption that $\mathcal{L}$ is positive-definite.) Let $x_1, \dots, x_k$ be the distinct roots of $P_n(x)$ with odd multiplicities. Then $(x - x_1) \cdots (x - x_k)P_n(x) \ge 0$ for all $x \in \mathbb{R}$. Therefore $\mathcal{L}((x - x_1) \cdots (x - x_k)P_n(x)) > 0$. But by Theorem 2.2.3 this implies $k \ge n$. Clearly, $k \le n$ and we obtain $k = n$. This means that $P_n(x)$ has n distinct roots. □

Theorem 2.5.3. *Let $\mathcal{L}$ be a positive-definite linear functional with monic OPS $\{P_n(x)\}_{n\ge 0}$. Then $P_n(x)$ has n distinct real roots for all $n \ge 1$ and the zeros of $P_n(x)$ and $P_{n+1}(x)$ interlace. More precisely, if $x_{n,1} > x_{n,2} > \cdots > x_{n,n}$ are the zeros of $P_n(x)$, then*

$$x_{n+1,1} > x_{n,1} > x_{n+1,2} > x_{n,2} > \cdots > x_{n+1,n} > x_{n,n} > x_{n+1,n+1}. \tag{2.5.5}$$

Proof. The first part is proved in Lemma 2.5.2. For the second part, we substitute $x = x_{n,j}$ in (2.5.2) to get

$$0 < \sum_{k=0}^{n} \frac{P_k(x_{n,j})^2}{\lambda_1 \cdots \lambda_k} = \frac{P'_{n+1}(x_{n,j})P_n(x_{n,j}) - P_{n+1}(x_{n,j})P'_n(x_{n,j})}{\lambda_1 \cdots \lambda_n} = \frac{-P_{n+1}(x_{n,j})P'_n(x_{n,j})}{\lambda_1 \cdots \lambda_n}.$$

This implies that the sign of $P_{n+1}(x_{n,j})$ is the opposite of the sign of $P'_n(x_{n,j})$. Considering the graph of $y = P_n(x)$, the sign of $P'_n(x_{n,j})$ is $(-1)^{j-1}$, see Figure 2.1. Thus the sign of $P_{n+1}(x_{n,j})$, for $j = 1, 2, \dots, n$, is $(-1)^j$ as indicated by the red dots in Figure 2.1. This means that $P_{n+1}(x)$ has a root between each interval $(x_{n,j+1}, x_{n,j})$ for $j = 1, \dots, n - 1$. Considering the limits $\lim_{x\to\infty} P_{n+1}(x) = \infty$ and $\lim_{x\to-\infty} P_{n+1}(x) = (-1)^{n+1}\infty$, we can see that $P_{n+1}(x)$ also has one root in $(x_{n,1}, \infty)$ and one root in $(-\infty, x_{n,n})$. Thus we obtain (2.5.5). □

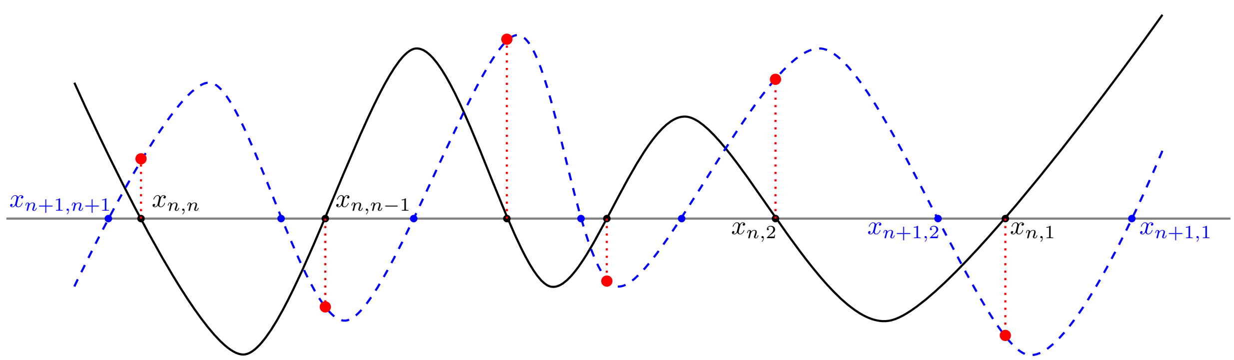

Figure 2.1: The interchanging zeros of $P_n(x)$ and $P_{n+1}(x)$.

Chapter 3

Basics of enumerative combinatorics

In this chapter we review fundamental objects in enumerative combinatorics. From now on we will use the notation $[n] := \{1, \dots, n\}$.

3.1 Formal power series and generating functions

In this section, we study basics of formal power series and generating functions. See [4] for more details on this topic.

A **power series** is a series of the form

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots .$$

The quantity a_n is called the **coefficient of x^n** in $f(x)$. The **constant term** of $f(x)$ is a_0, which we also denote by $f(0)$.

If the coefficients a_n are real numbers, then $f(x)$ may be considered as a function on x whose domain is the set of real numbers x such that the above infinite series converges. For example, if

$$f(x) = 1 + x + x^2 + \cdots ,$$

then we have $f(x) = 1/(1-x)$ for $|x| < 1$. Thus we can write, for $|x| < 1$,

$$1 + x + x^2 + \cdots = \frac{1}{1-x}. \tag{3.1.1}$$

This, however, does not make sense if $|x| > 1$. Hence, in calculus, whenever we consider a power series we always have to mention for what values of x the series converges. But in formal power series the convergence is not needed.

Let R be a commutative ring with identity. Recall that $R[x]$ denotes the ring of polynomials in x with coefficients in R.

Definition 3.1.1. The **ring of formal power series** in x with coefficients in R is the set

$$R[[x]] = \{a_0 + a_1 x + a_2 x^2 + \cdots : a_0, a_1, a_2, \ldots \in R\},$$

with addition

$$\left(\sum_{n=0}^{\infty} a_n x^n\right) + \left(\sum_{n=0}^{\infty} b_n x^n\right) = \sum_{n=0}^{\infty} (a_n + b_n) x^n,$$

and multiplication

$$\left(\sum_{n=0}^{\infty} a_n x^n\right)\left(\sum_{n=0}^{\infty} b_n x^n\right) = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_k b_{n-k}\right) x^n.$$

So, roughly speaking, a formal power series is a polynomial of infinite degree.

The multiplicative identity of $R[[1]]$ is 1, that is, $1 + 0x + 0x^2 + \cdots$. For $f(x), g(x) \in R[[1]]$, if $f(x)g(x) = 1$, then we say that $f(x)$ is the **inverse** of $g(x)$ and write $f(x) = g(x)^{-1} = 1/g(x)$.

In the language of formal power series, (3.1.1) is a perfectly valid identity without any convergence considered because

$$(1 + x + x^2 + \cdots)(1 - x) = (1 + x + x^2 + \cdots) - x(1 + x + x^2 + \cdots) = 1.$$

An important aspect of formal power series is that the coefficient of x^n must be computed using a finitely many additions and multiplications in R.

Example 3.1.2. The series

$$e^{1+x} = \sum_{n\geq 0} \frac{(1+x)^n}{n!}$$

is not a formal power series in $\mathbb{R}[[x]]$ because the constant term (the coefficient of x^0) is $\sum_{n\geq 0} 1/n!$, which cannot be computed by a finite number of additions and multiplications in $\mathbb{R}$ (although we know $\sum_{n\geq 0} 1/n! = e$). On the other hand,

$$e \cdot e^x = \sum_{n\geq 0} \frac{ex^n}{n!}$$

is a formal power series in $\mathbb{R}[[x]]$.

Note that being a formal power series is all about how the series is presented rather than what values the series take as a function. Most of the time, we will not consider a formal power series as a function.

For two formal power series $f(x) = \sum_{n\geq 0} f_n x^n$ and $g(x) = \sum_{n\geq 0} g_n x^n$ with $g_0 = 0$, we define the **composition** $(f \circ g)(x) = f(g(x))$ of $f(x)$ and $g(x)$ by

$$f(g(x)) = \sum_{n\geq 0} f_n g(x)^n. \tag{3.1.2}$$

To see that the above sum is a formal power series, note that since $g_0 = 0$, every term in $f_n g(x)^n$ has degree at least n. Thus, for a fixed $m \geq 0$, the coefficient of x^m in $f(g(x))$ is the coefficient of x^m in the finite sum $\sum_{n=0}^{m} f_n g(x)^n$ of formal power series, which in turn can be computed in a finite number of additions and multiplications in R. Note also that if $g_0 \neq 0$, then the constant term in the sum (3.1.2) is an infinite sum $\sum_{n\geq 0} f_n g_0$, hence $f(g(x))$ is not a formal power series (unless $f(x)$ is a polynomial).

There is a simple criterion for the existence of an inverse of a formal power series.

Proposition 3.1.3. *Let R be a field. A formal power series $f(x) \in R[[x]]$ has an inverse if and only if $f(0) \neq 0$.*

Proof. ($\Rightarrow$) Let $g(x)$ be the inverse of $f(x)$. Suppose that $f(0) = 0$. Then the constant term of $f(x)g(x)$ is $f(0)g(0) = 0$, which is a contradiction to $f(x)g(x) = 1$. Thus we have $f(0) \neq 0$.

($\Leftarrow$) Let $f(x) = \sum_{n\geq 0} f_n x^n$. Then we can write $f(x)$ as

$$f(x) = f_0 (1 - h(x)), \qquad h(x) = \sum_{n\geq 1} h_n x^n, \qquad h_n = -f_0^{-1} f_n.$$

Then the inverse of $f(x)$ can be found in this way:

$$\frac{1}{f(x)} = \frac{1}{f_0} \cdot \frac{1}{1 - h(x)} = \frac{1}{f_0} \sum_{n\geq 0} h(x)^n.$$

Since the lowest degree term of $h(x)^n$ has degree at least n, the above infinite sum is a well-defined formal power series. $\square$

As in calculus we define the **derivative** of a formal power series $f(x) = \sum_{n\geq 0} f_n x^n$ by

$$f'(x) := \sum_{n\geq 1} n f_n x^{n-1} = \sum_{n\geq 0} (n+1) f_{n+1} x^n.$$

The usual differentiation rules hold.

Proposition 3.1.4. *For two formal power series $f(x)$ and $g(x)$, we have*

$$\begin{aligned}
(f(x)g(x))' &= f'(x)g(x) + f(x)g'(x), \\
\left(\frac{f(x)}{g(x)}\right)' &= \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}, \qquad g(x) \neq 0, \\
(f(g(x)))' &= f'(g(x))g'(x), \qquad g(0) = 0.
\end{aligned}$$

Proof. We can prove these identities using the formal definition of the derivative. We will only proof the first identity. Let $f(x) = \sum_{n\geq 0} f_n x^n$ and $g(x) = \sum_{n\geq 0} g_n x^n$. Then

$$(f(x)g(x))' = \left(\sum_{n\geq 0}\left(\sum_{k=0}^{n} f_k g_{n-k}\right) x^n\right)' = \sum_{n\geq 0}\left(\sum_{k=0}^{n} n f_k g_{n-k}\right) x^{n-1}.$$

On the other hand,

$$\begin{aligned}
f'(x)g(x) + f(x)g'(x) &= \sum_{n\geq 0} n f_n x^{n-1} \sum_{n\geq 0} g_n x^n + \sum_{n\geq 0} f_n x^n \sum_{n\geq 0} n g_n x^{n-1} \\
&= \sum_{n\geq 0}\left(\sum_{k=0}^{n} k f_k g_{n-k} + \sum_{k=0}^{n} f_k \cdot (n-k) g_{n-k}\right) x^{n-1} \\
&= \sum_{n\geq 0}\left(\sum_{k=0}^{n} n f_k g_{n-k}\right) x^{n-1}.
\end{aligned}$$

Thus we get the first identity. □

We can naturally extend the definition of formal power series to the multivariate case.

Definition 3.1.5. Let $\mathbf{x} = (x_1, x_2, \dots)$ be a sequence of variables. Let Z denote the set of sequences $I = (i_1, i_2, \dots) \in \mathbb{Z}_{\geq 0}^{\infty}$ such that $i_1 + i_2 + \cdots < \infty$. For $I = (i_1, i_2, \dots) \in Z$, we write $\mathbf{x}^I = x_1^{i_1} x_2^{i_2} \cdots$. The **ring of formal power series** in $x_1, x_2, \dots$ with coefficients in R is the set

$$R[[\mathbf{x}]] = \left\{\sum_{I\in Z} a_I \mathbf{x}^I : a_I \in R\right\},$$

with addition

$$\left(\sum_{I\in Z} a_I \mathbf{x}^I\right) + \left(\sum_{I\in Z} b_I \mathbf{x}^I\right) = \left(\sum_{I\in Z} (a_I + b_I)\mathbf{x}^I\right),$$

and multiplication

$$\left(\sum_{I\in Z} a_I \mathbf{x}^I\right)\left(\sum_{I\in Z} b_I \mathbf{x}^I\right) = \sum_{I\in Z}\left(\sum_{I_1, I_2 \in Z, I_1 + I_2 = I} a_{I_1} b_{I_2}\right)\mathbf{x}^I.$$

Again, rougly speaking, a multivariate formal power series is a multivariate polynomial of infinite degree.

Now we define the notion of generating functions.

Definition 3.1.6. The **generating function** for a sequences $\{a_n\}_{n\geq 0}$ is defined to be the formal power series

$$a_0 + a_1 x + a_2 x^2 + \cdots .$$

So, the generating function for $\{a_n\}_{n\geq 0}$ is nothing but a way of recording the sequence. One of the benefits of generating functions is that we can use many properties of formal power series.

Example 3.1.7. The generating function for $\{a_n = 2^n\}_{n\geq 0}$ is

$$\sum_{n\geq 0} 2^n x^n = \sum_{n\geq 0} (2x)^n = \frac{1}{1-2x}. \tag{3.1.3}$$

Example 3.1.8. Let's find the generating function for $\{a_n = n2^n\}_{n\geq 0}$. Differentiating both sides of (3.1.3), we get

$$\sum_{n\geq 0} n2^n x^{n-1} = \frac{2}{(1-2x)^2}.$$

Multiplying both sides by x, we obtain

$$\sum_{n\geq 0} n2^n x^n = \frac{2x}{(1-2x)^2}.$$

We can easily extend the definition of generating functions to accommodate arrays $\{a_I\}_{I\in Z}$ of elements $a_I \in R$ using multivariate formal power series. More generally, we will consider generating functions for arbitrary (combinatorial) objects.

Definition 3.1.9. Let A be a set of objects. A **weight** on A is a function $\mathrm{wt} : A \to R$, where R is any commutative ring. The **generating function** for A with respect to the weight function wt is the formal power series

$$\sum_{a\in A} \mathrm{wt}(a).$$

Example 3.1.10. Let $A = \{0, 1, 2, \dots\}$ and define a weight of A by $\mathrm{wt}(a) = x^a$. Then the generating function for A (with this weight) is

$$\sum_{a\in A} \mathrm{wt}(a) = \sum_{n=0}^{n} \mathrm{wt}(n) = \sum_{n=0}^{n} x^n = \frac{1}{1-x}.$$

Example 3.1.11. Let A be the set of subsets of $[n]$ and define a weight of A by $\mathrm{wt}(a) = x^{|a|}y^{n-|a|}$. Then the generating function for A (with this weight) is

$$\sum_{a\in A} \mathrm{wt}(a) = \sum_{a\subseteq [n]} x^{|a|}y^{n-|a|} = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = (x+y)^n.$$

Example 3.1.12. Let A be the set S_n of permutations of $[n]$ and define a weight of A by $\mathrm{wt}(a) = x^{\mathrm{cycle}(a)}$. Then it can be proved (see (3.4.4)) that the generating function for A (with this weight) is

$$\sum_{a\in A} \mathrm{wt}(a) = \sum_{\pi\in S_n} x^{\mathrm{cycle}(a)} = x(x+1)\cdots(x+n-1).$$

We will often use the term "generating function" in a flexible manner. For example, the generating function for the number of permutations would mean the generating function for the sequence $\{a_n = n!\}_{n\geq 0}$, that is, $\sum_{n\geq 0} n!x^n$.

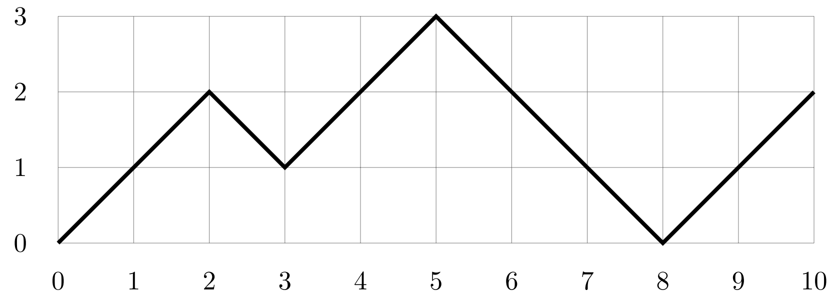

Figure 3.1: A Dyck path from $(0,0)$ to $(10,2)$.

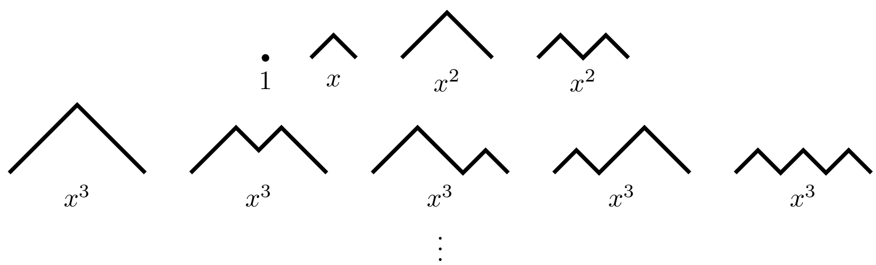

Figure 3.2: An illustration of the generating function for Dyck paths.

3.2 Dyck paths and Motzkin paths

In this section we introduce two important classes of lattice paths. These are fundamental objects in studying orthogonal polynomials combinatorially.

Definition 3.2.1. A **lattice path** from u to v is a sequence $\pi = (v_0, v_1, \dots, v_n)$ of points in $\mathbb{Z}\times\mathbb{Z}$ with $v_0 = u$ and $v_n = v$. Each pair (v_i, v_{i+1}) of consequence points is called a **step** of π.

A path $\pi = (v_0, v_1, \dots, v_n)$ is also considerd as a sequence $S_1 \cdots S_n$ of steps, where $S_i = (v_{i-1}, v_i)$. We will sometimes identify a step (v_i, v_{i+1}) with $v_{i+1} - v_i \in \mathbb{Z}\times\mathbb{Z}$.

Definition 3.2.2. A **Dyck path** is a lattice path consisting of **up steps** $(1,1)$ and **down steps** $(1,-1)$ that stays on or above the x-axis, see Figure 3.1. Denote by $\mathrm{Dyck}(u \to v)$ the set of Dyck paths from u to v. We also define $\mathrm{Dyck}_{2n} = \mathrm{Dyck}((0,0)\to(2n,0))$.

Let's enumerate the Dyck paths in Dyck_{2n} using generating functions. To do this let

$$C(x) = \sum_{n\geq 0} |\,\mathrm{Dyck}_{2n}\,| x^n.$$

Then we can also write

$$C(x) = \sum_{\pi\in\mathrm{Dyck}} \mathrm{wt}(\pi),$$

where Dyck is the set of all Dyck paths from $(0,0)$ to $(2n,0)$ for some $n \geq 0$ and $\mathrm{wt}(\pi) = x^{d(\pi)}$, where $d(\pi)$ is the number of down steps in π. It is helpful to imagine the generating function $C(x)$ as a picture of all Dyck paths, where each Dyck path has its weight attached to it as shown in Figure 3.2.

A Dyck path $\pi \in \mathrm{Dyck}$ can be considered as a sequence of up steps and down steps. For example, the Dyck path in Figure 3.3 is $\pi = UUDUUDDDUDUD$. Every nonempty Dyck path $\pi \in \mathrm{Dyck}$ is uniquely decomposed into $\pi = U\tau D\rho$ for some $\tau, \rho \in \mathrm{Dyck}$. For our running example,

$$\pi = UUDUUDDDUDUD = U(UDUUDDD)(UDUD), \tag{3.2.1}$$

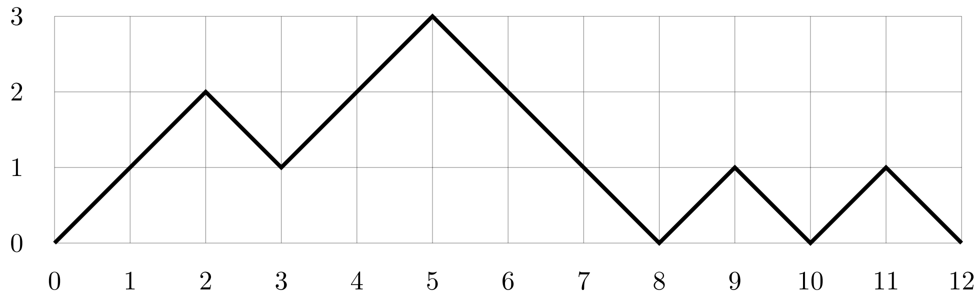

Figure 3.3: A Dyck path π from $(0,0)$ to $(12,0)$.

so we have $\tau = UDUUDDD$ and $\rho = UDUD$. This argument shows that

$$C(x) = 1 + C(x)xC(x). \tag{3.2.2}$$

Solving this quadratic equation for $C(x)$, we get

$$C(x) = \frac{1 \pm \sqrt{1-4x}}{2x}. \tag{3.2.3}$$

We must choose the correct sign here. First, by setting $x = 0$, we obtain that the constant term of $\sqrt{1-4x}$ is 1. Thus (3.2.3) is a valid formal power series only for the minus sign. This implies that

$$\sum_{n\geq 0} |\operatorname{Dyck}_{2n}| x^n = \frac{1-\sqrt{1-4x}}{2x}.$$

Now we can use the **binomial theorem**

$$(1+x)^\alpha := \sum_{n\geq 0} \binom{\alpha}{n} x^n,$$

where

$$\binom{\alpha}{n} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}.$$

By the binomial theorem, we have

$$\begin{aligned}\sqrt{1-4x} = (1-4x)^{1/2} &= \sum_{n\geq 0}\binom{1/2}{n}(-4x)^n = 1 + \sum_{n\geq 1}\frac{\frac{1}{2}\frac{-1}{2}\frac{-3}{2}\cdots\frac{-2n+3}{2}}{n!}(-1)^n 4^n x^n \\ &= 1 - \sum_{n\geq 1}\frac{1\cdot 3\cdot\cdots\cdot(2n-3)}{n!}2^n x^n = 1 - \sum_{n\geq 1}\frac{2(2n-2)!}{n!(n-1)!}x^n.\end{aligned}$$

Therefore,

$$\sum_{n\geq 0} |\operatorname{Dyck}_{2n}| x^n = \frac{1-\sqrt{1-4x}}{2x} = \sum_{n\geq 1}\frac{1}{n}\binom{2n-2}{n-1}x^{n-1} = \sum_{n\geq 0}\frac{1}{n+1}\binom{2n}{n}x^n.$$

Comparing the coefficient of x^n in both sides we obtain the following result.

Proposition 3.2.3. *We have*

$$|\operatorname{Dyck}_{2n}| = \frac{1}{n+1}\binom{2n}{n}. \tag{3.2.4}$$

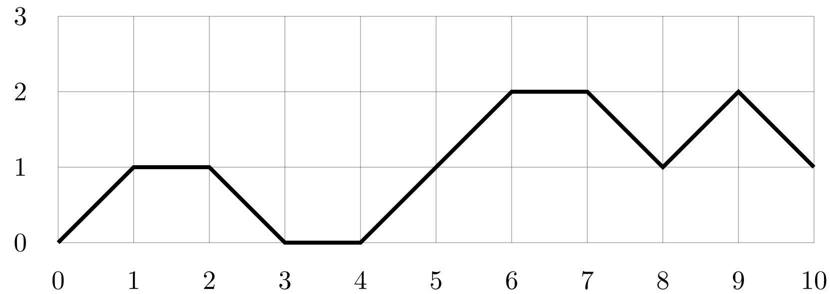

Figure 3.4: A Motzkin path from $(0,0)$ to $(10,1)$.

Note that we proved (3.2.4) using generating functions, but this can also be proved by a standard reflection principle.

The **Catalan number** C_n is defined by

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

The first few Catalan numbers are

$$1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, \dots.$$

There are many combinatorial objects counted by the Catalan number. Stanley [2] collected more than 200 such "Catalan objects". Dyck paths are one of the most well-known Catalan objects. Some of other well known Catalan objects are triangulations of an $(n+2)$-gon, ballot sequences of length $2n$, and plane binary trees with n vertices.

The Catalan numbers satisfy the following recurrence:

$$C_0 = 1, \qquad C_n = \sum_{k=0}^{n} C_k C_{n-1-k}, \qquad n \geq 1. \tag{3.2.5}$$

This recurrence can be proved similarly as (3.2.2) using the decomposition (3.2.1).

Now we consider lattice paths with three kinds of steps. These lattice paths will play a fundamental role in Viennot's theory of orthogonal polynomials.

Definition 3.2.4. A **Motzkin path** is a lattice path consisting of **up steps** $(1,1)$, **horizontal steps** $(1,0)$, and **down steps** $(1,-1)$ that stays on or above the x-axis, see Figure 3.4. Denote by $\mathrm{Motz}(u \to v)$ the set of Dyck paths from u to v. We also define $\mathrm{Motz}_n = \mathrm{Motz}((0,0) \to (n,0))$.

Considering the positions of horizontal steps, we can relate the number of Motzkin paths and that of Dyck paths.

Proposition 3.2.5. *We have*

$$|\mathrm{Motz}_n| = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} C_k.$$

Proposition 3.2.6. *Let $M(x) = \sum_{n\geq 0} |\mathrm{Motz}_n| x^n$. Then*

$$M(x) = \frac{1 - x - \sqrt{1 - 2x - 3x^2}}{2x^2}.$$

Proof. By a similar argument used to prove (3.2.2), we have

$$M(x) = 1 + xM(x) + M(x)x^2M(x).$$

Solving the equation we obtain the desired formula. □

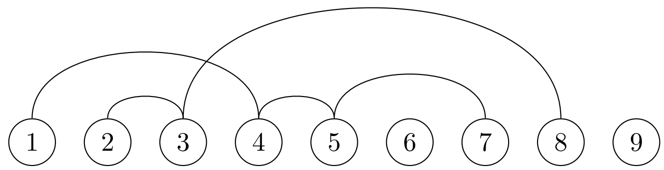

Figure 3.5: A visualization of a set partition $\{\{1,4,5,7\},\{2,3,8\},\{6\},\{9\}\}$ of $[9]$.

3.3 Set partitions and matchings

In this section we study set partitions and matchings. They will be used to give combinatorial interpretations for moments of Charilier polynomials and Hermite polynomials.

Definition 3.3.1. A **set partition** of a set X is a collection $\pi = \{B_1, \dots, B_k\}$ of subsets of X such that

(1) $B_i \ne \emptyset$ for all i,

(2) $B_i \cap B_j = \emptyset$ for all $i \ne j$, and

(3) $B_1 \cup \cdots \cup B_k = X$.

Each B_i is called a **block** of π.

A set partition can be visualized by connecting consecutive elements in each block see Figure 3.5.

We denote by Π_n the set of all set partitions of $[n]$. We also define $\Pi_{n,k}$ to be the set of all set partitions of $[n]$ with exactly k blocks. The **Stirling number of the second kind** $S(n,k)$ is the cardinality of $\Pi_{n,k}$.

We use the convention that $\emptyset$ is the only set partition of $\emptyset$, i.e., $\Pi_0 = \{\emptyset\}$. The following are immediate from the definition of set partitions:

- $S(n,0) = \delta_{n,0}$,
- $S(n,n) = 1$,
- $S(n,k) = 0$ if $k > n$.

We can compute the number $S(n,k)$ using the following recursion with the above initial conditions.

Proposition 3.3.2. *For integers $n, k \ge 1$, we have*

$$S(n,k) = S(n-1,k-1) + kS(n-1,k).$$

Proof. Let $\pi \in \Pi_{n,k}$. If n is in a singleton of π, then $\pi \setminus \{n\} \in \Pi_{n-1,k-1}$. Otherwise, π can be obtained from a set partition $\pi' \in \Pi_{n-1,k}$ by adding n to one of the k blocks of π'. This shows the recursion. $\square$

Proposition 3.3.3. *We have*

$$S(n,k) = \frac{1}{k!} \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} i^n.$$

Proof. The number of onto functions $f : [n] \to [k]$ is $k!S(n,k)$. By the principle of inclusion and exclusion, this number is equal to

$$k!S(n,k) = \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} i^n,$$

which implies the desired formula. $\square$

For an integer $n \geq 0$, a **falling factorial** $(x)_n$ is defined by

$$(x)_n = x(x-1)\cdots(x-n+1).$$

Proposition 3.3.4. *We have*

$$\sum_{k=0}^{n} S(n,k)(x)_k = x^n. \tag{3.3.1}$$

Proof. Since both sides are polynomials in x, it suffices to show that the identity holds for all positive integers x. So, let's assume that x is a positive integer. Then the right-hand side is the number of all functions $f : [n] \to [x]$.

Now, consider a function $f : [n] \to [x]$ such that the image $f([n])$ has exactly k elements. Let $f([n]) = \{a_1 < \cdots < a_k\}$. Then $\{f^{-1}(a_1), \dots, f^{-1}(a_k)\}$ is a set partition of $[n]$ with k blocks. Thus such a function f is obtained by first partitionining $[n]$ into k blocks $B_1, \dots, B_k$ and constructing a one-to-one map from $\{B_1, \dots, B_k\}$ to $[x]$. This shows that the number of such functions is $S(n,k)(x)_k$. Summing over all k gives the number of all functions $f : [n] \to [x]$.

Since both sides of the identity count the same number, they are equal. $\square$

Definition 3.3.5. A **matching** on a set X is a set partition $\pi = \{B_1, \dots, B_k\}$ of X in which every block has size 1 or 2. Each block of size 1 is called a **fixed point** and each block of size 2 is called an **edge** or an **arc** of π.

A matching is said to be **perfect** or **complete** if there are no fixed points.

Proposition 3.3.6. *The number of complete matchings of* $[2n]$ *is*

$$(2n-1)!! := 1 \cdot 3 \cdot \cdots \cdot (2n-1).$$

The number of matchings of $[n]$ *is*

$$\sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} (2k-1)!!.$$

Proof. The first identity can easily be proved by induction on n since there are $2n-1$ ways to form an edge with the last element $2n$ and another element.

The second identity follows from the observation that if a matching of $[n]$ has k edges, then these edges form a complete matching on a set of size $2k$. $\square$

3.4 Permutations

In this section we study permutations, which are one of the most fundamental objects in combinatorics. We will see later a connection between permutations and moments of Laguerre polynomials.

Definition 3.4.1. A **permutation** on $[n]$ is a bijection $\pi : [n] \to [n]$. The **symmetric group** $\mathfrak{S}_n$ is the group of permutations on $[n]$ with multiplication given by composition of functions.

For $\pi, \tau \in \mathfrak{S}_n$, we write $\pi\tau = \pi \circ \tau$, that is $\pi\tau$ is the permutation defined by $(\pi\tau)(i) = \pi(\tau(i))$.

Let $\pi : [n] \to [n]$ be a permutation. We will often write $\pi_i = \pi(i)$ and identify this permutation with a word

$$\pi = \pi_1\pi_2 \cdots \pi_n,$$

which is called the **one-line notation** of π. The **two-line notation** of π is the array

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & n \\ \pi_1 & \pi_2 & \cdots & \pi_n \end{pmatrix}.$$

Example 3.4.2. Let $\pi \in \mathfrak{S}_3$ be the permutation given by

$$\pi(1) = 2, \pi(2) = 3, \pi(3) = 1.$$

Then in one-line notation,

$$\pi = \pi_1\pi_2\pi_3 = 231$$

and in two-line notation,

$$\pi = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

We have

$$\pi^2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \qquad \pi^3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}.$$

A **cycle** of π is a sequence $(a_1, \dots, a_k)$ of distinct elements of $[n]$ such that

$$\pi(a_1) = a_2, \qquad \pi(a_2) = a_3, \qquad \dots, \qquad \pi(a_k) = a_1.$$

We denote by $\mathrm{cycle}(\pi)$ the number of cycles in π.

A cycle $(a_1, \dots, a_k)$ is considered to be the same as any of its **cyclic shift** $(a_j, \dots, a_k, a_1, \dots, a_{j-1})$. We also consider a cycle $\rho = (a_1, \dots, a_k)$ as a permutation of $[n]$ such that

$$\rho(i) = \begin{cases} i & \text{if } i \notin \{a_1, \dots, a_k\}, \\ a_{j+1} & \text{if } i = a_j, \end{cases}$$

where $a_{k+1} = a_1$.

A **cycle of length k** is a permutation (in some $\mathfrak{S}_n$) of the form $(a_1, \dots, a_k)$. A **transposition** is a cycle of length 2. A **simple transposition** is a transposition of the form $(i, i+1)$.

Note that for a permutation $\pi = \pi_1 \cdots \pi_n \in \mathfrak{S}_n$ and a transposition $\tau = (i, j) \in \mathfrak{S}_n$ with $i < j$, the product $\pi\tau$ is the permutation obtained from π by interchaning the values π_i and π_j at the positions i and j:

$$\pi\tau = \pi_1 \cdots \pi_{i-1}\pi_j\pi_{i+1} \cdots \pi_{j-1}\pi_i\pi_{j+1} \cdots \pi_n.$$

On the other hand, the product $\tau\pi$ is the permutation obtained from π by interchaning the values i and j. For example, if $\pi = \cdots i \cdots j \cdots$, then $\tau\pi = \cdots j \cdots i \cdots$.

Proposition 3.4.3. *Let $\pi \in \mathfrak{S}_n$. Then we can write $\pi = \rho_1 \cdots \rho_k$ for some disjoint cycles $\rho_1, \dots, \rho_k$ in $\mathfrak{S}_n$. Moreover, we can also write $\pi = s_1 \cdots s_r$ for some (not necessarily disjoint) simple transpositions $s_i \in \mathfrak{S}_n$.*

Proof. Let $\pi \in \mathfrak{S}_n$. Let $m = 1$ and consider the sequence $\pi(m), \pi^2(m), \dots$. Since this is an infinite sequence of integers in $[n]$, we must have $\pi^i(m) = \pi^j(m)$ for some $i < j$. By multiplying π^{-i}, we have $m = \pi^{j-i}(m)$. Thus we can find the smallest integer r such that $\pi^r(m) = m$. Let ρ_1 be the cycle $(k, \pi(k), \pi^2(k), \dots, \pi^{r-1}(k))$.

Now let m be the smallest integer in $[n]$ except those in ρ_1. We repeat this process and obtain cycles $\rho_1, \dots, \rho_k$ whose union as a set is $[n]$. These cycles are disjoint because if ρ_i and ρ_j have a common element then they must be the same cycle.

For the second statement, let $\pi = \pi_1 \cdots \pi_n$. Note that multiplying a simple transposition $(i, i+1)$ on the left of π interchanges π_i and π_{i+1}. Thus we can sort $\pi = \pi_1 \cdots \pi_n$ into the the identity permutation $12 \cdots n$ by multiplying simple transpositions $t_1, \dots, t_r$ on the left, i.e., $\pi t_1 \cdots t_r = id$. Then $\pi = t_r \cdots t_1$, which is a product of simple transpositions. $\square$

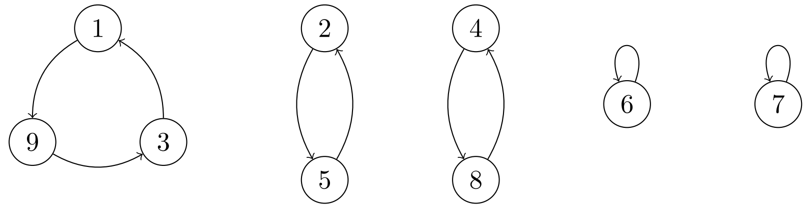

Figure 3.6: A visualization of a permutation $\pi = (1,9,3)(2,5)(4,8)(6)(7) \in \mathfrak{S}_9$.

By Proposition 3.4.3, we can write π in **cycle notation**, i.e., as a product of its disjoint cycles:

$$\pi = \rho_1 \cdots \rho_r.$$

Example 3.4.4. Let $\pi = 951826743 \in \mathfrak{S}_9$. In two-line notation,

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 9 & 5 & 1 & 8 & 2 & 6 & 7 & 4 & 3 \end{pmatrix}.$$

There are 5 disjoint cycles of π, namely, $(1,9,3)$, $(2,5)$, $(4,8)$, (6), and (7). Thus, in cycle notation,

$$\pi = (1,9,3)(2,5)(4,8)(6)(7).$$

Thus, $\mathrm{cycle}(\pi) = 5$. We sometime omit the cycles of length 1 and write

$$\pi = (1,9,3)(2,5)(4,8).$$

We can also visualize a permutation by drawing its cycles as shown in Figure 3.6.

Definition 3.4.5. A permutation $\pi \in \mathfrak{S}_n$ is called an **involution** if $\pi^2 = \iota$, where ι is the identity permutation on $[n]$. Let $\mathfrak{I}_n$ denote the set of involutions in $\mathfrak{S}_n$.

Proposition 3.4.6. *There is a bijection between $\mathfrak{I}_n$ and the set of matchings on $[n]$.*

Proof. A permutation $\pi \in \mathfrak{S}_n$ is an involution if and only if every cycle is of length 1 or 2. Thus, if π is an involution, changing each cycle of π into a block gives a matching on $[n]$. This is clearly a bijection. □

Definition 3.4.7. An **inversion** of a permutation $\pi \in \mathfrak{S}_n$ is a pair (i,j) of integers $1 \le i < j \le n$ such that $\pi(i) > \pi(j)$. We denote by $\mathrm{inv}(\pi)$ the number of inversions of π.

In other words, $\mathrm{inv}(\pi)$ is the pair of integers such that their relative positions are out of orders in π.

Proposition 3.4.8. *We have*

$$\sum_{\pi \in \mathfrak{S}_n} q^{\mathrm{inv}(\pi)} = (1+q)(1+q+q^2)\cdots(1+q+\cdots+q^{n-1}).$$

Proof. We leave this as an exercise. □

Definition 3.4.9. The **sign** of a permutation $\pi \in \mathfrak{S}_n$ is defined to be

$$\mathrm{sgn}(\pi) = (-1)^{\mathrm{inv}(\pi)}.$$

The notion of the sign of a permutation is very important when we study determinants. We will see several ways to compute the sign of a permutation. To this end we need some lemmas.

Figure 3.7: A cycle ρ with i and j on the left and the permutation $\rho\tau$ on the right, where $\tau = (i, j)$.

Lemma 3.4.10. *Let $\pi \in \mathfrak{S}_n$ and let $\tau = (i, j) \in \mathfrak{S}_n$. Then*

$$\operatorname{cycle}(\tau\pi) = \operatorname{cycle}(\pi\tau) = \begin{cases} \operatorname{cycle}(\pi) - 1 & \text{if } i \text{ and } j \text{ are in different cycles of } \pi, \\ \operatorname{cycle}(\pi) + 1 & \text{if } i \text{ and } j \text{ are in the same cycle of } \pi. \end{cases}$$

Proof. Suppose that i and j are in the same cycle, say ρ, of π. Then $\rho\tau$ becomes two cycles as shown in Figure 3.7. Thus in this case $\operatorname{cycle}(\pi\tau) = \operatorname{cycle}(\pi) + 1$. The other cases can be proved similarly.

□

Lemma 3.4.11. *Let $\pi \in \mathfrak{S}_n$ and let $\tau = (i, i+1) \in \mathfrak{S}_n$. Then*

$$\operatorname{sgn}(\pi\tau) = -\operatorname{sgn}(\pi).$$

Proof. Since

$$\pi\tau = \begin{pmatrix} \cdots & i & i+1 & \cdots \\ \cdots & \pi_{i+1} & \pi_i & \cdots \end{pmatrix},$$

we have $\operatorname{inv}(\pi\tau) = \operatorname{inv}(\pi) \pm 1$. This implies $\operatorname{sgn}(\pi\tau) = -\operatorname{sgn}(\pi)$. □

Lemma 3.4.12. *If $\pi \in \mathfrak{S}_n$ is a product of k simple transpositions, then*

$$\operatorname{sgn}(\pi) = (-1)^k.$$

Proof. Let $\pi = t_1 \cdots t_k$, where t_i's are simple transpositions. Then by Lemma 3.4.12,

$$\operatorname{sgn}(\pi) = \operatorname{sgn}(\iota t_1 \cdots t_k) = (-1)^k \operatorname{sgn}(\iota) = (-1)^k.$$

□

Proposition 3.4.13. *For two permutations $\pi, \sigma \in \mathfrak{S}_n$, we have*

$$\operatorname{sgn}(\pi\sigma) = \operatorname{sgn}(\pi)\operatorname{sgn}(\sigma).$$

Proof. Suppose $\pi = t_1 \cdots t_k$ and $\sigma = s_1 \cdots s_r$, where t_i's and s_r's are simple transpositions. Then since $\operatorname{sgn}(\pi) = (-1)^k$ $\operatorname{sgn}(\sigma) = (-1)^r$, we have

$$\operatorname{sgn}(\pi\sigma) = \operatorname{sgn}(t_1 \cdots t_k s_1 \cdots s_r) = (-1)^{k+r} = \operatorname{sgn}(\pi)\operatorname{sgn}(\sigma).$$ □

Proposition 3.4.14. *For $\pi \in \mathfrak{S}_n$, we have*

$$\operatorname{sgn}(\pi) = (-1)^{\operatorname{inv}(\pi)} = (-1)^{n-\operatorname{cycle}(\pi)} = (-1)^{\operatorname{evencycle}(\pi)},$$

where $\operatorname{evencycle}(\pi)$ *is the number of even cycles in π. In particular, if $\pi = t_1 \cdots t_k$, where t_i's are transpositions, then* $\operatorname{sgn}(\pi) = (-1)^t$.

Proof. Let $\pi = t_1 \cdots t_k$, where t_i's are simple transpositions. By the definition of $\mathrm{sgn}(\pi)$ and Lemma 3.4.12, we have $\mathrm{sgn}(\pi) = (-1)^{\mathrm{inv}(\pi)} = (-1)^k$. On the other hand, since $\pi = t_1 \cdots t_k \iota$, by Lemma 3.4.10, $(-1)^{\mathrm{cycle}(\pi)} = (-1)^{\mathrm{cycle}(\iota)+k} = (-1)^{n+k}$. Thus $\mathrm{sgn}(\pi) = (-1)^k = (-1)^{n-\mathrm{cycle}(\pi)}$.

Now let c_i be the number of cycles of length i in π. Then

$$(-1)^{n-\mathrm{cycle}(\pi)} = (-1)^{(1\cdot c_1+2\cdot c_2+\cdots+n\cdot c_n)-(c_1+\cdots+c_n)} = (-1)^{0\cdot c_1+1\cdot c_2+\cdots+(n-1)\cdot c_n} = (-1)^{n-\mathrm{cycle}(\pi)}.$$

The last statement follows from

$$\mathrm{sgn}(\pi) = \mathrm{sgn}(t_1 \cdots t_k) = \mathrm{sgn}(t_1) \cdots \mathrm{sgn}(t_k) = (-1)^k,$$

because the sign of a transposition τ is $\mathrm{sgn}(\tau) = (-1)^{\mathrm{evencycle}(\tau)} = (-1)^1 = -1$. □

The **signless Striling number of the first kind** $c(n,k)$ is defined to be the number of permutations on $[n]$ with k cycles. The **Striling number of the first kind** $s(n,k)$ is defined to by $s(n,k) = (-1)^{n-k}c(n,k)$. Note that $(-1)^{n-k}$ is the sign of a permutation on $[n]$ with k cycles.

Proposition 3.4.15. *For integers $n, k \geq 1$, we have*

$$c(n,k) = c(n-1,k-1) + (n-1)c(n-1,k). \tag{3.4.1}$$

Proof. A permutation $\pi \in \mathfrak{S}_n$ can be obtained from a permutation $\pi' \in \mathfrak{S}_{n-1}$ by creating a new cycle (n) of length 1 or by inserting n after any integer in a cycle of π'. For example, for $\pi' = (1,9,3)(2,5)(4,8)(6)(7) \in \mathfrak{S}_9$, if we insert 10 after 2, we get $\pi = (1,9,3)(2,10,5)(4,8)(6)(7) \in \mathfrak{S}_{10}$, if we insert 10 after 6, we get $\pi = (1,9,3)(2,5)(4,8)(6,10)(7) \in \mathfrak{S}_{10}$, and if we create a new cycle with 10, we get $\pi = (1,9,3)(2,5)(4,8)(6)(7)(10) \in \mathfrak{S}_{10}$. This shows the recursion. □

Proposition 3.4.16. *We have*

$$\sum_{k=0}^{n} s(n,k)x^k = (x)_n. \tag{3.4.2}$$

Equivalently,

$$\sum_{k=0}^{n} c(n,k)x^k = x(x+1)\cdots(x+n-1). \tag{3.4.3}$$

Proof. The equivalence of (3.4.2) and (3.4.3) is obtained by replacing x by $-x$ and multiplying $(-1)^n$ both sides. Thus it suffices to show (3.4.3). This can be proved by induction using (3.4.1). □

Note that (3.4.3) can be rewritten as

$$\sum_{\pi \in \mathfrak{S}_n} x^{\mathrm{cycle}(\pi)} = x(x+1)\cdots(x+n-1). \tag{3.4.4}$$

We can prove this bijectively.

A bijective proof of (3.4.4). We will construct an algorithm to construct a permutation $\pi \in \mathfrak{S}_n$. For $k = 1, \ldots, n$, we do the following.

Step 1 For $k = 1$, create a new cycle consisting of 1.

Step 2 Let $2 \leq k \leq n$ and suppose that the integers $1, \ldots, k-1$ form a permutation on $[k-1]$ in cycle notation. Then we either create a new cycle consisting of k or insert k after one of the integers $1, \ldots, k-1$.

For each $1 \le k \le n$, there are k choices: creating a new cycle (in one way) or inserting k into one of the existing cycles (in $k-1$ ways). The possible choices for k are exactly the same as the choices for the kth factor when we expand

$$x(x+1)(x+1+1)\cdots(x+\overbrace{1+1+\cdots+1}^{n-1}). \tag{3.4.5}$$

Moreover, the first choice (creating a new cycle) corresponds to multipying x. Thus, if π is a permutation obtained in this algoritm, then the same process in the exansion of (3.4.5) gives $x^{\mathrm{cycle}(\pi)}$. This means that the both sides of(3.4.4) match term-by-term, completing the proof of this identity. $\square$

Using (3.3.1) and (3.4.2) we obtain the following matrix identity, which is a duality between Stirling numbers of the first and second kinds.

Proposition 3.4.17. *We have*

$$\Big(S(n,k)\Big)_{n,k\ge 0}\Big(s(n,k)\Big)_{n,k\ge 0} = I, \tag{3.4.6}$$

where $I = (\delta_{n,k})_{n,k\ge 0}$ is the infinite identity matrix. Equivalently, for integers $n, m \ge 0$,

$$\sum_{k\ge 0} S(n,k)s(k,m) = \delta_{n,m}, \tag{3.4.7}$$

$$\sum_{k\ge 0} s(n,k)S(k,m) = \delta_{n,m}. \tag{3.4.8}$$

Proof. By (3.3.1) and (3.4.2), we have the change of basis identities between two bases $\{x^n\}_{n\ge 0}$ and $\{(x)_n\}_{n\ge 0}$ of the vector space of polynomials:

$$\Big(S(n,k)\Big)_{n,k\ge 0}\Big((x)_n\Big)_{n\ge 0} = \Big(x^n\Big)_{n\ge 0},$$
$$\Big(s(n,k)\Big)_{n,k\ge 0}\Big(x^n\Big)_{n\ge 0} = \Big((x)_n\Big)_{n\ge 0}.$$

Thus the two matrices $(S(n,k))_{n,k\ge 0}$ and $(s(n,k))_{n,k\ge 0}$ are inverse of each other, proving (3.4.6). $\square$

Chapter 4

Combinatorial models for OPS

From now one we will focus on the combinatorial approaches to orthogonal polynomials in Viennot's lecture notes [3]. A part of this chapter has some overlaps with Chapter 2.

The main goal of this chapter to give combinatorial interpretations for orthogonal polynomials and their moments. Using these combinatorial interpretations, we will reprove the orthogonality of orthogonal polynomials using combinatorics only.

4.1 Orthogonal polynomials and 3-term recurrences

In this section we recall basic definitions and properties of orthogonal polynomials. We then state the 3-term recurrence of orthogonal polynomials and Favard's theorem.

Let K be a field (we can also use a commutative ring for any result without using divisions). We denote by $K[x]$ the ring of polynomials in x with coefficients in K. A **linear functional** is a linear transformation $\mathcal{L} : K[x] \to K$, i.e., a function satisfying $\mathcal{L}(af(x) + bg(x)) = a\mathcal{L}(f(x)) + b\mathcal{L}(g(x))$ for all $f(x), g(x) \in K[x]$ and $a, b \in K$. The nth **moment** of $\mathcal{L}$ is defined to be $\mu_n = \mathcal{L}(x^n)$.

Definition 4.1.1. Let $\mathcal{L}$ be a linear functional defined on the space of polynomials in x. A sequence of polynomials $\{P_n(x)\}_{n\geq 0}$ is called an **orthogonal polynomial sequence (OPS)** with respect to $\mathcal{L}$ if the following conditions hold:

(1) $\deg P_n(x) = n$, $n \geq 0$,

(2) $\mathcal{L}(P_m(x)P_n(x)) = 0$ for $m \neq n$,

(3) $\mathcal{L}(P_m(x)^2) \neq 0$ for $m \geq 0$.

We also say that $\{P_n(x)\}_{n\geq 0}$ is orthogonal for the moments $\{\mu_n\}_{n\geq 0}$.

Orthogonal polynomials in the above definition are called "formal" or "general" orthogonal polynomials because the field K can be anything. For instance, it may contain arbitrary formal variables such as a, b, c, d. Then the polynomials $P_n(x)$ and the moments μ_n can be treated as polynomials (or more complicated objects such as formal power series or rational functions) in these formal variables.

Proposition 4.1.2. *Suppose that $\{P_n(x)\}_{n\geq 0}$ is an OPS for $\mathcal{L}$.*

(1) $\{P_n(x)\}_{n\geq 0}$ is also orthogonal with respect to $\mathcal{L}'$ for any $\mathcal{L}' = a\mathcal{L}$ for $a \neq 0$.

(2) $\mathcal{L}$ is uniquely determined up to nonzero scalar multiplication.

(3) If we set $\mathcal{L}(1) = 1$, then $\mathcal{L}$ is uniquely determined.

(4) $\{a_nP_n(x)\}_{n\geq 0}$ is an OPS with respect to $\mathcal{L}$ for any sequence $\{a_n\}_{n\geq 0}$ with $a_n \neq 0$.

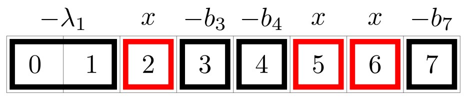

Figure 4.1: A Favard tiling $T \in \mathrm{FT}_8$ with $\mathrm{wt}(T) = \lambda_1 b_3 b_4 b_7 x^3$.

Proof. All statements are easy to check. For example, (2) can be seen by noticing that once the 0th moment $\mu_0 = \mathcal{L}(1)$ is determined, then the nth moment μ_n, for $n \geq 1$, is uniquely determined by the condition $\mathcal{L}(P_n(x)) = 0$. $\square$

By the above proposition we may assume that $\mathcal{L}(1) = 1$. **From now on we will always assume that $\deg P_n(x) = n$ and $\mathcal{L}(1) = 1$ unless otherwise stated.**

Recall from Theorem 2.4.1 that every OPS satisfies a 3-term recurrence.

Theorem 4.1.3 (3-term recurrence)**.** *Let $\mathcal{L}$ be a linear functional with monic OPS $\{P_n(x)\}_{n\geq 0}$. Then there are sequences $\{b_n\}_{n\geq 0}$ and $\{\lambda_n\}_{n\geq 1}$ such that $\lambda_n \neq 0$ and*

$$P_{n+1}(x) = (x - b_n)P_n(x) - \lambda_n P_{n-1}(x), \qquad n \geq 0,$$

where $P_{-1}(x) = 0$ and $P_0(x) = 1$.

The inverse of the above theorem is also true, which is one of the most important results in the theory of classical orthogonal polynomials.

Theorem 4.1.4 (Favard's theorem)**.** *Let $\{P_n(x)\}_{n\geq 0}$ be a sequence of monic polynomials. Then $\{P_n(x)\}_{n\geq 0}$ is an OPS for some linear functional $\mathcal{L}$ if and only if there are sequences $\{b_n\}_{n\geq 0}$ and $\{\lambda_n\}_{n\geq 1}$ such that $\lambda_n \neq 0$ and*

$$P_{n+1}(x) = (x - b_n)P_n(x) - \lambda_n P_{n-1}(x), \qquad n \geq 0, \tag{4.1.1}$$

where $P_{-1}(x) = 0$ and $P_0(x) = 1$.

The main goal of this chapter is to give combinatorial interpretations for the orthogonal polynomials $P_n(x)$ and their moments μ_n. Using these combinatorial interpretations we will prove Favard's theorem bijectively.

4.2 A model for orthogonal polynomials using Favard tilings

In this section we give a combinatorial interpretation for orthogonal polynomials using Favard tilings.

Definition 4.2.1. A **Favard tiling of size n** is a tiling of a $1 \times n$ square board with three types of tiles: red monominos, black monominos, and black dominos. The set of Favard tilings of size n is denoted by FT_n.

We label the squares in the $1 \times n$ board by $0, 1, \dots, n-1$ from left to right. Define the weight $\mathrm{wt}(T)$ of $T \in \mathrm{FT}_n$ to be the product of the weights of the tiles in T, where

(1) the weight of each red monomino is x,

(2) the weight of each black monomino containing a label i is $-b_i$, and

(3) the weight of each domino containing labels $i-1$ and i is $-\lambda_i$.

For example, see Figure 4.1. Note that the number u_n of Favard tilings of size n satisfies $u_{n+1} = 2u_n + u_{n-1}$ with $u_0 = 1$ and $u_1 = 2$. These numbers are called the Pell numbers.

The following proposition gives a combinatorial interpretation for orthogonal polynomials.

Proposition 4.2.2. *Suppose that* $\{P_n(x)\}_{n\geq 0}$ *is a sequence of polynomials satisfying* (4.1.1). *Then*

$$P_n(x) = \sum_{T\in \mathrm{FT}_n} \mathrm{wt}(T).$$

Proof. This is immediate from the recurrence (4.1.1). □

4.3 How to find a combinatorial model for moments

Moments are important quantities of a linear functional $\mathcal{L}$ because they have all the information of $\mathcal{L}$. In this section we will find a combinatorial interpretation for the moments of orthogonal polynomials. To do this we will first take a close look at the moments.

Suppose that $\mathcal{L}$ is a linear functional with monic OPS $\{P_n(x)\}_{n\geq 0}$, which satisfies the 3-term recurrence (4.1.1). Let's assume $\mathcal{L}(1) = 1$. Then, using the orthogonality, we have

$$\mathcal{L}(P_n(x)) = \delta_{n,0}. \tag{4.3.1}$$

This relation in fact completely determines the moments μ_n. For example, since

$$\begin{aligned}
&P_0(x) = 1,\\
&P_1(x) = (x-b_0)P_0(x) - \lambda_0 P_{-1}(x) = x - b_0,\\
&P_2(x) = (x-b_1)P_1(x) - \lambda_1 P_0(x) = x^2 - (b_1+b_0)x + b_0b_1 - \lambda_1,
\end{aligned}$$

we have

$$\begin{aligned}
\mu_0 &= \mathcal{L}(1) = 1,\\
\mu_1 &= \mathcal{L}(x) = \mathcal{L}(P_1(x) + b_0) = b_0,\\
\mu_2 &= \mathcal{L}(x^2) = \mathcal{L}(P_2(x) + (b_0+b_1)x - b_0b_1 + \lambda_1) = (b_0+b_1)b_0 - b_0b_1 + \lambda_1 = b_0^2 + \lambda_1.
\end{aligned}$$

In this way, we can compute a few more moments:

$$\begin{aligned}
\mu_3 &= b_0^3 + 2b_0\lambda_1 + b_1\lambda_1,\\
\mu_4 &= b_0^4 + 3b_0^2\lambda_1 + 2b_0b_1\lambda_1 + b_1^2\lambda_1 + \lambda_1^2 + \lambda_1\lambda_2,\\
\mu_5 &= b_0^5 + 4b_0^3\lambda_1 + 3b_0^2b_1\lambda_1 + 2b_0b_1^2\lambda_1 + b_1^3\lambda_1 + 3b_0\lambda_1^2 + 2b_1\lambda_1^2 + 2b_0\lambda_1\lambda_2 + 2b_1\lambda_1\lambda_2 + b_2\lambda_1\lambda_2.
\end{aligned}$$

The above experiments clearly suggest that μ_n would be a polynomial in b_i's and λ_i's with nonnegative integer coefficients. How can we prove such a conjecture? A satisfying answer to this question is to find combinatorial objects whose generating function is μ_n. That is to find a set X of combinatorial objects and a weight $\mathrm{wt}(A)$ of each element $A \in X$ such that

$$\mu_n = \sum_{A\in X} \mathrm{wt}(A),$$

and $\mathrm{wt}(A)$ is a polynomial (preferably a monomial) in b_i's and λ_i's with nonnegative integer coefficients.

But how can we find such combinatorial objects? Suppose that such combinatorial objects exist with monomial weight $\mathrm{wt}(A)$ for each $A \in X$. Then if we set $b_i = \lambda_i = 1$ for all i then μ_n would be the number of elements in X. If we compute μ_n with this substitution for $n = 0, 1, 2, \ldots$, then we obtain the following sequence:

$$1, 1, 2, 4, 9, 21, 51, 127, 323, 835, 2188, 5798, 15511, \ldots.$$

There is a very useful webpage `https://oeis.org/` where you can search integer sequences. If you search the above sequence, the webpage will tell you that this is the sequence of the number

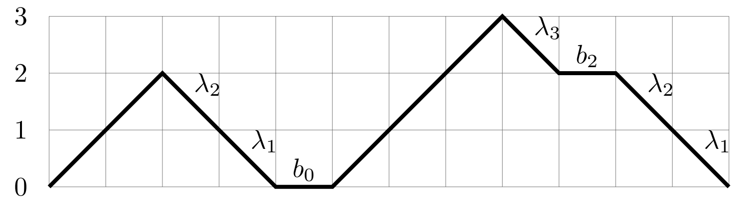

Figure 4.2: A Motzkin path π from $(0,0)$ to $(12,0)$ with $\mathrm{wt}(\pi) = b_0 b_2 \lambda_1^2 \lambda_2^2 \lambda_3$.

of Motzkin paths. So we can guess that there must be a close connection with the moments of orthogonal polynomials and Motzkin paths.

After spending enough time of trials and errors, we can come up with the following combinatorial model for the moments of orthogonal polynomials.

Recall that $\mathrm{Motz}(u \to v)$ is the set of Motzkin paths from u to v. We define the weight $\mathrm{wt}(\pi)$ of a Motzkin path π to be the product of the weights of the steps in π, where

(1) the weight of an up step is 1,

(2) the weight of a horizontal step starting at height i is b_i, and

(3) the weight of a down step starting at height i is λ_i.

See Figure 4.2.

We are now ready state Viennot's combinatorial interpretation for moments of orthogonal polynomials.

Theorem 4.3.1. *Suppose that $\{P_n(x)\}_{n\geq 0}$ is a monic OPS for a linear functional $\mathcal{L}$ with $\mathcal{L}(1) = 1$. Suppose that $\{P_n(x)\}_{n\geq 0}$ satisfy the 3-term recurrence*

$$P_{n+1}(x) = (x - b_n)P_n(x) - \lambda_n P_{n-1}(x), \qquad n \geq 0.$$

Then the moments $\mu_n = \mathcal{L}(x^n)$ are given by

$$\mu_n = \sum_{\pi \in \mathrm{Motz}_n} \mathrm{wt}(\pi).$$

More generally, we will prove a combinatorial interpretation for mixed moments.

Definition 4.3.2. Let $\{P_n(x)\}_{n\geq 0}$ be a monic OPS for a linear functional $\mathcal{L}$. For integers $n, r, s \geq 0$, the **mixed moments** $\mu_{n,r,s}$ and $\mu_{n,k}$ of this OPS are defined by

$$\mu_{n,r,s} = \frac{\mathcal{L}(x^n P_r(x) P_s(x))}{\mathcal{L}(P_s(x)^2)},$$
$$\mu_{n,k} = \mu_{n,0,k} = \frac{\mathcal{L}(x^n P_k(x))}{\mathcal{L}(P_k(x)^2)}.$$

Note that $\mu_n = \mu_{n,0,0}$.

Let $\mathrm{Motz}_{n,r,s}$ denote the set of Motzkin paths from $(0, r)$ to (n, s).

Theorem 4.3.3. *Following the notation in Theorem 4.3.1, we have*

$$\mu_{n,r,s} = \sum_{\pi \in \mathrm{Motz}_{n,r,s}} \mathrm{wt}(\pi).$$

Proof. We proceed by induction on n. Suppose $n = 0$. By the orthogonality of $\{P_n(x)\}_{n\geq 0}$, we have

$$\mu_{0,r,s} = \frac{\mathcal{L}(P_r(x) P_s(x))}{\mathcal{L}(P_s(x)^2)} = \delta_{r,s}.$$

Since $\mathrm{Motz}_{0,r,s} = \emptyset$ if $r = s$ and $\mathrm{Motz}_{0,r,s}$ has only one (empty) Motzkin path if $r = s$, we also have $\sum_{\pi\in\mathrm{Motz}_{n,r,s}} \mathrm{wt}(\pi) = \delta_{r,s}$.

Let $n \geq 1$ and suppose that the theorem holds for $n-1$. Then by the 3-term recurrence,

$$xP_r(x) = P_{r+1}(x) + b_r P_r(x) + \lambda_r P_{r-1}(x).$$

Thus

$$\begin{aligned}
\mu_{n,r,s} &= \frac{\mathcal{L}(x^n P_r(x)P_s(x))}{\mathcal{L}(P_s(x)^2)} = \frac{\mathcal{L}(x^{n-1}(xP_r(x))P_s(x))}{\mathcal{L}(P_s(x)^2)}\\
&= \frac{\mathcal{L}(x^{n-1}(P_{r+1}(x) + b_r P_r(x) + \lambda_r P_{r-1}(x))P_s(x))}{\mathcal{L}(P_s(x)^2)}\\
&= \frac{\mathcal{L}(x^{n-1}P_{r+1}(x)P_s(x))}{\mathcal{L}(P_s(x)^2)} + b_r\frac{\mathcal{L}(x^{n-1}P_r(x)P_s(x))}{\mathcal{L}(P_s(x)^2)} + \lambda_r\frac{\mathcal{L}(x^{n-1}P_{r-1}(x)P_s(x))}{\mathcal{L}(P_s(x)^2)}\\
&= \mu_{n-1,r+1,s} + b_r\mu_{n-1,r,s} + \lambda_r\mu_{n-1,r-1,s}\\
&= \sum_{\pi\in\mathrm{Motz}_{n-1,r+1,s}} \mathrm{wt}(\pi) + b_r \sum_{\pi\in\mathrm{Motz}_{n-1,r,s}} \mathrm{wt}(\pi) + \lambda_r \sum_{\pi\in\mathrm{Motz}_{n-1,r-1,s}} \mathrm{wt}(\pi),
\end{aligned}$$

where the last equation follows from the induction hypothesis.

On the other hand, considering the first step of each $\pi \in \mathrm{Motz}_{n,r,s}$, it is easy to check that

$$\sum_{\pi\in\mathrm{Motz}_{n,r,s}} \mathrm{wt}(\pi) = \sum_{\pi\in\mathrm{Motz}_{n-1,r+1,s}} \mathrm{wt}(\pi) + b_r \sum_{\pi\in\mathrm{Motz}_{n-1,r,s}} \mathrm{wt}(\pi) + \lambda_r \sum_{\pi\in\mathrm{Motz}_{n-1,r-1,s}} \mathrm{wt}(\pi).$$

Hence the theorem holds for n and we are done by induction. □

4.4 A bijective proof of Favard's theorem

We have a combinatorial interpretation for both orthogonal polynomials and their moments. In this section we will prove Favard's theorem bijectively using these combinatorial models.

Suppose that $\{P_n(x)\}_{n\geq 0}$ is a sequence of polynomials satisfying the 3-term recurrence

$$P_{n+1}(x) = (x - b_n)P_n(x) - \lambda_n P_{n-1}(x).$$

To prove Favard's theorem, we need to find a linear functional $\mathcal{L}$ for which $\{P_n(x)\}_{n\geq 0}$ are orthogonal. We simply define $\mathcal{L}$ so that the moments are given by

$$\mathcal{L}(x^n) = \sum_{\pi\in\mathrm{Motz}_n} \mathrm{wt}(\pi). \tag{4.4.1}$$

It is enough to show that

$$\mathcal{L}(P_r(x)P_s(x)) = \lambda_1\cdots\lambda_s\delta_{r,s}.$$

More generally, we will prove

$$\mathcal{L}(x^n P_r(x)P_s(x)) = \lambda_1\cdots\lambda_s \sum_{\pi\in\mathrm{Motz}_{n,r,s}} \mathrm{wt}(\pi). \tag{4.4.2}$$

We first need to give a combinatorial meaning to the the left-hand side of (4.4.2). For a Favard tiling T with k marked monominos, let $\mathrm{wt}'(T) = \mathrm{wt}(T)/x^k$. By Proposition 4.2.2 and the definition (4.4.1) of $\mathcal{L}(x^n)$, we have

$$\mathcal{L}(x^n P_r(x)P_s(x)) = \sum_{(T_1,T_2,\pi)\in X} \mathrm{wt}'(T_1)\,\mathrm{wt}'(T_2)\,\mathrm{wt}(\pi),$$

where X is the set of triples (T_1, T_2, π) such that

(1) $T_1 \in \mathrm{FT}_r$ has i marked monominos for some $0 \le i \le r$,

(2) $T_2 \in \mathrm{FT}_s$ has j marked monominos for some $0 \le j \le s$, and

(3) $\pi \in \mathrm{Motz}_{n+i+j}$.

Let Y be the set of $\pi \in \mathrm{Motz}_{n+r+s}$ such that the first r steps are up steps and the last s step s are down steps. Then the right-hand side of (4.4.2) is equal to $\sum_{\pi \in Y} \mathrm{wt}(\pi)$.

By the above observation, (4.4.2) is equivalent to the following theorem.

Theorem 4.4.1. *For the sets X and Y defined above, we have*

$$\sum_{(T_1,T_2,\pi)\in X} \mathrm{wt}'(T_1)\,\mathrm{wt}'(T_2)\,\mathrm{wt}(\pi) = \sum_{\pi\in Y} \mathrm{wt}(\pi). \tag{4.4.3}$$

Proof. We will find a sign-reversing involution on X.

Consider $(T_1, T_2, \pi) \in X$. We write $\pi = S_1S_2\cdots S_m$ as a sequence of steps and suppose

- a is the largest integer such that T_1 starts with a marked monominos,
- b is the largest integer such that T_2 starts with a marked monominos,
- u is the largest integer such that π starts with u up steps,
- v is the largest integer such that π ends with v down steps.

We define $\phi(T_1, T_2, \pi) = (T_1', T_2', \pi')$ as follows.

Case 1 $u < a$. In this case we set $T_2' = T_2$. There are two subcases.

Case 1-1 S_{u+1} is a horizontal step. Let

$$\pi' = S_1 \cdots \widehat{S_{u+1}} \cdots S_m,$$

and define T_1' to be the Favard tiling obtained from T_1 by replacing the $(u+1)$st marked monomino (at position u) by a unmarked monomino. Here the notation $\widehat{S_{u+1}}$ means that S_{u+1} is removed from the sequence. See Figure 4.3.

Case 1-2 S_{u+1} is a down step. Let

$$\pi' = S_1 \cdots \widehat{S_u}\widehat{S_{u+1}} \cdots S_m,$$

and define T_1' to be the Favard tiling obtained from T_1 by replacing the uth and $(u+1)$st marked monominos (at positions $u-1$ and u) by a domino. See Figure 4.5.

Case 2 $u \ge a \ne r$. In this case we set $T_2' = T_2$. Let A be the $(a+1)$st tile in T_1 (A starts at position a). There are two subcases.

Case 2-1 A is a unmarked monomino. In this case let

$$\pi' = S_1 \cdots S_u H S_{u+1} \cdots S_m,$$

and define T_1' to be the Favard tiling obtained from T_1 by replacing A by a marked monomino. See Figure 4.3.

Case 2-2 A is a domino. In this case let

$$\pi' = S_1 \ldots S_u U D S_{u+1} \ldots S_m,$$

and define T_1' to be the Favard tiling obtained from T_1 by replacing A by two marked monominos. See Figure 4.5.

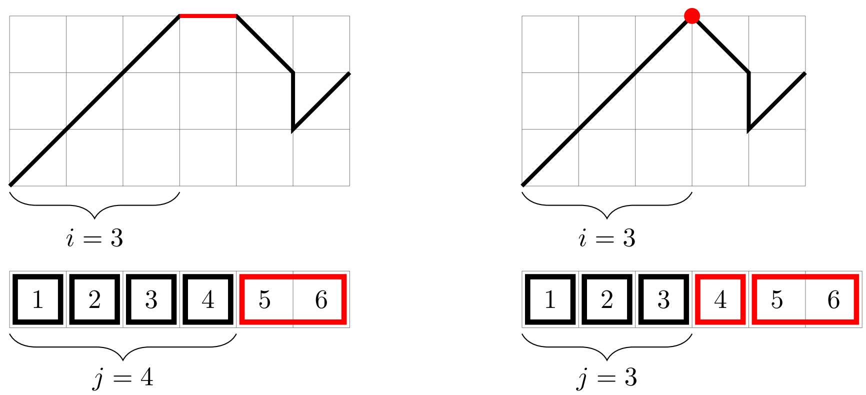

Figure 4.3: A pair $(\pi, T) \in X$ in Case 1-a on the left and the corresponding pair (π', T') in Case 2-a on the right, for $(n, m, \ell) = (2, 6, 2)$. The horizontal step starting at $(3, 3)$ in π is collapsed to a point.

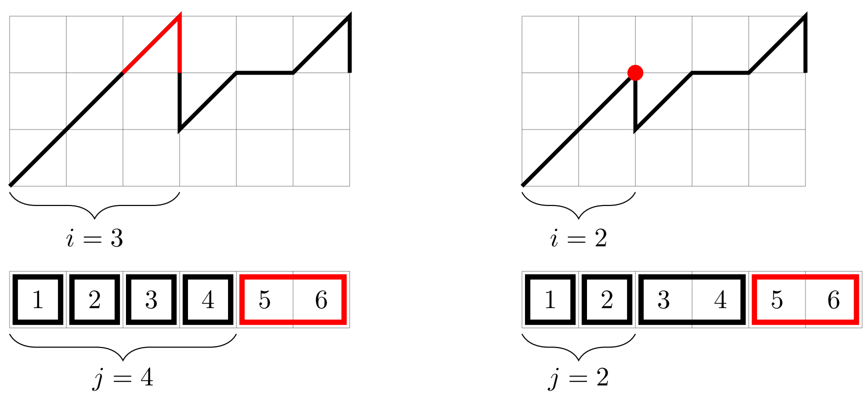

Figure 4.4: A pair $(\pi, T) \in X$ in Case 1-b on the left and the corresponding pair (π', T') in Case 2-b on the right, for $(n, m, \ell) = (2, 6, 2)$. The peak (U, V) starting at $(2, 2)$ in π is collapsed to a point.

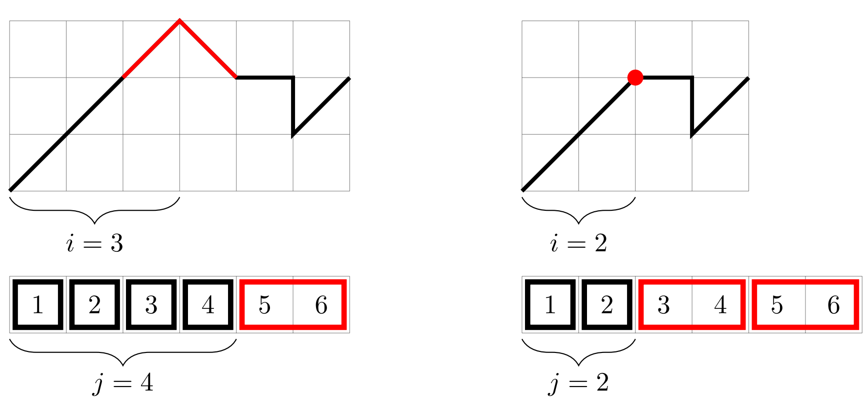

Figure 4.5: A pair $(\pi, T) \in X$ in Case 1-c on the left and the corresponding pair (π', T') in Case 2-c on the right, for $(n, m, \ell) = (2, 6, 2)$. The peak (U, D) starting at $(2, 2)$ in π is collapsed to a point.

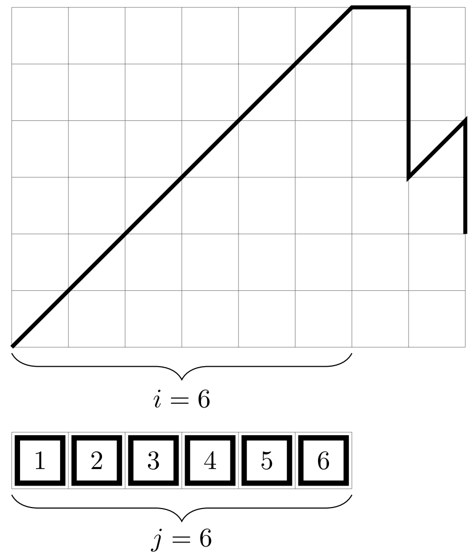

Figure 4.6: A pair $(\pi, T) \in X$ in Case 3 for $(n, m, \ell) = (2, 6, 2)$. In this case $(\pi, T) = (\pi', T')$ is a fixed point.

Case 3 $u \geq a = r$ and $v < b$. This can be done similarly as Case 1. The only difference is that we set $T_1' = T_1$ and consider the steps of π from the right.

Case 4 $u \geq a = r$ and $v \geq b \neq s$. This can be done similarly as Case 2.

Case 5 $u \geq a = r$ and $v \geq b = s$. In this case we set $(T_1', T_2', \pi') = (T_1, T_2, \pi)$.

It is not hard to check that the map $\phi(T_1, T_2, \pi) = (T_1', T_2', \pi')$ is a sign-reversing involution on X with fixed points $(\emptyset, \emptyset, \pi)$ where $\pi \in Y$. $\square$

Chapter 5

Moments of classical orthogonal polynomials

In this section we consider Tchebyshev polynomials of the 1st and 2nd kinds, Laguerre polynomials. Hermite polynomials, Charlier polynomials, and Meixner polynomials of the 1st and 2nd kinds.

Note that an OPS $\{P_n(x)\}_{n\geq 0}$ can be defined in many ways, namely, one of the following determines the orthogonal polynomials:

(1) the coefficients $a_{n,k}$ of $P_n(x)$,

(2) the generating function $\sum_{n\geq 0} P_n(x)t^n$ or $\sum_{n\geq 0} P_n(x)t^n/n!$,

(3) the moments $\{\mu_n\}_{n\geq 0}$,

(4) the 3-term recurrence coefficients $\{b_n\}_{n\geq 0}$ and $\{\lambda_n\}_{n\geq 1}$.

For each OPS, we will show bijectively the equivalence of (3) and (4).

For example, in the case that b_k and λ_k are integers, we interprete $\mathrm{wt}(\alpha)$ as a certain number of "histories". To each history we associate bijectively a certain combinatorial object ξ of a finite set M_n. Each of b_k and λ_k is considered as the number of possible choices in a stage of a construction of the object, where each stage corresponds to an elementary step of α. Then it remains to show that $|M_n| = \mu_n$.

If $P_n(x)$ depend on some parameters, it will be sufficient to consider the histories and the combinatorial objects in M_n.

5.1 Tchebyshev polynomials

Appendix A

Sign-reversing involutions

Definition A.0.1. A **sign** of a set X is a function $\operatorname{sgn} : X \to \{+1, -1\}$. A **sign-reversing involution** on X is an involution $\phi : X \to X$ such that

(1) $\operatorname{sgn}(x) = 1$ for all $x \in \operatorname{Fix}(\phi)$;

(2) $\operatorname{sgn}(\phi(x)) = -\operatorname{sgn}(x)$ for all $x \in X \setminus \operatorname{Fix}(\phi)$,

where $\operatorname{Fix}(\phi)$ is the set of **fixed points** of ϕ, i.e., $\operatorname{Fix}(\phi) = \{x \in X : \phi(x) = x\}$.

It is easy to see that if ϕ is a sign-reversing involution on X, then

$$\sum_{x \in X} \operatorname{sgn}(X) = |\operatorname{Fix}(\phi)|. \tag{A.0.1}$$

Example A.0.2. Let's prove the following identity using sign-reversing involutions:

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0. \tag{A.0.2}$$

To this end we need to construct a set X and a sign-reversing involution ϕ on X such that (A.0.1) becomes (A.0.2).

Let X be the set of all subsets of $[n] := \{1, \dots, n\}$ and for $A \in X$, define $\operatorname{sgn}(A) = (-1)^{|A|}$. Then it suffices to construct a sign-reversing involution on X with no fixed points. This can be done by letting $\phi(A) = A\Delta\{1\}$, where $A\Delta B := (A \cup B) \setminus (A \cap B)$.

Example A.0.3. Recall that we proved the following identitiy, which was stated in (2.1.4), using generating functions:

$$\sum_{k \geq 0} P_m(k) P_n(k) \frac{a^k}{k!} = \frac{e^a a^n}{n!} \delta_{n,m}, \tag{A.0.3}$$

where $P_n(x)$ are the Charlier polynomials defined by

$$P_n(x) = \sum_{k=0}^{n} \binom{x}{k} \frac{(-a)^{n-k}}{(n-k)!}.$$

We will prove this identity using sign-reversing involutions. To do this, we will consider (A.0.3)

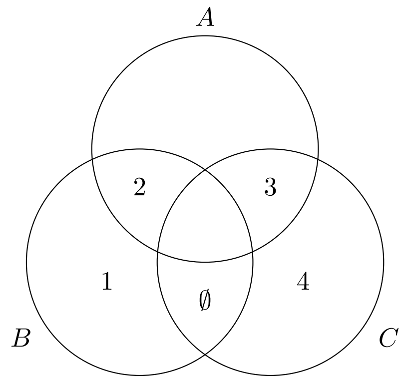

Figure A.1: The triple (A, B, C).

as a power series in a. Note that

$$\begin{aligned}\sum_{k\geq 0} P_m(k)P_n(k)\frac{a^k}{k!} &= \sum_{k\geq 0}\sum_{i=0}^{m}\binom{k}{i}\frac{(-a)^{m-i}}{(m-i)!}\sum_{j=0}^{n}\binom{k}{j}\frac{(-a)^{n-j}}{(n-j)!}\frac{a^k}{k!}\\ &= \sum_{k\geq 0}\sum_{i=0}^{m}\sum_{j=0}^{n}\binom{k}{m-i}\frac{(-a)^i}{i!}\binom{k}{n-j}\frac{(-a)^j}{j!}\frac{a^k}{k!}\\ &= \sum_{N\geq 0}\frac{a^N}{N!}\sum_{i+j+k=N}(-1)^{i+j}\frac{N!}{i!j!k!}\binom{k}{m-i}\binom{k}{n-j},\end{aligned}$$

where $\binom{r}{s} = 0$ if $s < 0$.

For a fixed N,

$$\sum_{i+j+k=N}(-1)^{i+j}\binom{N}{i,j,k}\binom{k}{m-i}\binom{k}{n-j} = \sum_{(A,B,C)\in X}(-1)^{|B\setminus A|+|C\setminus A|},$$

where X is the set of triples (A, B, C) such that $A\cup B\cup C = \{1, \ldots, N\}$, $|A| = k, |B| = m, |C| = n$, $(B \cap C) \setminus A = \emptyset$. Define $\operatorname{sgn}(A, B, C) = (-1)^{|B\setminus A|+|C\setminus A|}$. We will find a sign-reversing involution on X toggling the smallest integer in regions 1 and 2 or in regions 3 and 4 in Figure A.1.

To be precise, for $(A, B, C) \in X$, define $\phi(A, B, C)$ as follows.

Case 1 The regions $1, 2, 3, 4$ are all empty. In this case we define $\phi(A, B, C) = (A, B, C)$.

Case 2 At least one of the regions $1, 2, 3, 4$ is nonempty. Let s be the smallest integer in $(B\cap C)\setminus A$. If s is in region 1 (respectively 2, 3, 4), then move this integer to region 2 (respectively 1, 4, 3). Then let $\phi(A, B, C) = (A', B', C')$, where A', B', C' are the resulting sets.

By the construction, ϕ is a sign-reversing involution on X whose fixed points are the triples (A, B, C) such that the regions $1, 2, 3, 4$ are all empty, that is, $B = C \subseteq A$. If $B = C \subseteq A$, then $A = [N]$, so the number of such triples (A, B, C) is $\binom{N}{n}$ if $m = n$ and 0 otherwise. Thus

$$\sum_{(A,B,C)\in X}(-1)^{|B\setminus A|+|C\setminus A|} = |\operatorname{Fix}(\phi)| = \delta_{m,n}\binom{N}{n}.$$

This implies

$$\sum_{k\geq 0} P_m(k)P_n(k)\frac{a^k}{k!} = \delta_{m,n}\sum_{N\geq 0}\frac{a^N}{N!}\binom{N}{n} = \frac{e^a a^n}{n!}\delta_{n,m}.$$

Bibliography

[1] T. S. Chihara. *An introduction to orthogonal polynomials*. Gordon and Breach Science Publishers, New York, 1978. Mathematics and its Applications, Vol. 13.

[2] R. P. Stanley. *Catalan numbers*. Cambridge University Press, New York, 2015.

[3] G. Viennot. Une théorie combinatoire des polynômes orthogonaux généraux. Lecture Notes, UQAM, 1983.

[4] H. S. Wilf. *generatingfunctionology*. A K Peters, Ltd., Wellesley, MA, third edition, 2006.