Lecture notes on combinatorics of orthogonal polynomials

Jang Soo Kim

November 9, 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 [2]. We will then focus on the combinatorial approach of orthogonal polynomials, which will be based on Viennot's lecture notes [9]. 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 [2].

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 $n \geq 0$,

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

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 $\mu_0, \mu_1, \ldots$. Not every sequence $\mu_0, \mu_1, \ldots$ 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 $\{P_n(x)\}_{n\geq 0}$ 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) $\mathcal{L}(P_m(x)^2) \neq 0$ for $m \geq 0$,

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

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

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

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

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

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 $k = 0, 1, 2, \ldots$ 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 ***n*th 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) $\Rightarrow$ (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) $\Rightarrow$ (3): Trivial. (2) $\Rightarrow$ (3): Trivial. $\square$

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. $\square$

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 $Q_n(x) = c_nP_n(x)$. $\square$

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. $\square$

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. $\square$

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 $\mathcal{L}$. $\square$

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 nondecreasing function such that $\{x : \psi(x+\epsilon) - \psi(x-\epsilon) > 0 \text{ for all } \epsilon > 0\}$ is an infinite set. It is known [2, 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 $\mathcal{L}$. $\square$

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. □

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 $u^T Au > 0$ 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. □

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

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

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 $\{P_n(x)\}_{n\ge0}$. 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). □

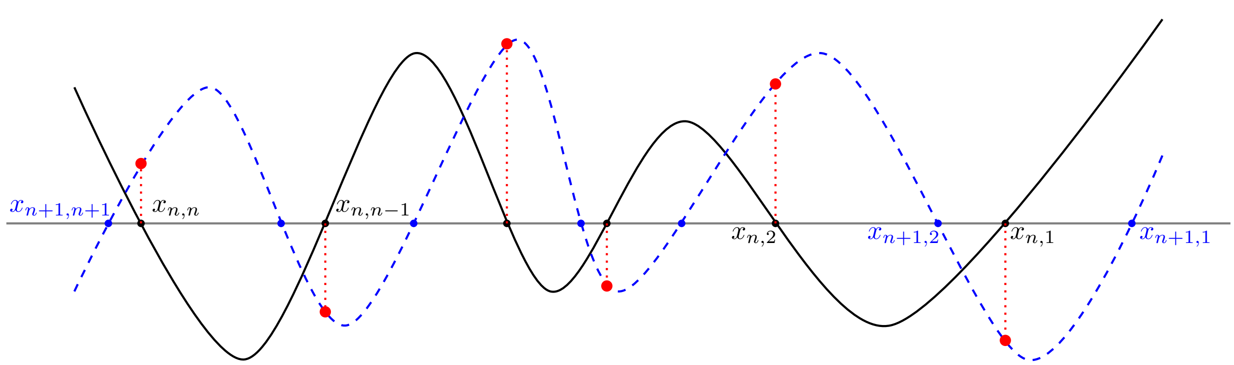

Figure 2.1: The interchanging zeros of $P_n(x)$ and $P_{n+1}(x)$.

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 [10] 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

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

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 $\boldsymbol{x} = (x_1, x_2, \dots)$ 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 $\boldsymbol{x}^I = x_1^{i_1} x_2^{i_2} \cdots$. The **ring of formal power series** in $x_1, x_2, \dots$ with coefficients in R is the set

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

with addition

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

and multiplication

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

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

$$a_0 + a_1x + a_2x^2 + \cdots .$$

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, $\sum_{n\geq 0} n!x^n$.

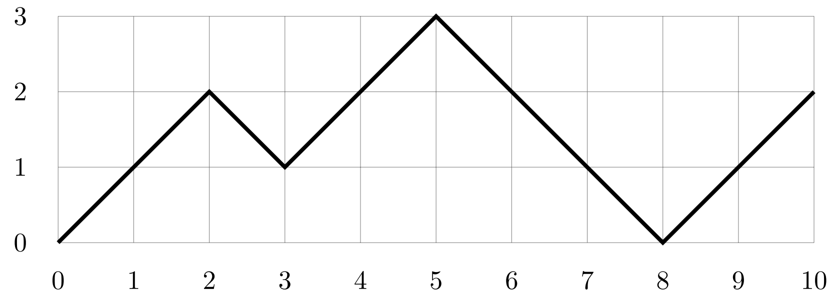

Figure 3.1: A Dyck path from $(0,0)$ to $(10,2)$.

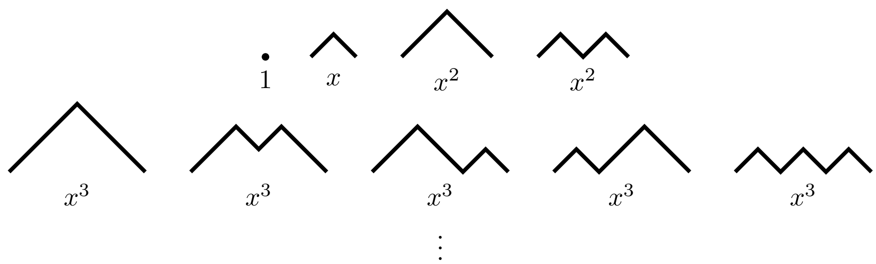

Figure 3.2: An illustration of the generating function for Dyck paths.

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

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

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 $n \ge 0$ 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,

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

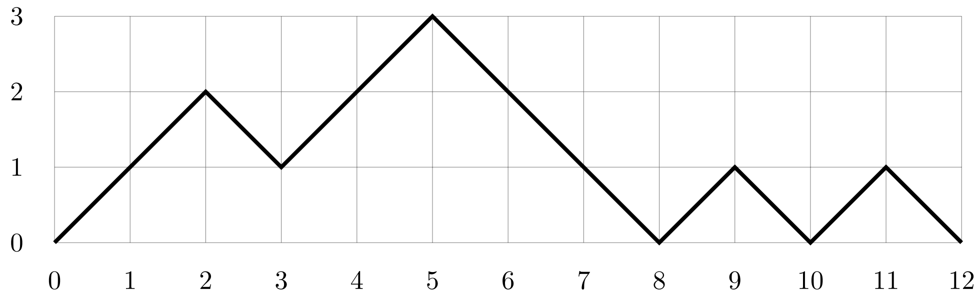

Figure 3.3: A Dyck path π from $(0, 0)$ to $(12, 0)$.

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

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

Now we can use the **binomial theorem**

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

where

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

By the binomial theorem, we have

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

Therefore,

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

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

Proposition 3.2.3. *We have*

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

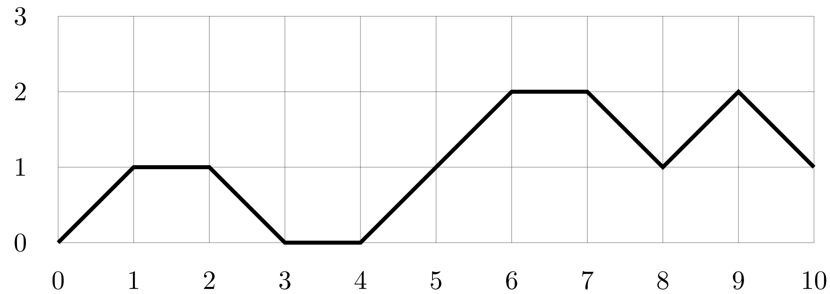

Figure 3.4: A Motzkin path from $(0, 0)$ to $(10, 1)$.

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 [8] 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. $\square$

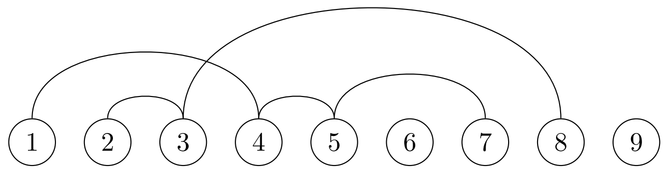

Figure 3.5: A visualization of a set partition $\{\{1,4,5,7\},\{2,3,8\},\{6\},\{9\}\}$ of $[9]$.

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 Charlier 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) $B_i \neq \emptyset$ for all i,

(2) $B_i \cap B_j = \emptyset$ for all $i \neq j$, 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 $n, k \geq 1$, 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. □

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. □

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 $2k$. □

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. $\square$

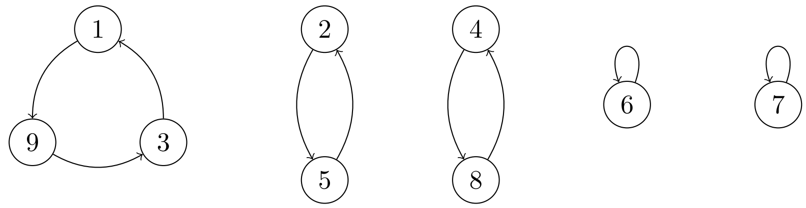

Figure 3.6: A visualization of a permutation $\pi = (1, 9, 3)(2, 5)(4, 8)(6)(7) \in \mathfrak{S}_9$.

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, $\operatorname{cycle}(\pi) = 5$. 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 $\operatorname{inv}(\pi)$ the number of inversions of π.

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

Proposition 3.4.8. *We have*

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

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

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

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 Stirling number of the first kind** $c(n,k)$ is defined to be the number of permutations on $[n]$ with k cycles. The **Stirling 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. ☐

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). ☐

Chapter 4

Combinatorial models for OPS

From now one we will focus on the combinatorial approaches to orthogonal polynomials in Viennot's lecture notes [9]. 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 $\{P_n(x)\}_{n\ge 0}$ 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) $\mathcal{L}(P_m(x)P_n(x)) = 0$ for $m \ne n$,

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

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

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 $\{P_n(x)\}_{n\ge 0}$ is an OPS for $\mathcal{L}$.*

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

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

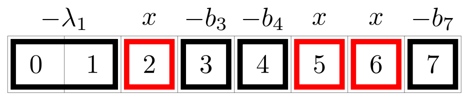

Figure 4.1: A Favard tiling $T \in \mathrm{FT}_8$ with $\mathrm{wt}(T) = \lambda_1 b_3 b_4 b_7 x^3$.

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 $\mathcal{L}(P_n(x)) = 0$. □

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 polynomials satisfying $P_{-1}(x) = 0$, $P_0(x) = 1$, and*

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

for some sequences $\{b_n\}_{n\geq 0}$ and $\{\lambda_n\}_{n\geq 1}$ with $\lambda_n \neq 0$. Then $\{P_n(x)\}_{n\geq 0}$ is an OPS for some linear functional $\mathcal{L}$.

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 $0, 1, \ldots, n-1$ 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 $n = 0, 1, 2, \dots$, then we obtain the following sequence:

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

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

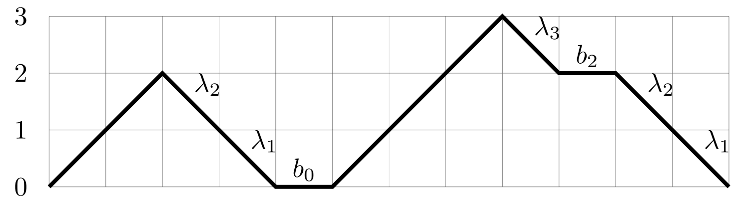

Figure 4.2: A Motzkin path π from $(0,0)$ to $(12,0)$ with $\mathrm{wt}(\pi) = b_0 b_2 \lambda_1^2 \lambda_2^2 \lambda_3$.

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

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

where the second to last equation follows from the induction hypothesis and the last equation follows from considering the first step of each $\pi \in \mathrm{Motz}_{n,r,s}$. Hence the theorem holds for n and we are done by induction. □

Corollary 4.3.4. *Following the notation in Theorem 4.3.1, we have*

$$\mathcal{L}(P_n(x)^2) = \lambda_1 \cdots \lambda_n.$$

Proof. Since $P_n(x)$ is monic, we can write $P_n(x) = x^n + Q(x)$ for some polynomial $Q(x)$ of degree less than n. Thus by Theorem 4.3.3,

$$\mathcal{L}(P_n(x)^2) = \mathcal{L}((x^n + Q(x))P_n(x)) = \mathcal{L}(x^n P_n(x)) = \sum_{\pi \in \mathrm{Motz}_{n,n,0}} \mathrm{wt}(\pi) = \lambda_1 \cdots \lambda_n,$$

as desired. Here, the last equality follows from the fact that there is only one Motzkin path in $\mathrm{Motz}_{n,n,0}$, namely, the path from $(0, n)$ to $(n, 0)$ consisting of n down steps. □

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 red monominos, let $\mathrm{wt}'(T) = \mathrm{wt}(T)/x^k$. In other words, $\mathrm{wt}'(T)$ is the same as $\mathrm{wt}(T)$ except we only consider the weights of black monominos and black dominos. Then Proposition 4.2.2 can be restated as

$$P_n(x) = \sum_{T \in \mathrm{FT}_n} \mathrm{wt}'(T) \cdot x^{(\text{number of red monominos in } T)}.$$

Thus, by the definition of $\mathcal{L}(x^n)$ in (4.4.1), 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 for some $0 \le i \le r$ and $0 \le j \le s$,

(1) $T_1 \in \mathrm{FT}_r$ has i red monominos,

(2) $T_2 \in \mathrm{FT}_s$ has j red monominos, 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 $\sum_{\pi \in Y} \mathrm{wt}(\pi)$. Therefore (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 weight-preserving involution on X with fixed point set $\{(\emptyset, \emptyset, \pi) : \pi \in Y\}$. Consider $(T_1, T_2, \pi) \in X$. We write $\pi = S_1 S_2 \cdots S_m$ as a sequence of steps. Let a, b, u, v be the integers defined as follows:

- a is the largest integer such that T_1 starts with a red monominos,
- b is the largest integer such that T_2 starts with b red 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 now define $\phi(T_1, T_2, \pi) = (T_1', T_2', \pi')$ in the following way. Here, a', b', u', v' are the integers defined similarly as above using T_1', T_2', and π'.

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 red monomino (at position u) by a black monomino. Here the notation $\widehat{S_{u+1}}$ means that S_{u+1} is removed from the sequence. See Figure 4.3. Observe that since $\mathrm{wt}'(T_1') = -b_u \,\mathrm{wt}'(T_1)$ and $\mathrm{wt}'(\pi') = \mathrm{wt}(\pi)/b_u$, we have

$$\mathrm{wt}'(T_1')\,\mathrm{wt}'(T_2')\,\mathrm{wt}(\pi') = -\,\mathrm{wt}'(T_1)\,\mathrm{wt}'(T_2)\,\mathrm{wt}(\pi).$$

Moreover, we always have $u' \ge u$ and $a' = u < a \le r$, hence $u' \ge a' \ne r$.

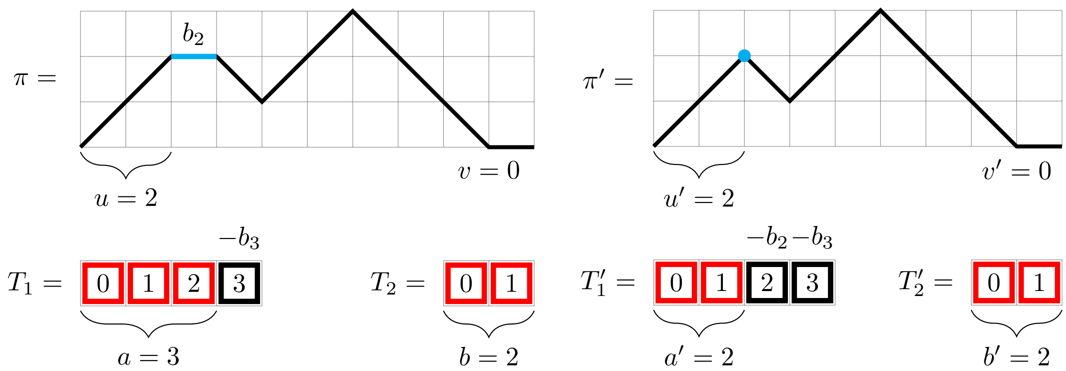

Figure 4.3: A triple $(\pi, T_1, T_2) \in X$ in Case 1-1 on the left and the corresponding triple $(\pi', T_1', T_2') \in X$ in Case 2-1 on the right, where $r = 4$, $s = 2$, and $n = 5$. The horizontal step $(2, 2) \to (3, 3)$ in π is collapsed to a point.

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 red monominos (at positions $u - 1$ and u) by a domino. See Figure 4.4. Observe that since $\mathrm{wt}'(T_1') = -\lambda_u \, \mathrm{wt}'(T_1)$ and $\mathrm{wt}'(\pi') = \mathrm{wt}'(\pi)/\lambda_u$, we have

$$\mathrm{wt}'(T_1') \, \mathrm{wt}'(T_2') \, \mathrm{wt}(\pi') = - \, \mathrm{wt}'(T_1) \, \mathrm{wt}'(T_2) \, \mathrm{wt}(\pi).$$

Moreover, we always have $u' \geq u - 1$ and $a' = u - 1 < a \leq r$, hence $u' \geq a' \neq r$.

Case 2 $u \geq a \neq r$. 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 black monomino. In this case let

$$\pi' = S_1 \cdots S_a H S_{a+1} \cdots S_m,$$

and define T_1' to be the Favard tiling obtained from T_1 by replacing A by a red monomino. See Figure 4.3 (with the roles of (T_1, T_2, π) and (T_1', T_2', π') interchanged).

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

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

and define T_1' to be the Favard tiling obtained from T_1 by replacing A by two red monominos. See Figure 4.4 (with the roles of (T_1, T_2, π) and (T_1', T_2', π') interchanged).

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 $(T_1', T_2', \pi') = (T_1, T_2, \pi)$. See Figure 4.5.

By the construction, Case 1 corresponds to Case 2 and Case 3 corresponds to Case 4. Thus the map $\phi(T_1, T_2, \pi) = (T_1', T_2', \pi')$ is a sign-reversing weight-preserving involution on X with fixed points $(\emptyset, \emptyset, \pi)$ where $\pi \in Y$. This completes the proof. □

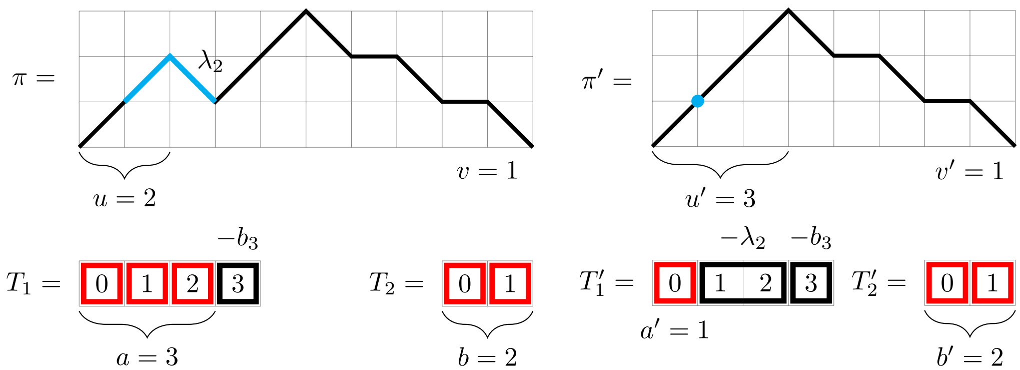

Figure 4.4: A triple $(\pi, T_1, T_2) \in X$ in Case 1-2 on the left and the corresponding triple $(\pi', T'_1, T'_2) \in X$ in Case 2-2 on the right, where $r = 4$, $s = 2$, and $n = 5$. The peak (an up-step followed by a down step) $(1, 1) \to (2, 2) \to (3, 1)$ in π is collapsed to a point.

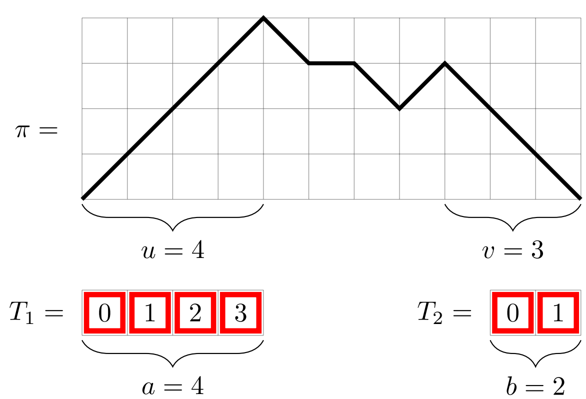

Figure 4.5: A triple $(\pi, T_1, T_2) \in X$ in Case 5, where $r = 4$, $s = 2$, and $n = 5$.

Chapter 5

Moments of classical orthogonal polynomials

In this chapter we apply the combinatorial interpretation of moments for Tchebyshev polynomials of the 1st and 2nd kinds, Hermite polynomials, Charlier polynomials, and Laguerre polynomials.

Note that a monic 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 $\{a_{n,k}\}_{n,k\geq 0}$ of $P_n(x) = \sum_{k=0}^{n} a_{n,k}x^k$,

(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 $\{\lambda_n\}_{n\geq 1}$.

5.1 Tchebyshev polynomials

In this section we will compute the moments of Tchebyshev polynomials using Theorem 4.3.1. We will first consider Tchebyshev polynomials of the second kind since they are simpler than the first kind in our approach.

The **Tchebyshev polynomials of the second kind** $U_n(x)$ are defined by

$$U_n(x) = \frac{\sin(n+1)\theta}{\sin\theta}, \qquad x = \cos\theta, \qquad n \geq 0.$$

They satisfy

$$U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x), \qquad n \geq 0,$$

where $U_{-1}(x) = 0$ and $U_0(x) = 1$. Using calculus we can prove that

$$\int_{-1}^{1} U_m(x)U_n(x)(1-x^2)^{1/2}dx = \frac{\pi}{2}\delta_{m,n}.$$

Let $\mathcal{L}$ be the linear functional defined by

$$\mathcal{L}(f(x)) = \frac{2}{\pi}\int_{-1}^{1} f(x)(1-x^2)^{1/2}dx.$$

Then $\{U_n(x)\}_{n\geq 0}$ is an OPS for $\mathcal{L}$ and $\mathcal{L}(1) = 1$.

Since $U_n(x)$ has leading coefficient 2^n, the monic Tchebyshev polynomials $\hat{U}_n(x)$ are given by $\hat{U}_n(x) = 2^{-n}U_n(x)$ and

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

where $b_n = 0$ and $\lambda_n = 1/4$.

Note that $\{\hat{U}_n(x)\}_{n\geq 0}$ is also an OPS for $\mathcal{L}$. Using calculus we can prove that the moments

$$\mu_n = \mathcal{L}(x^n) = \frac{2}{\pi}\int_{-1}^{1} x^n(1-x^2)^{1/2}dx$$

are given by

$$\mu_{2n} = \frac{1}{4^n}C_n, \qquad \mu_{2n+1} = 0. \tag{5.1.1}$$

We will prove this combinatorially using the combinatorial interpretation for μ_n.

By Theorem 4.3.1,

$$\mu_n = \sum_{\pi\in \mathrm{Motz}_n} \mathrm{wt}(\pi) = \sum_{\pi\in\mathrm{Dyck}_n} \left(\frac{1}{4}\right)^{n/2}.$$

Thus

$$\mu_{2n} = \frac{1}{4^n}|\,\mathrm{Dyck}_{2n}\,|, \qquad \mu_{2n+1} = 0.$$

This is the same as (5.1.1).

Now we consider the Tchebyshev polynomials of the first kind, $T_n(x) = \cos n\theta$, $x = \cos\theta$. Recall that

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

Let $\mathcal{L}$ be the linear functional defined by

$$\mathcal{L}(f(x)) = \frac{1}{\pi}\int_{-1}^{1} f(x)(1-x^2)^{-1/2}dx.$$

Then $\{T_n(x)\}_{n\geq 0}$ is an OPS for $\mathcal{L}$ and $\mathcal{L}(1) = 1$. The moments

$$\mu_n = \mathcal{L}(x^n) = \frac{1}{\pi}\int_{-1}^{1} x^n(1-x^2)^{-1/2}dx$$

are given by

$$\mu_{2n} = \frac{1}{2^{2n}}\binom{2n}{n}, \qquad \mu_{2n+1} = 0. \tag{5.1.2}$$

We will prove this combinatorially.

The monic Tchebyshev polynomials of the first kind are given by $\hat{T}_0(x) = 1$ and $\hat{T}_n(x) = 2^{1-n}T_n(x)$ for $n \geq 1$. We have

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

where $b_n = 0$ for $n \geq 0$, $\lambda_1 = 1/2$, and $\lambda_n = 1/4$ for $n \geq 2$.

By Theorem 4.3.1,

$$\mu_n = \sum_{\pi\in\mathrm{Motz}_n} \mathrm{wt}(\pi) = \left(\frac{1}{4}\right)^{n/2} \sum_{\pi\in\mathrm{Dyck}_n} 2^{a(\pi)},$$

where $a(\pi)$ is the number of down steps in π touching the x-axis. Thus (5.1.2) is a consequence of the following proposition.

Proposition 5.1.1. *We have*

$$\sum_{\pi\in\mathrm{Dyck}_{2n}} 2^{a(\pi)} = \binom{2n}{n}. \tag{5.1.3}$$

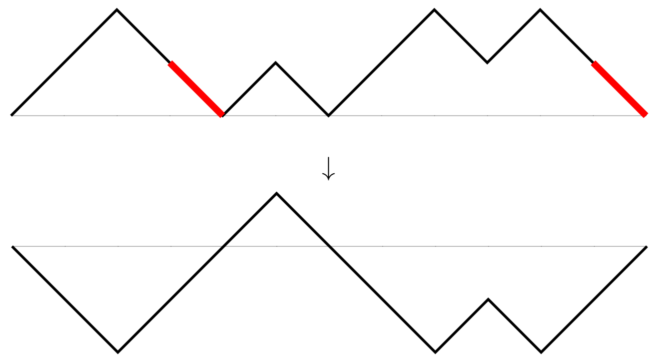

Figure 5.1: The map reflecting each path with a red down step.

Proof. Define a **colored Dyck path** to be a Dyck path in Dyck_{2n} such that every down step touching the x-axis is colored red or black. The left-hand side of the equation is the number of colored Dyck paths in Dyck_{2n}. Thus it suffices to show that this number is equal to $\binom{2n}{n}$.

Note that every Dyck path $\pi \in \mathrm{Dyck}_{2n}$ is decomposed into

$$\pi = (U\pi_1 D)(U\pi_2 D)\cdots(U\pi_k D),$$

where $\pi_i \in \mathrm{Dyck}_{2t_i}$ for some $t_i \in \mathbb{Z}_{\geq 0}$. Each down step D after π_i touches the x-axis and no other down steps have this property. Thus a colored Dyck path can be identified with a sequence of the form

$$\pi = (U\pi_1 D_1)(U\pi_2 D_2)\cdots(U\pi_k D_k),$$

where each D_i is colored red or black. If D_i is colored red, reflect the subpath $U\pi_i D$ about the x-axis. Then we get a path from $(0,0)$ to $(2n,0)$ consisting of up steps and down steps (which may go below the x-axis).

This map gives a bijection between the colored Dyck paths from $(0,0)$ to $(2n,0)$ to any path from $(0,0)$ to $(2n,0)$ consisting of up steps and down steps. Since there are $\binom{2n}{n}$ such paths, we obtain the result. $\square$

5.2 Hermite polynomials

The **Hermite polynomials** $H_n(x)$ are defined by $H_{-1}(x) = 0$, $H_0(x) = 1$, and

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x), \qquad n \geq 1.$$

Since the leading coefficient of $H_n(x)$ is 2^n, we can make it monic by letting $\hat{H}_n(x) = 2^{-n}H_n(x)$. Then $\hat{H}_{-1}(x) = 0$, $\hat{H}_0(x) = 1$, and

$$\hat{H}_{n+1}(x) = x\hat{H}_n(x) - \frac{n}{2}\hat{H}_{n-1}(x), \qquad n \geq 1. \tag{5.2.1}$$

For the combinatorial study of orthogonal polynomials, it is more convenient if the recurrence coefficients b_n and λ_n are integers. We can rescale orthogonal polynomials using the following lemma.

Lemma 5.2.1 (Rescaling OPS). *Suppose that $\{P_n(x)\}_{n\geq 0}$ is a monic OPS such that $P_{-1}(x) = 0$, $P_0(x) = 1$, and*

$$P_{n+1}(x) = (x - b_n)P_n(x) - \lambda_n P_{n-1}(x), \qquad n \geq 1. \tag{5.2.2}$$

Let $\widetilde{P}_n(x) = a^n P_n(x/a)$, where $a \neq 0$. Then $\widetilde{P}_{-1}(x) = 0$, $\widetilde{P}_0(x) = 1$, and

$$\widetilde{P}_{n+1}(x) = (x - ab_n)\widetilde{P}_n(x) - a^2\lambda_n\widetilde{P}_{n-1}(x), \qquad n \geq 1. \tag{5.2.3}$$

Proof. Replacing x by x/a and multiplying both sides by a^{n+1} in (5.2.2) yields (5.2.3). $\square$

Define the **rescaled Hermite polynomials** $\widetilde{H}_n(x)$ by $\widetilde{H}_n(x) = \sqrt{2}^n \hat{H}_n(x/\sqrt{2})$. By Lemma 5.2.1 and (5.2.1), $\widetilde{H}_{-1}(x) = 0$, $\widetilde{H}_1(x) = 1$, and

$$\widetilde{H}_{n+1}(x) = x\widetilde{H}_n(x) - n\widetilde{H}_{n-1}(x), \qquad n \geq 1. \tag{5.2.4}$$

We note that $H_n(x)$ are called "physicist's Hermite polynomials" and $\widetilde{H}_n(x)$ are called "probabilist's Hermite polynomials".

The moment μ_n of $\{\widetilde{H}_n(x)\}_{n\geq 0}$ is given by

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

where $\mathrm{wt}(\pi)$ is determined by $b_n = 0$ and $\lambda_n = n$. Since $b_n = 0$, we have $\mu_{2n+1} = 0$ and

$$\mu_{2n} = \sum_{\pi \in \mathrm{Dyck}_{2n}} \mathrm{wt}(\pi).$$

Note that for each $\pi \in \mathrm{Dyck}_{2n}$, its weight $\mathrm{wt}(\pi)$ is a positive integer. It is thus natural to ask what combinatorial objects $\mathrm{wt}(\pi)$ counts.

Definition 5.2.2. A **Hermite history** is a Dyck path where each down step starting at height k has a label in $\{1, \ldots, k\}$. Let HH_{2n} denote the set of Hermite histories whose underlying Dyck paths are from $(0, 0)$ to $(2n, 0)$.

Let $\pi \in \mathrm{Dyck}_{2n}$. For each down step of π starting at height k, there are k ways to assign a label from $\{1, \ldots, k\}$. Thus $\mathrm{wt}(\pi)$ is the number of Hermite histories with underlying Dyck path π. This implies $\mu_{2n} = |\,\mathrm{HH}_{2n}\,|$.

Let CM_{2n} be the set of complete matchings on $[2n]$.

Proposition 5.2.3. *There is a bijection between* HH_{2n} *and* CM_{2n}*.*

Proof. Let $\pi \in \mathrm{HH}_{2n}$. We construct a complete matching ρ as follows. For $k = 1, \ldots, 2n$, if the kth step of π is an up step, then make the kth vertex of ρ to be an **opener**, which means it will be connected to a vertex to its right. If the kth step of π is a down step, then make the kth vertex of ρ to be a **closer**, which means it will be connected to a vertex to its left. If the kth step of π is a down step with label a_k, then connected the vertex at k with the kth closest available opener. For example, see Figure 5.2.

Observe that the height of the starting point of the kth down step is equal to the number of available openers for the vertex k. Therefore the map $\pi \mapsto \rho$ is well-defined. The inverse map $\rho \mapsto \pi$ is straighforward to construct. Hence the map $\pi \mapsto \rho$ is a bijection from HH_{2n} and CM_{2n}. $\square$

Since the number of complete matchings on $[2n]$ is $(2n-1)!!$, we obtain the following result.

Corollary 5.2.4. *The* $2n$*th moment of rescaled Hermite polynomials* $\widetilde{H}_n(x)$ *is*

$$\mu_{2n} = (2n-1)!!.$$

5.3 Charlier polynomials

The (normalized) **Charlier polynomials** $C_n(x; a)$ are defined by $C_{-1}(x; a) = 0$, $C_1(x; a) = 1$, and

$$C_{n+1}(x; a) = (x - n - a)C_n(x; a) - anC_{n-1}(x; a), \qquad n \geq 1.$$

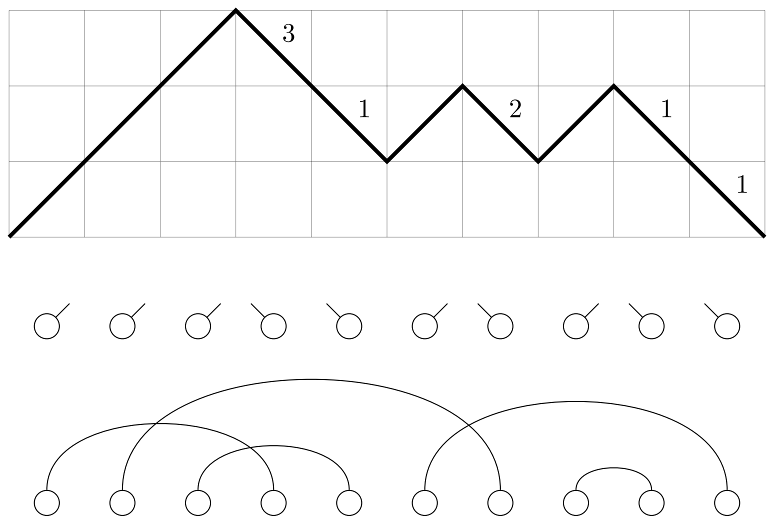

Figure 5.2: A Hermite history (top), the openers and closers (middle), and the corresponding matching (bottom).

Definition 5.3.1. A **Charlier history** is a Motzkin path where each horizontal step at height k has a label in $\{0, 1, \ldots, k\}$ and each down step starting at height k has a label in $\{1, \ldots, k\}$. Let CH_n denote the set of Charlier histories whose underlying Motzkin paths are from $(0, 0)$ to $(n, 0)$. For $\rho \in \mathrm{CH}_n$, define $t(\rho)$ to the number of horizontal steps with label 0 plus the number of down steps.

Since $b_k = k + a$ and $\lambda_k = ak$, by the definition of the Charlier histories, the moment μ_n of the Charlier polynomials is given by

$$\mu_n = \sum_{\pi \in \mathrm{Motz}_n} \mathrm{wt}(\pi) = \sum_{\rho \in \mathrm{CH}_n} a^{t(\rho)}.$$

For a set partition σ, let $\mathrm{block}(\sigma)$ denote the number of blocks in σ.

Theorem 5.3.2. *The moment μ_n of the Charlier polynomials is given by*

$$\mu_n = \sum_{\sigma \in \Pi_n} a^{\mathrm{block}(\sigma)}.$$

Proof. We will construct a weight-preserving bijection $\phi : \mathrm{CH}_n \to \Pi_n$. Let $\rho \in \mathrm{CH}_n$. We construct the corresponding set partition $\phi(\rho) = \sigma \in \Pi_n$ as follows. For $k = 1, \ldots, n$,

- if the kth step of ρ is an up step, then make the kth vertex of σ to be an opener,
- if the kth step of ρ is a down step, then make the kth vertex of ρ to be a closer,
- if the kth step of ρ is a horizontal step, then make the kth vertex of ρ to be a **transient**, which means that it is connected to a vertex to its left and also a vertex to its right.

If the kth step of ρ is a down step with label a_k, then connected the vertex at k with the kth closest available opener or transient. If the kth step of ρ is a horizontal step with label a_k, then connected the vertex at k with the kth closest available opener or transient. Here, if $a_k = 0$, we the vertex k is connected to itself making it a singleton. For example, see Figure 5.3.

It is not hard to see that $\phi : \mathrm{CH}_n \to \Pi_n$ is a bijection. Moreover, if $\phi(\rho) = \sigma$, then $t(\rho) = \mathrm{block}(\sigma)$. This can be seen from the observation that every block of ρ is either a singleton (corresponding to a horizontal step with label 0) or a block with exactly one closer (corresponding to a down step). This completes the proof. $\square$

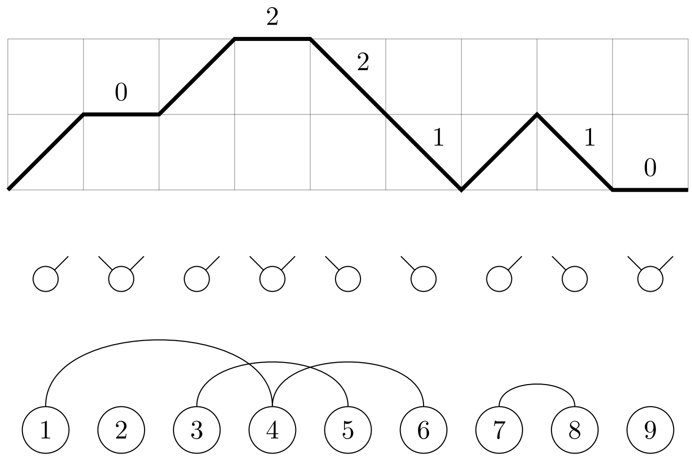

Figure 5.3: A Charlier history (top), the openers, closers, and transients (middle), and the corresponding set partition (bottom).

5.4 Laguerre polynomials

The (normalized) **Laguerre polynomials** $L_n^{(\alpha)}(x)$ are defined by[1] $L_{-1}^{(\alpha)}(x) = 0$, $L_1^{(\alpha)}(x) = 1$, and

$$L_{n+1}^{(\alpha)}(x) = (x - 2n - \alpha)L_n^{(\alpha)}(x) - n(n - 1 + \alpha)L_{n-1}^{(\alpha)}(x), \qquad n \geq 1.$$

Lemma 5.4.1. *Suppose that* $\{P_n(x)\}_{n\geq 0}$ *is a monic OPS such that*

$$P_{n+1}(x) = (x - b_n)P_n(x) - a_{n-1}c_nP_{n-1}(x).$$

Then

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

where $\mathrm{wt}'(\pi)$ *is the product of the weights of the steps in* π *and*

- *the weight of an up step starting at height* k *is* a_k*,*
- *the weight of a horizontal step at height* k *is* b_k*,*
- *the weight of a down step starting at height* k *is* c_k*.*

Proof. We know that

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

where $\mathrm{wt}(\pi)$ is the product of b_k for each horizontal step of height k and $\lambda_k = a_kc_k$ for each down step starting at height k. Observe that for $\pi \in \mathrm{Motz}_n$ every down step corresponds to a unique up step. More precisely, if we write π as a sequence of steps $S_1 \cdots S_n$ and if $S_i = D$, then there is a unique index j such that $S_j = U$ and $S_{j+1} \cdots S_{i-1}$ is a (translated) Dyck path in Dyck_{i-j-1}.

Thus we can split the weight $\lambda_k = a_kc_k$ on a down step starting at height k into the weight a_k of the corresponding up step and the weight c_k of the down step as shown in Figure 5.4. This proves the lemma. ☐

By Lemma 5.4.1, the moment of the Laguerre polynomials is given by

$$\mu_n = \sum_{\pi\in \mathrm{Motz}_n} \mathrm{wt}'(\pi), \tag{5.4.1}$$

where $a_k = k + \alpha$, $b_k = 2k + \alpha$, and $c_k = k$.

[1] In the literature it is more common to define the Laguerre polynomials with α replaced by $\alpha+1$ in our definition.

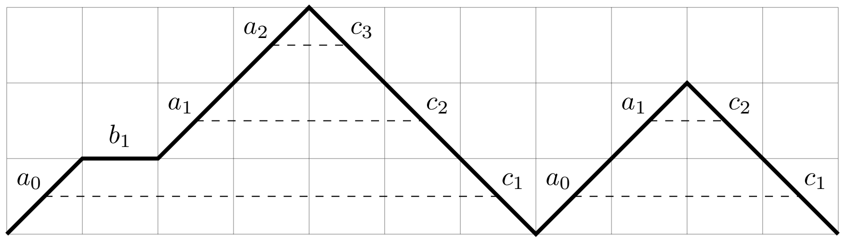

Figure 5.4: Splitting the weight $\lambda_k = a_{k-1}c_k$ into a_{k-1} and c_k.

Definition 5.4.2. A **Laguerre history** is a Motzkin path where

- each up step starting at height k has a label in $\{0, 1, \dots, k\}$,
- each horizontal step at height k has a label in $\{-k, \dots, -1, 0, 1, \dots, k\}$,
- each down step starting at height k has a label in $\{1, 2, \dots, k\}$,

Let LH_n denote the set of Laguerre histories whose underlying Motzkin paths are from $(0,0)$ to $(n, 0)$. For $\rho \in \mathrm{LH}_n$, define $\mathrm{zero}(\rho)$ to be the number of labels equal to 0 in ρ.

Then (5.4.1) is equivalent to

$$\mu_n = \sum_{\rho \in \mathrm{LH}_n} \alpha^{\mathrm{zero}(\rho)}.$$

There are several bijections between $\mathfrak{S}_n$ and LH_n due to Françon–Viennot [5], Foata–Zeilberger [4], see also [3, Algorithm 7]. To prove the following theorem we use a slight modification of the bijection due to Françon–Viennot [5].

Theorem 5.4.3. *We have*

$$\mu_n = \sum_{\pi \in \mathfrak{S}_n} \alpha^{\mathrm{cycle}(\pi)}.$$

Proof. It suffices to find a bijection $\phi : \mathrm{LH}_n \to \mathfrak{S}_n$ such that if $\phi(\rho) = \pi$, then $\mathrm{zero}(\rho) = \mathrm{cycle}(\pi)$. Consider $\pi \in \mathrm{LH}_n$ and let S_k be the kth step of π and let ℓ_k be its label.

For each $k = 0, 1, \dots, n$, we will construct a list A_k of cycles of integers and dots, $\bullet$'s. First we set $A_0 = \emptyset$. We then construct A_k recursively as follows.

Case 1: S_k is an up step.

Case 1-1: $\ell_k = 0$. Create a new cycle "$(k\bullet)$" at the beginning:

$$A_k = (k\bullet)A_{k-1}.$$

Case 1-2: $\ell_k = i > 0$. Replace the ith dot in A_{k-1} by "$\bullet k \bullet$".

Case 2: S_k is a horizontal step.

Case 2-1: $\ell_k = 0$. Create a new cycle "(k)" at the beginning:

$$A_k = (k)A_{k-1}.$$

Case 2-2: $\ell_k = i > 0$. Replace the ith dot in A_{k-1} by "$k\bullet$".

Case 2-3: $\ell_k = -i < 0$. Replace the ith dot in A_{k-1} by "$\bullet k$".

Case 3: S_k is a down step. Then $\ell_k = i > 0$. Replace the ith dot in A_{k-1} by "k".

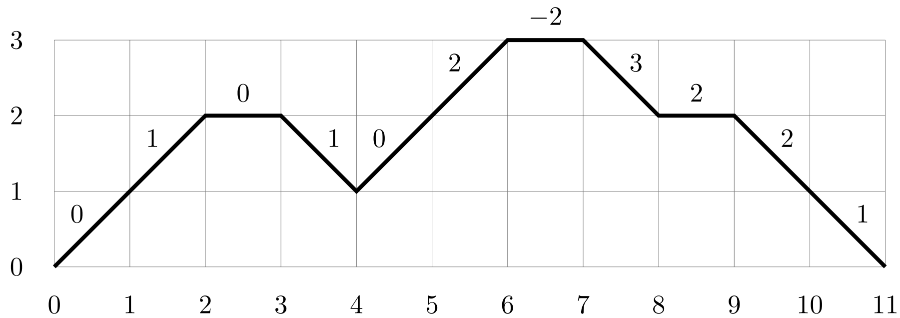

Figure 5.5: A Laguerre history in LH_{11}.

Then we define $\phi(\rho)$ to the permutation π whose cycle decomposition is A_n.

For example, if ρ is the Laguerre history in Figure 5.5, then

$$\begin{aligned}
A_0 &= \emptyset,\\
A_1 &= (1, \bullet),\\
A_2 &= (1, \bullet, 2, \bullet),\\
A_3 &= (3), (1, \bullet, 2, \bullet),\\
A_4 &= (3), (1, 4, 2, \bullet),\\
A_5 &= (5, \bullet), (3), (1, 4, 2, \bullet),\\
A_6 &= (5, \bullet), (3), (1, 4, 2, \bullet, 6, \bullet),\\
A_7 &= (5, \bullet), (3), (1, 4, 2, \bullet, 7, 6, \bullet),\\
A_8 &= (5, \bullet), (3), (1, 4, 2, \bullet, 7, 6, 8),\\
A_9 &= (5, \bullet), (3), (1, 4, 2, 9, \bullet, 7, 6, 8),\\
A_{10} &= (5, \bullet), (3), (1, 4, 2, 9, 10, 7, 6, 8),\\
A_{11} &= (5, 11), (3), (1, 4, 2, 9, 10, 7, 6, 8).
\end{aligned}$$

The number of dots in A_{k-1} is always equal to the ending height, say h_{k-1}, of $S_1 \cdots S_{k-1}$. Since h_{k-1} is the height of the starting point of S_k, we have $|\ell_k| \le h_{k-1}$. Hence, if $|\ell_k| = i \ne 0$, we can always find the ith dot in A_{k-1} and the above construction is well defined.

We create a cycle if and only if $\ell_k = 0$. Thus, $\mathrm{zero}(\rho) = \mathrm{cycle}(\pi)$ as desired.

To prove that the map $\phi : \mathrm{LH}_n \to \mathfrak{S}_n$ is invertible, we construct its inverse map. Let $\pi \in \mathfrak{S}_n$. Then we write the cycles of $\mathfrak{S}_n$ so that each cycle starts with its smallest element and the cycles are listed in the decreasing order of their smallest elements. Let A_n be the list of cycles obtained in this way. For $k = n, n-1, \dots, 1, 0$, we define A_{k-1} to be the configuration obtained from A_k by replacing k together with every dot "$\bullet$" adjacent to it by a single dot "$\bullet$". If k forms a cycle of length 1, then we delete the whole cycle "(k)". Once the sequence $A_0, A_1, \dots, A_n$ is constructed, we can define the corresponding Laguerre history ρ whose kth step is S_k with label ℓ_k as follows.

For each $k = 1, \dots, n$, we compare A_k with A_{k-1}.

Case 1: $A_k = (k\bullet)A_{k-1}$. Then define $S_k = U$ and $\ell_k = 0$.

Case 2: A_k is obtained from A_{k-1} by replacing the ith dot by "$\bullet k \bullet$". Then $S_k = U$ and $\ell_k = i$.

Case 3: $A_k = (k)A_{k-1}$. Then define $S_k = H$ and $\ell_k = 0$.

Case 4: A_k is obtained from A_{k-1} by replacing the ith dot by "$k\bullet$". Then $S_k = D$ and $\ell_k = i$.

Case 5: A_k is obtained from A_{k-1} by replacing the ith dot by "$\bullet k$". Then $S_k = D$ and $\ell_k = -i$.

Then we define $\psi(\pi)$ to be the resulting Laguerre history ρ.

By the construction, ψ is the inverse map of ϕ, and the proof is completed. □

By Theorem 5.4.3 and (3.4.4), we obtain the simple formula for μ_n.

Corollary 5.4.4. *The nth moment of Laguerre polynomials is*

$$\mu_n = \alpha(\alpha+1)\cdots(\alpha+n-1).$$

Note that the bijection $\phi : \mathrm{LH}_n \to \mathfrak{S}_n$ induces two bijections $\phi_1 : \mathrm{CH}_n \to \Pi_n$ and $\phi_2 : \mathrm{HH}_n \to \mathrm{CM}_n$.

To see this, observe that a Laguerre history becomes a Charlier history if every up step has the zero label, and every horizontal step has a nonnegative label. Then in the corresponding list A_n of cycles, every cycle is an increasing list of integers. Hence the cycles can be identified with blocks giving a set partition.

Similalry, a Laguerre history becomes a Hermite history if every up step has the zero label, and there is no horizontal step. Then in the corresponding list A_n of cycles, every cycle is a pair (i, j) of integers $i < j$. Hence the cycles can be identified with arcs giving a complete matching.

At this point the reader may wonder if there is another bijection between Laguerre histories and permutations similar to the bijections in the previous sections using arcs. Indeed, there is such a bijection due to Foata and Zeilberger [4]. We will briefly describe this bijection. For simplicity we consider the case $\alpha = 1$. In this map we use the usual Motzkin weight which gives a weight $b_k = 2k+1$ for a horizontal step starting at height k and a weight $\lambda_k = k^2$ for a down step starting at height k.

A **modified Laguerre history** of length n is a Motzkin path from $(0,0)$ to $(n,0)$ in which every horizontal step with starting height k is labeled by an integer in $\{-k,\dots,-1,0,1,\dots,k\}$ every down step with starting height k is labeled by a pair (i, j) of integers in $\{1,\dots,k\}$.

Let ρ be a modified Laguerre history. For $k = 1,\dots,n$, we construct a diagram on n vertices as follows.

(1) The kth vertex is an **opener**, a **closer**, a **fixed point**, an **upper transient**, or a **lower transient** if the kth step of ρ is an up step, a down step, a horizontal step with label 0, a horizontal step with label 0, a horizontal step with positive label, or a horizontal step with negative label, respectively.

(2) Using the label ℓ_k of the kth step of ρ, we connect a closer, an upper transient, or a lower transient similarly to the bijections in the previous sections.

For example, see Figure 5.6.

Then the resulting diagram represent a permutation π where $\pi(i) = j$ if $i < j$ and i is connected to j with an upper arc or $i > j$ and i is connected to j with a lower arc. If there is no arc connecting i, then $\pi(i) = i$. For example, the diagram in Figure 5.6 represent the following permutation:

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 4 & 8 & 3 & 2 & 9 & 11 & 5 & 7 & 10 & 1 & 6 \end{pmatrix}.$$

It is not hard to show that the above map $\rho \mapsto \pi$ is a bijection from the set of modified Laguerre histories of length n to $\mathfrak{S}_n$. Moreover, this map has the property $\max(\rho) = \mathrm{LRmax}(\pi)$, where $\max(\rho)$ is the number of maximum possible labels of a horizontal step plus the number of maximum possible labels in the first component of a down step and $\mathrm{LRmax}(\pi)$ is the number of left-to-right maxima of $\pi = \pi_1 \cdots \pi_n$, that is π_i such that $\pi_i = \max\{\pi_1,\dots,\pi_i\}$. This implies that

$$\mu_n = \sum_{\pi\in\mathfrak{S}_n} \alpha^{\mathrm{LRmax}(\pi)}.$$

For example, suppose ρ and π are as in Figure 5.6. Then the ith step is a horizontal step with maximum possible label or a down step with maximum possible lable in the first component for $i = 4, 8, 9, 11$. The left-to-right maxima of π are exactly $4, 8, 9, 11$.

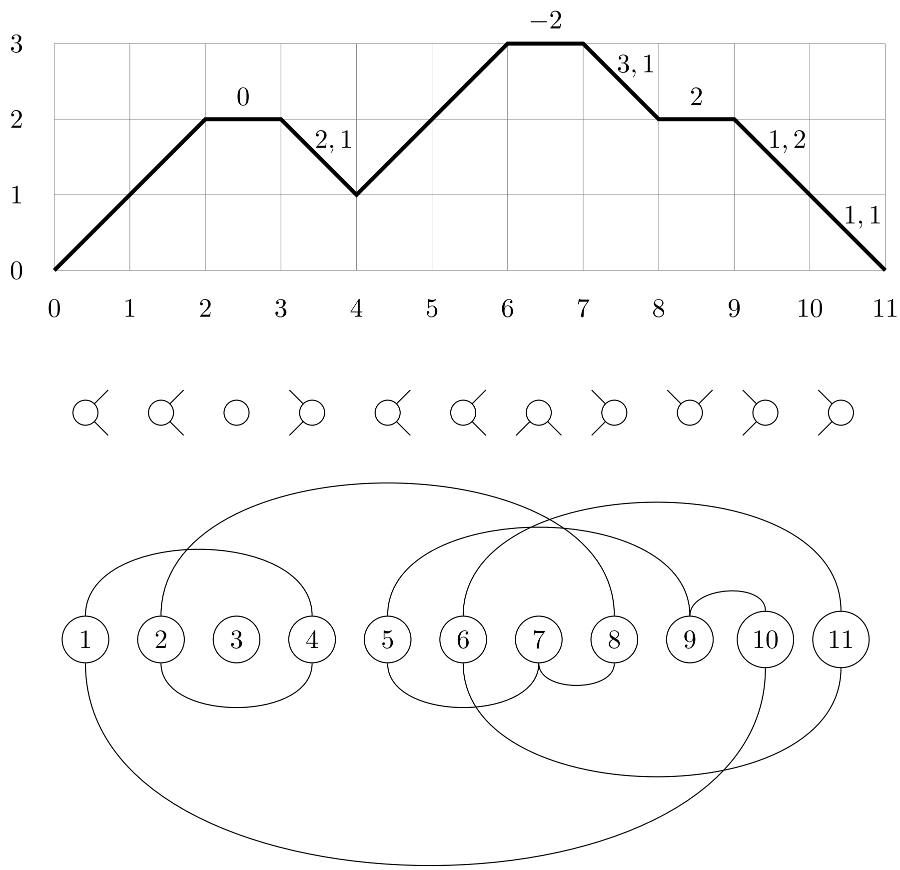

Figure 5.6: A modified Laguerre history in LH_{11} and the corresponding diagram representing a permutation.

Chapter 6

Duality between mixed moments and coefficients

Suppose that $\{P_n(x)\}_{n\geq 0}$ is a monic OPS given by

$$P_n(x) = \sum_{k=0}^{n} \nu_{n,k} x^k.$$

In this chapter we will show that

$$x^n = \sum_{k=0}^{n} \mu_{n,k} P_k(x),$$

where $\mu_{n,k} = \mathcal{L}(x^n P_k(x))/\mathcal{L}(P_k(x)^2)$ is the mixed moment. We then show the duality between the mixed moments $\mu_{n,k}$ and the coefficients $\nu_{n,k}$ combinatorially. As special cases we obtain various known dualities among binomial coefficients, q-binomial coefficients, Stirling numbers, and elementary and homogeneous symmetric functions.

6.1 Mixed moments and coefficients

As before suppose that $\{P_n(x)\}_{n\geq 0}$ is a monic OPS with a linear functional $\mathcal{L}$ given by

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

Recall from Definition 4.3.2 and Theorem 4.3.3 that the mixed moment $\mu_{n,k}$ has the following combinatorial interpretation:

$$\mu_{n,k} = \frac{\mathcal{L}(x^n P_k(x))}{\mathcal{L}(P_k(x))} = \sum_{\pi \in \mathrm{Motz}_{n,k}} \mathrm{wt}(\pi).$$

where $\mathrm{Motz}_{n,k}$ is the set of Motzkin paths from $(0,0)$ to (n,k).

Proposition 6.1.1. *We have*

$$x^n = \sum_{k=0}^{n} \mu_{n,k} P_k(x).$$

Proof. Let

$$x^n = \sum_{k=0}^{n} \sigma_{n,k} P_k(x).$$

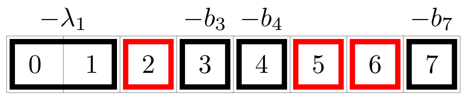

Figure 6.1: A Favard tiling $T \in \mathrm{FT}_{8,3}$ with $\mathrm{wt}'(T) = \lambda_1 b_3 b_4 b_7$.

Multiplying $P_k(x)$ and taking $\mathcal{L}$ both sides, we obtain

$$\mathcal{L}(x^n P_k(x)) = \sigma_{n,k}\mathcal{L}(P_k(x)^2).$$

Hence

$$\sigma_{n,k} = \frac{\mathcal{L}(x^n P_k(x))}{\mathcal{L}(P_k(x)^2)} = \mu_{n,k},$$

as desired. □

Now let $\nu_{n,k}$ be the coefficient of x^k in $P_n(x)$ so that

$$x^n = \sum_{k=0}^{n} \mu_{n,k} P_k(x),$$
$$P_n(x) = \sum_{k=0}^{n} \nu_{n,k} x^k.$$

Since $\{P_n(x)\}_{n\geq 0}$ and $\{x^n\}_{n\geq 0}$ are bases of the space of polynomials, we have the following matrix identities:

$$(\nu_{n,k})_{n,k\geq 0} (\mu_{n,k})_{n,k\geq 0} = (\mu_{n,k})_{n,k\geq 0} (\nu_{n,k})_{n,k\geq 0} = I.$$

Equivalently,

$$\sum_{k\geq 0} \nu_{n,k}\mu_{k,m} = \delta_{n,m}, \tag{6.1.1}$$
$$\sum_{k\geq 0} \mu_{n,k}\nu_{k,m} = \delta_{n,m}. \tag{6.1.2}$$

Since we have combinatorial interpretations for $\mu_{n,k}$ and $\nu_{n,k}$, we can prove the above matrix identities combinatorially. In fact, in the next section we will prove these combinatorially without the assumption that $\lambda_k \neq 0$.

6.2 Combinatorial proof of duality

Suppose that $\{b_n\}_{n\geq 0}$ and $\{\lambda_n\}_{n\geq 1}$ are arbitrary sequences (λ_n may be zero). Then we can take the following as definitions:

$$\mu_{n,k} = \sum_{\pi \in \mathrm{Motz}_{n,k}} \mathrm{wt}(\pi),$$
$$\nu_{n,k} = \sum_{T \in \mathrm{FT}_{n,k}} \mathrm{wt}'(T),$$

where $\mathrm{FT}_{n,k}$ is the set of Favard tilings in FT_n with k red monominos, see Figure 6.1. Observe that $\mu_{n,k} = \nu_{n,k} = 0$ if $n < k$.

The following theorem can be proved similarly to the proof of Theorem 4.4.1. We give a proof of this theorem to compare it with that of the next theorem.

Theorem 6.2.1. *For nonnegative integers n and m, we have*

$$\sum_{k\geq 0}\nu_{n,k}\mu_{k,m}=\delta_{n,m}.$$

Proof. Since the proof is similar to that of Theorem 4.4.1, we only give a sketch. Note that the sum in the theorem is 0 if $n<m$ because $\nu_{n,k}\mu_{k,m}=0$ unless $n\geq k\geq m$. Thus we may assume $n\geq m$.

Let X be the set of pairs (T,π) of a Favard tiling $T\in \mathrm{FT}_{n,k}$ and a Motzkin path $\pi\in \mathrm{Motz}_{k,m}$, for some $m\leq k\leq n$. Let Y be the set of pairs (T,π) of a Favard tiling $T\in \mathrm{FT}_{n,n}$ and a Motzkin path $\pi\in \mathrm{Motz}_{n,m}$ such that T has red monominos only and π has up steps only. Then Y has a unique element if $n=m$; and $Y=\emptyset$ if $n\neq m$.

Our goal is to find a weight-preserving sign-reversing involution $\phi:X\to X$ with fixed point set Y. To do this, let $(T,\pi)\in X$ with $T\in\mathrm{FT}_{n,k}$ and $\pi\in\mathrm{Motz}_{k,m}$. We write $\pi=S_1S_2\cdots S_k$ as a sequence of steps. Let a,u be the integers defined as follows:

- a is the largest integer such that T starts with a red monominos,
- u is the largest integer such that π starts with u up steps.

We define $\phi(T,\pi)=(T',\pi')$ in the following way. Here, a',u' are the integers defined similarly as above using T' and π'.

Case 1 $u<a$. There are two subcases.

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

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

and define T' to be the Favard tiling obtained from T by replacing the $(u+1)$st red monomino by a black monomino. See Figure 6.2.

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

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

and define T' to be the Favard tiling obtained from T by replacing the uth and $(u+1)$st red monominos by a domino. See Figure 6.3.

Case 2 $u\geq a\neq n$. Let A be the $(a+1)$st tile in T. There are two subcases.

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

$$\pi'=S_1\cdots S_aHS_{a+1}\cdots S_k,$$

and define T' to be the Favard tiling obtained from T by replacing A by a red monomino. See Figure 6.2 (with the roles of (T,π) and (T',π') interchanged).

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

$$\pi'=S_1\ldots S_aUDS_{a+1}\ldots S_k,$$

and define T' to be the Favard tiling obtained from T by replacing A by two red monominos. See Figure 6.3 (with the roles of (T,π) and (T',π') interchanged).

Case 3 $u\geq a=n$. Since $u\leq k\leq n$, we must have $u=a=n$. Then T has only red monominos and π has only up steps. We define $(T',\pi')=(T,\pi)$.

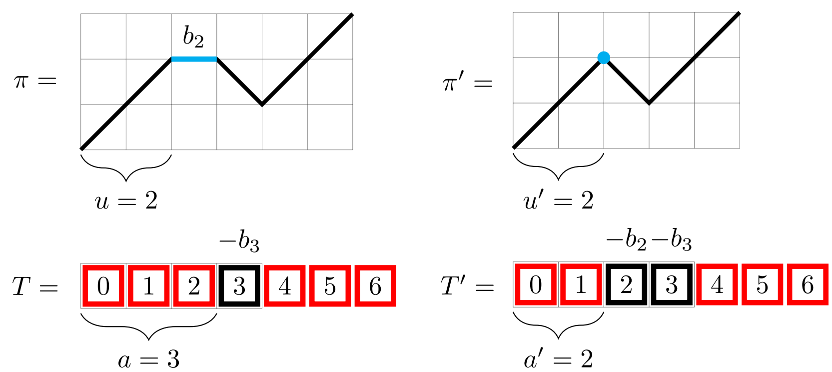

Figure 6.2: A pair $(T, \pi) \in X$ in Case 1-1 on the left and the corresponding triple $(T', \pi') \in X$ in Case 2-1 on the right.

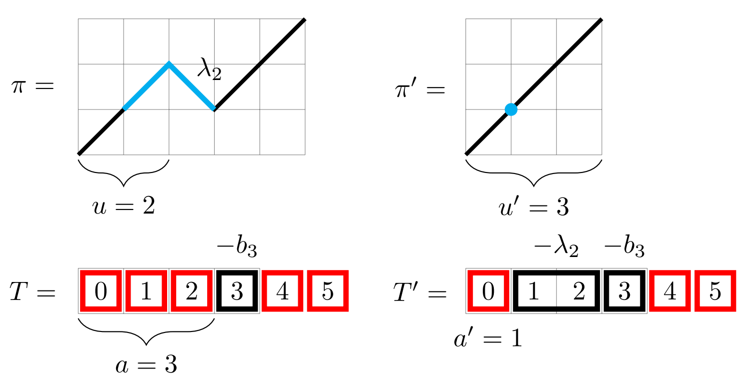

Figure 6.3: A pair $(T, \pi) \in X$ in Case 1-2 on the left and the corresponding triple $(T', \pi') \in X$ in Case 2-2 on the right.

The map $\phi : X \to X$ is a weight-preserving sign-reversing involution with fixed point set Y. Hence

$$\sum_{(T,\pi)\in X} \mathrm{wt}'(T)\,\mathrm{wt}(\pi) = \sum_{(T,\pi)\in Y} \mathrm{wt}'(T)\,\mathrm{wt}(\pi) = \delta_{n,m}.$$

□

Now we prove the other duality.

Theorem 6.2.2. *For nonnegative integers n and m, we have*

$$\sum_{k\geq 0} \mu_{n,k}\nu_{k,m} = \delta_{n,m}.$$

Proof. The proof is quite similar to that of Theorem 6.2.1 except that we consider the end of a Motzkin instead of its beginning. Again, we may assume $n \geq m$.

Let X be the set of pairs (π, T) of a Motzkin path $\pi \in \mathrm{Motz}_{n,k}$ and a Favard tiling $T \in \mathrm{FT}_{k,m}$, for some $m \leq k \leq n$. Let Y be the set of pairs (π, T) of a Motzkin path $\pi \in \mathrm{Motz}_{n,m}$ and a Favard tiling $T \in \mathrm{FT}_{m,m}$ such that π has up steps only and T has red monominos only. Then Y has a unique element if $n = m$; and $Y = \emptyset$ if $n \neq m$.

Our goal is to find a weight-preserving sign-reversing involution $\phi : X \to X$ with fixed point set Y. To do this, let $(\pi, T) \in X$ with $\pi \in \mathrm{Motz}_{n,k}$ and $T \in \mathrm{FT}_{k,m}$ with $m \leq k \leq n$. We write $\pi = S_n S_{n-1} \cdots S_1$ as a sequence of steps. Let a, u be the integers defined as follows:

- a is the largest integer such that T end with a red monominos,
- u is the largest integer such that π ends with u up steps.

We define $\phi(\pi, T) = (\pi', T')$ in the following way. Here, a', u' are the integers defined similarly as above using π' and T'.

Case 1 $n \neq u \leq a$. There are two subcases.

Case 1-1 S_{u+1} is a horizontal step. Let π' be the Motzkin path obtained from π by replacing S_{u+1} by U and define T' to be the Favard tiling obtained from T by inserting a black monomino before the last u red monominos. See Figure 6.4.

Case 1-2 S_{u+1} is a down step. Let π' be the Motzkin path obtained from π by replacing S_{u+1} by U and define T' to be the Favard tiling obtained from T by inserting a black domino before the last u red monominos. See Figure 6.5.

Case 2 $u > a$. Since $a < u \leq k$, we can let A be the $(a+1)$st tile from the right in T. There are two subcases.

Case 2-1 A is a black monomino. Let π' be the Motzkin path obtained from π by replacing S_{a+1} by H and define T' to be the Favard tiling obtained from T by deleting the $(a+1)$st tile from the right in T. See Figure 6.4. (with the roles of (π, T) and (π', T') interchanged).

Case 2-2 A is a domino. Let π' be the Motzkin path obtained from π by replacing S_{a+1} by D and define T' to be the Favard tiling obtained from T by deleting the $(a+1)$st tile from the right in T. See Figure 6.5. (with the roles of (π, T) and (π', T') interchanged).

Case 3 $n = u \leq a$. Since $n \geq m \geq a$, we must have $u = a = n$. Then T has only red monominos and π has only up steps. We define $(\pi', T') = (\pi, T)$.

The map $\phi : X \to X$ is a weight-preserving sign-reversing involution with fixed point set Y. Hence

$$\sum_{(\pi,T)\in X} \mathrm{wt}'(T)\,\mathrm{wt}(\pi) = \sum_{(\pi,T)\in Y} \mathrm{wt}'(T)\,\mathrm{wt}(\pi) = \delta_{n,m}.$$

□

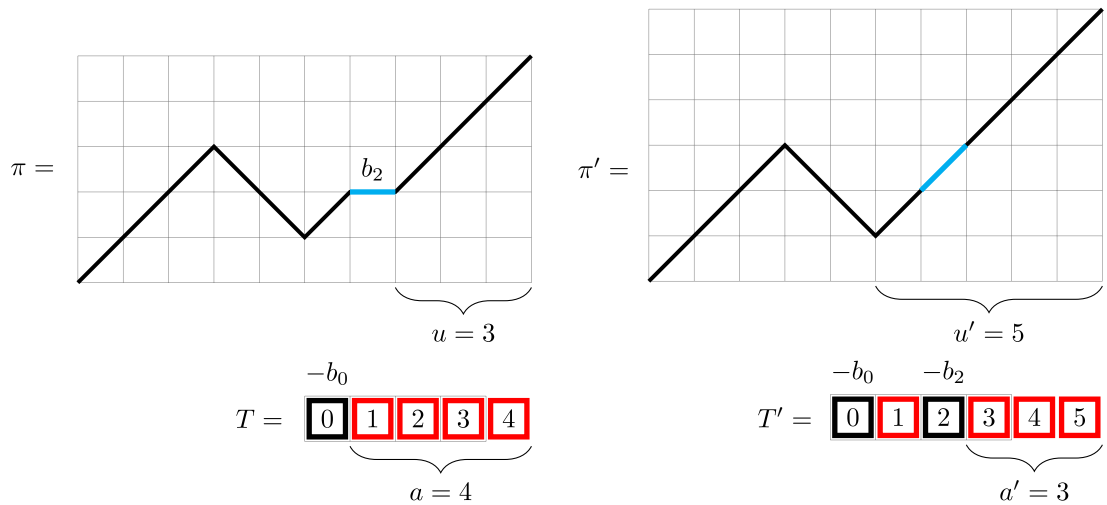

Figure 6.4: A pair $(T, \pi) \in X$ in Case 1-1 on the left and the corresponding triple $(T', \pi') \in X$ in Case 2-1 on the right.

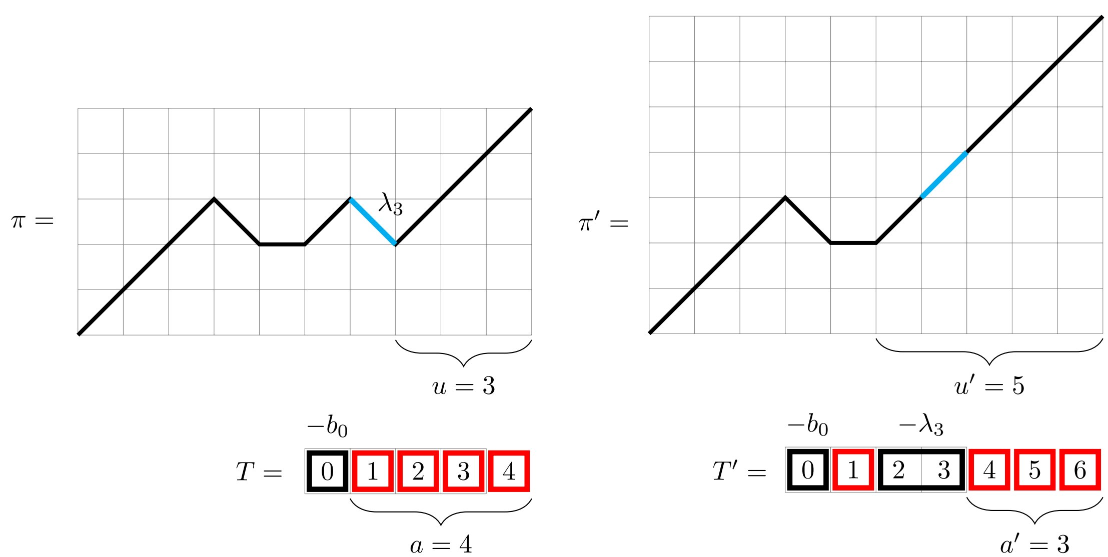

Figure 6.5: A pair $(T, \pi) \in X$ in Case 1-2 on the left and the corresponding triple $(T', \pi') \in X$ in Case 2-2 on the right.

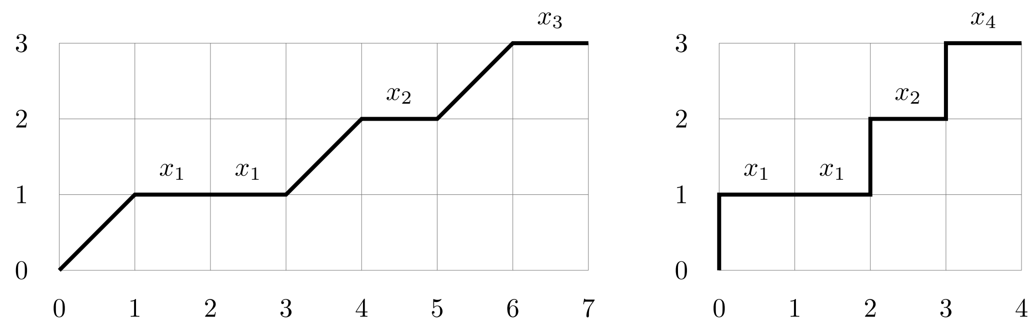

Figure 6.6: A path with steps $U = (1,1)$ and $H = (1,0)$ on the left and the corresponding path with steps $U' = (0,1)$ and $H = (1,0)$ on the right.

6.3 Special case: elementary and homogeneous symmetric functions

For the rest of this chapter we will study several special cases of the duality between $\mu_{n,k}$ and $\nu_{n,k}$. In this section we consider the case that b_k are arbitrary and $\lambda_k = 0$. The next sections will cover some interesting special cases of this.

Suppose that $b_k = x_k$ and $\lambda_k = 0$, where $x_0, x_1, \ldots$ are indeterminates. Observe that if $\lambda_k = 0$, then $\pi \in \mathrm{Motz}_{n,k}$ is a path from $(0,0) \to (n,k)$ consisting of up steps $U = (1,1)$ and horizontal steps $H = (1,0)$. By replacing each $U = (1,1)$ by $U' = (1,0)$, we can identify π as a path from $(0,0)$ to $(n-k,k)$, see Figure 6.6.

Therefore we can write

$$\mu_{n,k} = \sum_{\pi:(0,0)\to(n-k,k)} \mathrm{wt}(\pi),$$

where the sum is over all paths π from $(0,0)$ to $(n-k,k)$ with steps U' and H, and $\mathrm{wt}(\pi)$ is the product of x_k for each horizontal step of height k. Such a path is completely determined by its weight. For example, the path in Figure 6.6 is the unique path from $(0,0)$ to $(4,3)$ with weight $x_1^2 x_2 x_3$. Moreover, every weight is of the form $x_{i_1} \cdots x_{i_{n-k}}$ with $0 \le i_1 \le \cdots \le i_{n-k} \le k$. Thus

$$\mu_{n,k} = \sum_{0 \le i_1 \le \cdots \le i_{n-k} \le k} x_{i_1} \cdots x_{i_{n-k}}. \tag{6.3.1}$$

This is a homogeneous symmetric polynomial.

Definition 6.3.1. Let $x = (x_0, x_1, \ldots)$ be a sequence of variables. A power series $f(x_0, x_1, \ldots)$ in the variables x is called a **symmetric function** if it is invariant under permuting variables. A **homogeneous symmetric function** h_k is defined by

$$h_k = \sum_{i_1 \le \cdots \le i_k} x_{i_1} \cdots x_{i_k}.$$

An **elementary symmetric function** e_k is defined by

$$e_k = \sum_{i_1 < \cdots < i_k} x_{i_1} \cdots x_{i_k}.$$

We define $h_0 = e_0 = 1$ and $h_k = e_k = 0$ if $k < 0$. A **homogeneous symmetric polynomial** $h_k(x_0, x_1, \ldots, x_n)$ is defined by

$$h_k(x_0, x_1, \ldots, x_n) = \sum_{0 \le i_1 \le \cdots \le i_k \le n} x_{i_1} \cdots x_{i_k}.$$

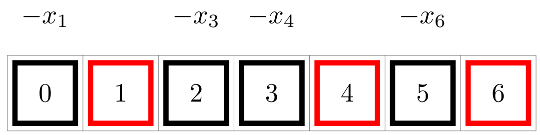

Figure 6.7: A Favard tiling with red monominos and black monominos.

An **elementary symmetric polynomial** $e_k(x_0, x_1, \dots, x_n)$ is defined by

$$e_k(x_0, x_1, \dots, x_n) = \sum_{0 \le i_1 < \cdots < i_k \le n} x_{i_1} \cdots x_{i_k}.$$

For example,

$$\begin{aligned} h_1(x_0, x_1, x_2) &= e_1(x_0, x_1, x_2) = x_0 + x_1 + x_2, \\ h_2(x_0, x_1, x_2) &= x_0^2 + x_1^2 + x_2^2 + x_0x_1 + x_0x_2 + x_1x_2, \\ e_2(x_0, x_1, x_2) &= x_0x_1 + x_0x_2 + x_1x_2, \\ e_3(x_0, x_1, x_2) &= x_0x_1x_2. \end{aligned}$$

Then we can rewrite (6.3.1) as follows.

Theorem 6.3.2. *Suppose that $b_k = x_k$ and $\lambda_k = 0$. Then*

$$\mu_{n,k} = h_{n-k}(x_0, x_1, \dots, x_k).$$

Now we consider

$$\nu_{n,k} = \sum_{T \in \mathrm{FT}_{n,k}} \mathrm{wt}'(T).$$

Since $\lambda_i = 0$, every $T \in \mathrm{FT}_{n,k}$ has red monominos and black monominos only. Hence, T is determined by choosing $n-k$ squares for black monominos in a $1 \times n$ board. Moreover, $\mathrm{wt}'(T)$ is of the form $(-1)^{n-k} x_{i_1} \cdots x_{i_{n-k}}$ for some $0 \le i_1 < \cdots < i_{n-k} \le n-1$, see Figure 6.7. This shows that

$$\nu_{n,k} = \sum_{0 \le i_1 < \cdots < i_{n-k} \le n-1} (-1)^{n-k} x_{i_1} \cdots x_{i_{n-k}}.$$

This can be restated as follows.

Theorem 6.3.3. *Suppose that $b_k = x_{k+1}$ and $\lambda_k = 0$. Then*

$$\nu_{n,k} = (-1)^{n-k} e_{n-k}(x_0, x_1, \dots, x_{n-1}).$$

By the duality (Theorem 6.2.1 and Theorem 6.2.2) and Theorem 6.3.2 and Theorem 6.3.3, we obtain the following corollary.

Corollary 6.3.4. *The following matrix identities hold:*

$$\begin{aligned} &(h_{n-k}(x_0, x_1, \dots, x_k))_{n,k \ge 0} \left((-1)^{n-k} e_{n-k}(x_0, x_1, \dots, x_{n-1})\right)_{n,k \ge 0} = I, \\ &\left((-1)^{n-k} e_{n-k}(x_0, x_1, \dots, x_{n-1})\right)_{n,k \ge 0} (h_{n-k}(x_0, x_1, \dots, x_k))_{n,k \ge 0} = I. \end{aligned}$$

Equivalently, for fixed integers $n, m \ge 0$,

$$\sum_{k \ge 0} h_{n-k}(x_0, x_1, \dots, x_k)(-1)^{k-m} e_{k-m}(x_0, x_1, \dots, x_{k-1}) = \delta_{n,m}, \tag{6.3.2}$$

$$\sum_{k \ge 0} (-1)^{n-k} e_{n-k}(x_0, x_1, \dots, x_{n-1}) h_{k-m}(x_0, x_1, \dots, x_m) = \delta_{n,m}. \tag{6.3.3}$$

Suppose that $N = n - m \geq 0$. Then, by shifting the index $k \mapsto k + m$ in (6.3.3), we have

$$\sum_{k\geq 0}(-1)^{N-k} e_{N-k}(x_0, x_1, \dots, x_{m+N-1}) h_k(x_0, x_1, \dots, x_m) = \delta_{N,0}.$$

If we let $m \to \infty$, we obtain the well-known identity

$$\sum_{k\geq 0}(-1)^k e_{N-k} h_k = \delta_{N,0}.$$

6.4 Special case: binomial coefficients

In this section we consider the case $b_k = 1$ and $\lambda_k = 0$. We will show that the duality of this case is related to the principle of inclusion and exclusion.

Suppose $b_k = 1$ and $\lambda_k = 0$. As observed in the previous section, in this case $\mu_{n,k}$ is the number of paths from $(0,0)$ to $(n-k, k)$ using steps $(1,0)$ and $(0,1)$. Thus $\mu_{n,k} = \binom{n}{k}$. On the other hand, $\nu_{n,k}$ is $(-1)^{n-k}$ times the number of Favard tilings of size n with k red monominos and $n-k$ black monominos. Hence $\nu_{n,k} = (-1)^{n-k}\binom{n}{k}$.

Proposition 6.4.1. *If $b_k = 1$ and $\lambda_k = 0$, we have*

$$\mu_{n,k} = \binom{n}{k}, \qquad \nu_{n,k} = (-1)^{n-k}\binom{n}{k}.$$

As a corollary, we obtain the following duality between binomial coefficients.

Corollary 6.4.2. *We have*

$$\left(\binom{n}{k}\right)_{n,k\geq 0}\left((-1)^{n-k}\binom{n}{k}\right)_{n,k\geq 0} = \left((-1)^{n-k}\binom{n}{k}\right)_{n,k\geq 0}\left(\binom{n}{k}\right)_{n,k\geq 0} = I.$$

Equivalently,

$$\left(\binom{n}{k}\right)_{n,k\geq 0}^{-1} = \left((-1)^{n-k}\binom{n}{k}\right)_{n,k\geq 0}. \tag{6.4.1}$$

Equation (6.4.1) has an interesting connection with the principle of inclusion and exclusion. To see this, suppose that $A_1, \dots, A_n$ are subsets of a set X. For a subset $I \subseteq [n]$, we define

$$A_{=I} = \{x \in X : x \in A_i \text{ if and only if } i \in I\},$$
$$A_{\geq I} = \{x \in X : x \in A_i \text{ for all } i \in I\}.$$

In other words, $A_{=I}$ is the set of elements x which are contained in exactly those A_i for $i \in I$ and $A_{\geq I}$ is the set of elements x which are contained in at least those A_i for $i \in I$.

Note that

$$A_{\geq I} = \bigcap_{i\in I} A_i = \bigcup_{J\supseteq I} A_{=J}.$$

Thus

$$|A_{\geq I}| = \sum_{J\supseteq I} |A_{=J}|. \tag{6.4.2}$$

We can invert this equation as follows, which is a form of the principle of inclusion and exclusion.

Lemma 6.4.3. *For any subset I of $[n]$, we have*

$$|A_{=I}| = \sum_{J\supseteq I} (-1)^{|J-I|} |A_{\geq J}|. \tag{6.4.3}$$

Proof. We will prove this by considering the contribution of each element $x \in X$ to both sides of the equation. Note that every $x \in X$ is contained in A_K for a unique subset K of $[n]$. We consider the following three cases.

Case 1: $K = I$. Then the contribution of x in both sides is 1.

Case 2: $K \supsetneq I$. Then the contribution of x to the left-hand side is 0. The contribution of x to the right-hand side is

$$\sum_{I \subseteq J \subseteq K} (-1)^{|J-I|} = \sum_{J \subseteq (K-I)} (-1)^{|J|} = \sum_{j=0}^{|K-I|} (-1)^j \binom{|K-I|}{j} = 0.$$

Case 3: $K \not\supseteq I$. In this case the contribution of x in both sides is 0.

Since the contribution of x to both sides is always the same, the identity holds. □

If $I = \emptyset$, we obtain the following common form of the principle of inclusion and exclusion:

$$\begin{aligned}
|A_1^c \cap \cdots \cap A_n^c| &= \sum_{k=0}^{n} (-1)^k \sum_{i_1 < \cdots < i_k} |A_{i_1} \cap \cdots \cap A_{i_k}| \\
&= |X| - |A_1| - \cdots - |A_n| \\
&\quad + |A_1 \cap A_2| + |A_1 \cap A_3| + \cdots + |A_{n-1} \cap A_n| \\
&\quad - \cdots \\
&\quad + (-1)^n |A_1 \cap \cdots \cap A_n|.
\end{aligned}$$

Example 6.4.4. A derangement is a permutation without fixed points. Let d_n be the number of derangements in $\mathfrak{S}_n$. To compute d_n, we define $X = \mathfrak{S}_n$ and $A_i = \{\pi \in X : \pi(i) = i\}$. Then

$$d_n = |A_{=\emptyset}| = \sum_{J \subseteq [n]} (-1)^{|J|} |A_{\geq J}|.$$

If $J = \{j_1, \dots, j_k\}$, then $|A_{\geq J}| = (n-k)!$. Thus

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

Note that

$$\frac{d_n}{n!} = \sum_{k=0}^{n} \frac{(-1)^k}{k!} \approx \frac{1}{e} = 0.367879441171 \cdots .$$

In the previous example, $A_{=I}$ and $A_{\geq I}$ depend only on the cardinality of I. In this case let $a_k = |A_{=I}|$ and $b_k = |A_{\geq I}|$ for any subset I of cardinality k. Then (6.4.2) and (6.4.3) can be written as

$$\begin{aligned}
b_k &= \sum_{j=k}^{n} \binom{n-k}{j-k} a_j = \sum_{j=0}^{n-k} \binom{n-k}{j} a_{j+k}, \\
a_k &= \sum_{j=k}^{n} (-1)^{j-k} \binom{n-k}{j-k} b_j = \sum_{j=0}^{n-k} (-1)^j \binom{n-k}{j} b_{j+k}.
\end{aligned}$$

To make things look nicer, let $b'_k := b_{n-k}$ and $a'_k := a_{n-k}$. Then the above equations can be rewritten as

$$b'_{n-k} = \sum_{j=0}^{n-k} \binom{n-k}{n-k-j} a'_{n-k-j} = \sum_{j=0}^{n-k} \binom{n-k}{j} a'_j,$$

$$a'_{n-k} = \sum_{j=0}^{n-k} (-1)^j \binom{n-k}{n-k-j} b'_{n-k-j} = \sum_{j=0}^{n-k} (-1)^{n-k-j} \binom{n-k}{j} b'_j.$$

Finally, replacing k by $n-k$, we obtain

$$b'_k = \sum_{j=0}^{k} \binom{k}{j} a'_j,$$

$$a'_k = \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} b'_j.$$

Equivalently,

$$\begin{pmatrix} b'_0 \\ \vdots \\ b'_n \end{pmatrix} = \left(\binom{i}{j} \right)_{i,j=0}^{n} \begin{pmatrix} a'_0 \\ \vdots \\ a'_n \end{pmatrix}$$

$$\begin{pmatrix} a'_0 \\ \vdots \\ a'_n \end{pmatrix} = \left((-1)^{i-j} \binom{i}{j} \right)_{i,j=0}^{n} \begin{pmatrix} b'_0 \\ \vdots \\ b'_n \end{pmatrix}$$

Since a'_i and b'_i can be anything, we have the following matrix identity:

$$\left(\binom{i}{j} \right)_{i,j=0}^{n} = \left((-1)^{i-j} \binom{i}{j} \right)_{i,j=0}^{n},$$

which is equivalent to (6.4.1).

6.5 Special case: q-binomial coefficients

In this section we consider the case $b_k = q^k$ and $\lambda_k = 0$. This case gives q-binomial coefficients. We first need some definitions. From now on, we treat q as an indeterminate.

Definition 6.5.1. For a nonnegative integer n, the **q-integer** $[n]_q$ is defined by

$$[n]_q = \frac{1-q^n}{1-q} = 1 + \cdots + q^{n-1}.$$

The **q-factorial** $[n]_q!$ and the **q-binomial coefficient** $\begin{bmatrix} n \\ k \end{bmatrix}_q$, for $0 \le k \le n$, are defined by

$$[n]_q! = [1]_q [2]_q \cdots [n]_q, \qquad \begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{[n]_q!}{[k]_q! [n-k]_q!}.$$

We also define $\begin{bmatrix} n \\ k \end{bmatrix}_q = 0$ if $n < k$.

Note that if $q = 1$, then

$$[n]_q = n, \qquad [n]_q! = n!, \qquad \begin{bmatrix} n \\ k \end{bmatrix}_q = \binom{n}{k}.$$

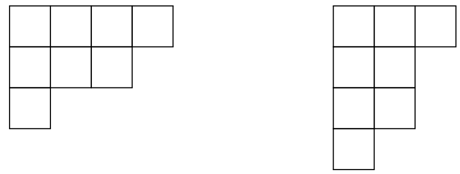

Figure 6.8: The Young diagram of the partition $\lambda = (4,3,1)$ and its transpose $\lambda' = (3,2,2,1)$.

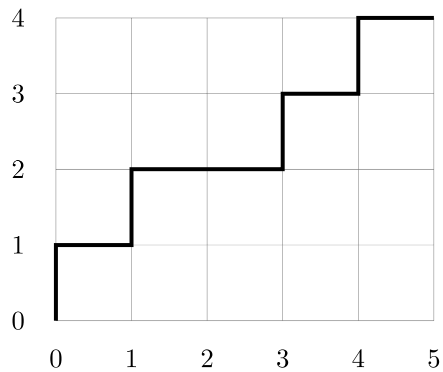

Figure 6.9: The path corresponding to a partition $\lambda = (4,3,1)$ contained in a rectangle (5^4).

Lemma 6.5.2. *For $0 \le k \le n$, we have*

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = q^{n-k} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q + \begin{bmatrix} n-1 \\ k \end{bmatrix}_q . \tag{6.5.1}$$

Proof. We compute

$$\begin{aligned} q^{n-k} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q + \begin{bmatrix} n-1 \\ k \end{bmatrix}_q &= \frac{q^{n-k}[n-1]_q!}{[n-k]_q![k-1]_q!} + \frac{[n-1]_q!}{[n-1-k]_q![k]_q!}, \\ &= \frac{[n-1]_q!}{[n-k]_q![k]_q!} \left(q^{n-k}[k]_q + [n-k]_q \right), \\ &= \frac{[n-1]_q!}{[n-k]_q![k]_q!} [n]_q = \begin{bmatrix} n \\ k \end{bmatrix}_q . \end{aligned}$$

□

Definition 6.5.3. A **partition** is a sequence of nonnegative integers $\lambda = (\lambda_1, \dots, \lambda_\ell)$ with $\lambda_1 \ge \dots \ge \lambda_\ell$. Each λ_i is called a **part** of λ. The **size** of λ is defined to be $|\lambda| = \lambda_1 + \dots + \lambda_\ell$. The **Young diagram** of λ is a left-justified array of squares where the ith row has λ_i squares. The **transpose** λ' of λ is the partition whose Young diagram is obtained by reflecting the Young diagram of λ along the diagonal as shown in Figure 6.8.

Let (a^b) denote the partition with b parts equal to a. For two partitions λ and μ, we write $\mu \subseteq \lambda$ to mean that the Young diagram of μ is contained in that of λ. Note that considering the Young diagram, $\lambda \subseteq ((n-k)^k)$ can be identified with a path from $(0,0)$ to $(n-k,k)$, see Figure 6.9. The following proposition shows that the q-binomial coefficient $\begin{bmatrix} n \\ k \end{bmatrix}_q$ is always a polynomial in q.

Proposition 6.5.4. *For $0 \le k \le n$, we have*

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \sum_{\lambda \subseteq ((n-k)^k)} q^{|\lambda|}. \tag{6.5.2}$$

Proof. We prove this by induction on n. If $n = 0$, then $k = 0$, and both sides are equal to 1. Let $n \ge 1$ and suppose that the statement holds for $n-1$. To prove the statement for n, it suffices to show that the right-hand side of the equation satisfies the same recurrence as (6.5.1).

To this end consider $\lambda = (\lambda_1, \dots, \lambda_{n-k}) \subseteq ((n-k)^k)$. If $\lambda_1 = n-k$, then $(\lambda_2, \dots, \lambda_{n-k}) \subseteq ((n-k)^{k-1}) = ((n-1-(k-1))^{k-1})$. If $\lambda_1 \le n-k-1$, then $(\lambda_1, \dots, \lambda_{n-k}) \subseteq ((n-k-1)^k)$. Thus

$$\sum_{\lambda \subseteq ((n-k)^k)} q^{|\lambda|} = q^{n-k} \sum_{\lambda \subseteq ((n-1-(k-1))^{k-1})} q^{|\lambda|} + \sum_{\lambda \subseteq ((n-1-k)^k)} q^{|\lambda|},$$

which is the same recurrence as (6.5.1). □

Note that by taking the transpose of λ, we can rewrite (6.5.2) as

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \sum_{\lambda \subseteq (k^{n-k})} q^{|\lambda|}. \tag{6.5.3}$$

Now we are ready to consider the mixed moments and coefficients when $b_k = q^k$ and $\lambda_k = 0$.

Proposition 6.5.5. *If $b_k = q^k$ and $\lambda_k = 0$, we have*

$$\mu_{n,k} = \begin{bmatrix} n \\ k \end{bmatrix}_q, \qquad \nu_{n,k} = (-1)^{n-k} q^{\binom{n-k}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q.$$

Proof. By Theorem 6.3.2 and Theorem 6.3.3, we have

$$\begin{aligned} \mu_{n,k} &= h_{n-k}(x_0, x_1, \dots, x_k), \\ \nu_{n,k} &= (-1)^{n-k} e_{n-k}(x_0, x_1, \dots, x_{n-1}), \end{aligned}$$

where $x_i = q^i$. Thus, by (6.5.2),

$$\mu_{n,k} = \sum_{0 \le i_1 \le \cdots \le i_{n-k} \le k} q^{i_1 + \cdots + i_{n-k}} = \sum_{\mu \subseteq (k^{n-k})} q^{|\mu|} = \begin{bmatrix} n \\ k \end{bmatrix}_q.$$

The second identity follows from

$$\begin{aligned} e_{n-k}(1, q, \dots, q^n) &= \sum_{0 \le i_1 < \cdots < i_{n-k} \le n-1} q^{i_1 + \cdots + i_{n-k}} \\ &= \sum_{0 \le j_1 \le \cdots \le j_{n-k} \le k} q^{0+1+\cdots+(n-k-1)} q^{j_1 + \cdots + j_{n-k}} \\ &= q^{\binom{n-k}{2}} \sum_{\mu \subseteq (k^{n-k})} q^{|\mu|} = q^{\binom{n-k}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q, \end{aligned}$$

where the change of indices $(i_1, i_2, \dots, i_{n-k}) = (j_1 + 0, j_2 + 1, \dots, j_{n-k} + n - k + 1)$ is used. □

6.6 Special case: Stirling numbers

In this section we consider the case $b_k = k$ and $\lambda_k = 0$. In this case we obtain the duality between Stirling numbers of the first kind and second kind.

Theorem 6.6.1. *If $b_k = k$ and $\lambda_k = 0$, then*

$$\mu_{n,k} = S(n,k), \qquad \nu_{n,k} = s(n,k).$$

Proof. We have

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

Recall that if $b_k = k+1$ and $\lambda_k = k$, then μ_n is the number of Charlier histories from $(0,0)$ to $(n,0)$. In our case $b_k = k$ and $\lambda_k = 0$, the same bijection shows that $\mu_{n,k}$ is the number of paths

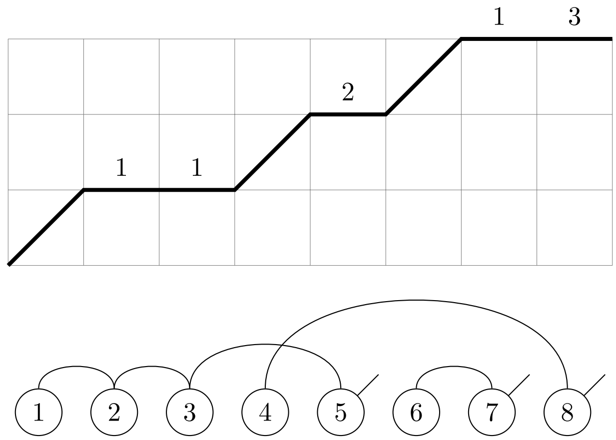

Figure 6.10: A partial Charlier history (top) and the corresponding set partition, where every block has a half edge attached at the end (bottom).

from $(0, 0)$ to (n, k) consisting of $H = (1, 0)$ and $U = (1, 1)$ in which every horizontal step at height h has a label in $\{1, \dots, h\}$. We can apply the same bijection between Charlier histories and set partitions to these (partial) Charlier as shown in Figure 6.10. Since there are no singleton blocks and no closers, we obtain that such (partial) Charlier histories are in bijection with set partitions of $[n]$ into k blocks. This shows that $\mu_{n,k} = S(n, k)$.

Since $b_k = k + 1$ and $\lambda_k = k$, we have

$$\sum_{k=0}^{n} \nu_{n,k} x^k = \sum_{T \in \mathrm{FT}_n} \mathrm{wt}(T) = x(x-1)(x-2)\cdots(x-n+1).$$

On the other hand, by (3.4.4), we have

$$\sum_{k=0}^{n} s(n, k) x^k = x(x-1)(x-2)\cdots(x-n+1).$$

This proves the second identity also holds. $\square$

Chapter 7

Determinants of moments

We have learned that μ_n can be written using b_n and λ_n. Since a monic OPS $\{P_n(x)\}_{n\ge0}$ (and hence the recurrence coefficients b_n and λ_n as well) is uniquely determined by μ_n, there must be a way to express b_n and λ_n using μ_n. In this chapter we find such an expression using nonintersecting lattice paths and the Lindström–Gessel–Viennot lemma.

7.1 Computing the 3-term recurrence coefficients

Let us first see how one can compute μ_n using b_n and λ_n for small n. Recall

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

Thus, the first few moments are

$$\begin{aligned}
\mu_1 &= b_0,\\
\mu_2 &= b_0^2 + \lambda_1,\\
\mu_3 &= b_0^3 + 2b_0\lambda_1 + b_1\lambda_1.
\end{aligned}$$

Then we can solve for b_n and λ_n:

$$\begin{aligned}
b_0 &= \mu_1,\\
\lambda_1 &= \mu_2 - b_0^2 = \mu_2 - \mu_1^2,\\
b_1 &= (\mu_3 - b_0^3 + 2b_0\lambda_1)/\lambda_1 = (\mu_3 - \mu_1^3 + 2\mu_1(\mu_2-\mu_1^2))/(\mu_2-\mu_1^2),
\end{aligned}$$

and so on. Although, the formula gets very complicated, we can convince ourselves that this always gives a formula for b_n and λ_n. We prove this rigorously as follows.

Suppose that we have computed $b_0,\dots,b_{n-1}$ and $\lambda_1,\dots,\lambda_n$. Then in the sum

$$\mu_{2n+1} = \sum_{\pi\in\mathrm{Motz}_{2n+1}} \mathrm{wt}(\pi),$$

b_n appears only once as $b^n\lambda_n\lambda_{n-1}\cdots\lambda_1$ for the Motzkin path U^nHD_n. Thus

$$b^n\lambda_n\lambda_{n-1}\cdots\lambda_1 = \mu_{2n+1} - \sum_{\pi\in\mathrm{Motz}_{2n+1},\pi\ne U^nHD_n} \mathrm{wt}(\pi). \tag{7.1.1}$$

Since the sum on the right-hand side of (7.1.1) has only $b_0,\dots,b_{n-1}$ and $\lambda_1,\dots,\lambda_n$, we can express it using μ_k's. Then dividing both sides of (7.1.1) by $\lambda_1\cdots\lambda_n$, which can also be written using μ_k's, we obtain a formula for b_n in terms of μ_k's.

Similarly, if we have computed $b_0, \dots, b_n$ and $\lambda_1, \dots, \lambda_n$, then we can find a formula for λ_{n+1} in terms of μ_k's using

$$\mu_{2n+2} = \sum_{\pi \in \mathrm{Motz}_{2n+2}} \mathrm{wt}(\pi),$$

because λ_{n+1} appears only once for the path $U^{n+1}D^{n+1}$.

The above algorithm shows that it is possible to express b_n and λ_n using μ_n. To find an explicit formula, we need to develop some interesting theory of lattice paths.

7.2 The Lindström–Gessel–Viennot lemma

The Lindström–Gessel–Viennot lemma [7, 6] is a very useful tool in combinatorics. This lemma is listed in the book called "Proofs from The Book" [1, Chapter 32], which tries to collect the most beautiful proofs in mathematics. We start with basic definitions on paths in a directed graph.

Definition 7.2.1. A **graph** is a pair $G = (V, E)$ of two sets V and E such that $E \subseteq V \times V$. Each element $v \in V$ is called a **vertex** and each element $(u, v) \in E$ is called an **edge**. We say that G is **undirected** if (u, v) is identified with (v, u). Otherwise, G is said to be **directed**.

A **path** from u to v is a sequence of vertices $(v_0, v_1, \dots, v_n)$ such that $v_0 = u$, $v_n = v$, and $(v_i, v_{i+1}) \in E$ for all $0 \le i \le n-1$. A **cycle** is a path from a vertex to itself. For two vertices u and v, we denote by $P(u \to v)$ the set of paths from u to v. If there is no cycle, G is said to be **acyclic**.

An **edge weight** of G is a function $w : E \to K$, for some commutative ring K. The **weight** of a path p is defined to be the product of $w(e)$ for every edge e in p.

We consider families of paths. For brevity, we define an **n-path** to be just an n-tuple $\boldsymbol{p} = (p_1, \dots, p_n)$ of paths. We say that two paths p and p' are **nonintersecting** if they do not have a common vertex. We also say that an n-path $\boldsymbol{p} = (p_1, \dots, p_n)$ is **nonintersecting** if p_i and p_j are nonintersecting for all $i \neq j$.

Definition 7.2.2. Let G be a directed graph with edge weight w. Let $\boldsymbol{A} = (A_1, \dots, A_n)$ and $\boldsymbol{B} = (B_1, \dots, B_n)$ be sequences of vertices of G. We denote by $P(\boldsymbol{A} \to \boldsymbol{B})$ the set of n-paths $\boldsymbol{p} = (p_1, \dots, p_n)$ such that $p_i \in P(A_i \to B_{\sigma(i)})$, $1 \le i \le n$, for some $\sigma \in \mathfrak{S}_n$. We define $w(\boldsymbol{p}) = w(p_1) \cdots w(p_n)$ and $\mathrm{sgn}(\boldsymbol{p}) = \mathrm{sgn}(\sigma)$. Finally, we define $\mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B})$ to be the set of all nonintersecting n-paths in $P(\boldsymbol{A} \to \boldsymbol{B})$.

We are now ready to state the Lindström–Gessel–Viennot lemma.

Theorem 7.2.3 (The Lindström–Gessel–Viennot lemma)**.** *Let G be a directed graph with edge weight w. Fix vertex sequences $\boldsymbol{A} = (A_1, \dots, A_n)$ and $\boldsymbol{B} = (B_1, \dots, B_n)$ and define the matrix $M = (M_{i,j})_{i,j=1}^n$ by*

$$M_{i,j} = \sum_{p \in P(A_i \to B_j)} w(p).$$

Then we have

$$\det M = \sum_{\boldsymbol{p} \in \mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B})} \mathrm{sgn}(\boldsymbol{p}) w(\boldsymbol{p}).$$

Before proving this theorem let us consider an example.

Example 7.2.4. Let G be the directed graph whose vertex set V and (directed) edge set E are given by

$$V = \{(i, j) : 0 \le i, j \le 2\},$$
$$E = \{(i, j) \to (i+1, j) : 0 \le i \le 1, 0 \le j \le 2\} \cup \{(i, j) \to (i, j+1) : 0 \le i \le 2, 0 \le j \le 1\}.$$

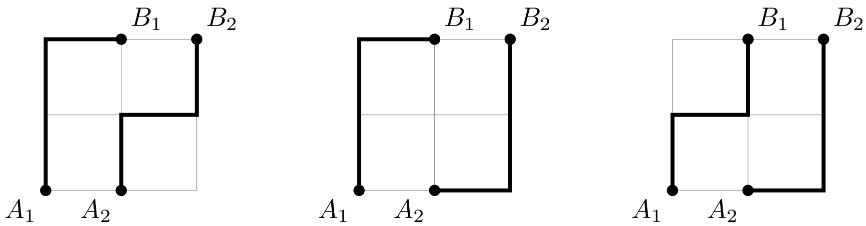

Figure 7.1: All nonintersecting 3-paths from A_1, A_2 to B_1, B_2.

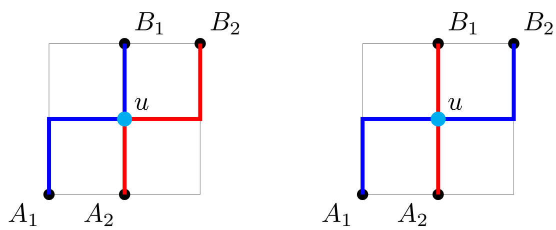

Figure 7.2: An intersecting 2-path (p_1, p_2) and the corresponding 2-path (q_1, q_2) obtained by exchanged the tails after the first intersection u.

Define the weight of every edge to be 1. Let $\boldsymbol{A} = (A_1, A_2)$ and $\boldsymbol{B} = (B_1, B_2)$, where $A_1 = (0,0)$, $A_2 = (1,0)$, $B_1 = (1,2)$, and $B_2 = (2,2)$. Then

$$\det M = \det \begin{pmatrix} \binom{3}{1} & \binom{4}{2} \\ \binom{2}{0} & \binom{3}{1} \end{pmatrix} = \det \begin{pmatrix} 3 & 6 \\ 1 & 3 \end{pmatrix} = 3.$$

On the other hand, there are exactly 3 nonintersecting 2-paths from $\boldsymbol{A}$ to $\boldsymbol{B}$ as shown in Figure 7.1.

Note that in the above example, we can compute the number of nonintersecting 2-paths $\boldsymbol{p} \in \mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B})$ as follows. First, observe that if $\boldsymbol{p} = (p_1, p_2) \in \mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B})$, then $p_1 \in P(A_1 \to B_1)$ and $p_2 \in P(A_2 \to B_2)$. Thus $\mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B})$ is contained in the set $P(A_1 \to B_1) \times P(A_2 \to B_2)$ whose cardinality is $\binom{3}{1}\binom{3}{1}$. So, if we subtract the number of intersecting 2-paths $(p_1, p_2) \in P(A_1 \to B_1) \times P(A_2 \to B_2)$, we would get the cardinality of $\mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B})$.

Suppose that $(p_1, p_2) \in P(A_1 \to B_1) \times P(A_2 \to B_2)$ is intersecting. Then we can find the first intersection point u of p_1 and p_2. Then we can write $p_1 = p_1' p_1''$ and $p_2 = p_2' p_2''$, where p_i' (resp. p_i'') is the part of p_i before u (resp. after u). By exchanging the tails p_1'' and p_2'' we can construct a new 2-path (q_1, q_2), that is, $q_1 = p_1' p_2''$ and $q_2 = p_2' p_1''$, see Figure 7.2.

Note that $(q_1, q_2) \in P(A_1 \to B_2) \times P(A_2 \to B_1)$, and any such 2-path is always intersecting and thus gives rise to an intersecting 2-path $(p_1, p_2) \in P(A_1 \to B_1) \times P(A_2 \to B_2)$ by the same process of exchanging the tails. Thus the number of intersecting 2-paths $(p_1, p_2) \in P(A_1 \to B_1) \times P(A_2 \to B_2)$ is equal to

$$|P(A_1 \to B_2) \times P(A_2 \to B_1)| = \binom{4}{2}\binom{2}{0}.$$

We have just shown that

$$|\mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B})| = \binom{3}{1}\binom{3}{1} - \binom{4}{2}\binom{2}{0} = \det \begin{pmatrix} \binom{3}{1} & \binom{4}{2} \\ \binom{2}{0} & \binom{3}{1} \end{pmatrix}.$$

This idea of canceling intersecting 2-paths can be extended to prove Theorem 7.2.3.

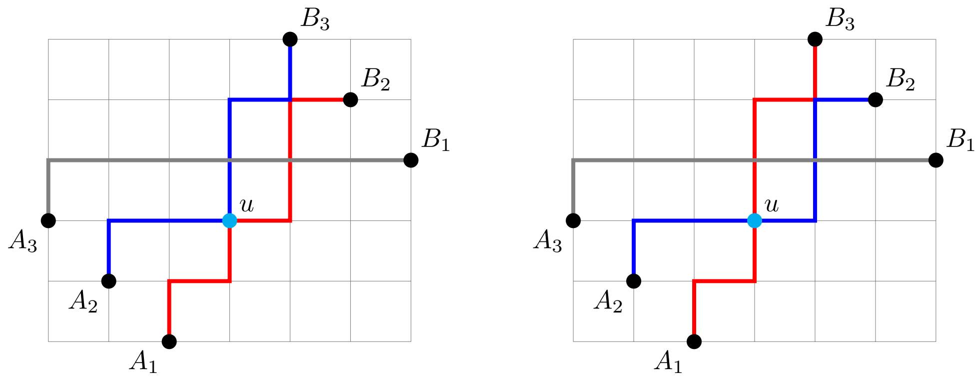

Figure 7.3: The involution ϕ exchanges the tails of p_1 and p_2 after their first intersection point u.

Proof of Theorem 7.2.3. By definition,

$$\det M = \sum_{\sigma\in\mathfrak{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} M_{i,\sigma(i)} = \sum_{\sigma\in\mathfrak{S}_n} \operatorname{sgn}(\sigma) \sum_{p\in P(A_i\to B_{\sigma(i)})} w(p) = \sum_{\boldsymbol{p}\in P(\boldsymbol{A}\to\boldsymbol{B})} \operatorname{sgn}(\boldsymbol{p}) w(\boldsymbol{p}).$$

Thus it suffices to find a sign-reversing and weight-preserving involution ϕ on $P(\boldsymbol{A}\to\boldsymbol{B})$ with fixed point set $\mathrm{NI}(\boldsymbol{A}\to\boldsymbol{B})$.

Consider an n-path $\boldsymbol{p} = (p_1,\dots,p_n) \in P(\boldsymbol{A}\to\boldsymbol{B})$ with $p_i \in P(A_i \to B_{\sigma(i)})$ for some $\sigma\in\mathfrak{S}_n$. If $\boldsymbol{p}$ is nonintersecting, then define $\phi(\boldsymbol{p}) = \boldsymbol{p}$. Suppose now that $\boldsymbol{p}$ is intersecting. Then we can find the lexicographically smallest pair (r,s) such that p_r and p_s are intersecting. We can then find the first intersection point, say u, of p_r and p_s. Let p_r' and p_s' to be the paths obtained from p_r and p_s by exchanging their tails after u. We define $\phi(\boldsymbol{p}) = \boldsymbol{q}$, where

$$\boldsymbol{q} = (q_1,\dots,q_n) = (p_1,\dots,p_{r-1},p_r',p_{r+1},\dots,p_{s-1},p_s',p_{s+1},\dots,p_n).$$

See Figure 7.3.

Since $\boldsymbol{p}$ and $\boldsymbol{q}$ have the same set of edges (with the same multiplicities), we have $w(\boldsymbol{p}) = w(\boldsymbol{q})$. Since $p_i \in P(A_i \to B_{\sigma(i)})$, we have $q_i \in P(A_i \to B_{\sigma'(i)})$, where σ' is the permutation obtained from $\sigma = \sigma_1\cdots\sigma_n$ by exchanging σ_i and σ_j. In other words, $\sigma' = \sigma(i,j)$, hence

$$\operatorname{sgn}(\boldsymbol{p}') = \operatorname{sgn}(\sigma') = \operatorname{sgn}(\sigma) = \operatorname{sgn}(\boldsymbol{p}).$$

Moreover, by the construction of ϕ, it is clearly an involution. Thus, ϕ is a sign-reversing and weight-preserving involution ϕ on $P(\boldsymbol{A}\to\boldsymbol{B})$ with fixed point set $\mathrm{NI}(\boldsymbol{A}\to\boldsymbol{B})$, which completes the proof. $\square$

Corollary 7.2.5. *Let G be a directed graph with edge weight w. Fix vertex sequences $\boldsymbol{A} = (A_1,\dots,A_n)$ and $\boldsymbol{B} = (B_1,\dots,B_n)$ and define the matrix $M = (M_{i,j})_{i,j=1}^n$ by*

$$M_{i,j} = \sum_{p\in P(A_i\to B_j)} w(p).$$

Suppose that every nonintersecting n-path $\boldsymbol{p} = (p_1,\dots,p_n) \in \mathrm{NI}(\boldsymbol{A}\to\boldsymbol{B})$ satisfies $p_i \in P(A_i \to B_i)$ for all $1\le i\le n$. Then

$$\det M = \sum_{\boldsymbol{p}\in \mathrm{NI}(\boldsymbol{A}\to\boldsymbol{B})} w(\boldsymbol{p}).$$

In particular, if $w(e) = 1$ for all $e\in E$, then the number of nonintersecting n-paths is

$$|\,\mathrm{NI}(\boldsymbol{A}\to\boldsymbol{B})| = \det M.$$

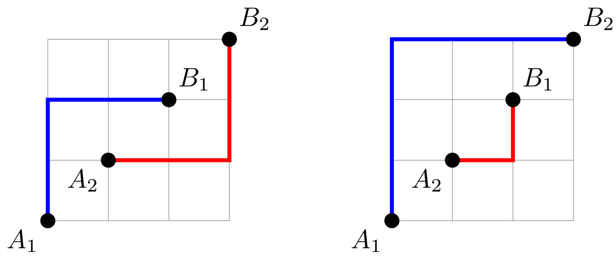

Figure 7.4: Two nonintersecting 2-paths (p_1, p_2) and (q_1, q_2) with different connecting patterns.

Example 7.2.6. The graph in Figure 7.3 and the vertices (A_1, A_2, A_3) and (B_1, B_2, B_3) satisfy the conditions in Corollary 7.2.5. Thus The number of nonintersecting 3-paths from (A_1, A_2, A_3) and (B_1, B_2, B_3) is

$$\det \begin{pmatrix} \binom{7}{4} & \binom{7}{3} & \binom{7}{2} \\ \binom{7}{5} & \binom{7}{4} & \binom{7}{3} \\ \binom{7}{6} & \binom{7}{5} & \binom{7}{4} \end{pmatrix} = 4116.$$

Example 7.2.7. Not every directed graph and vertex sequences satisfy the conditions in Corollary 7.2.5. See Figure 7.4. In this case,

$$\det M = \det \begin{pmatrix} \binom{4}{2} & \binom{6}{3} \\ \binom{2}{1} & \binom{4}{2} \end{pmatrix} = -4.$$

Note that the number of nonintersecting 2-paths (p_1, p_2) with $p_i \in P(A_i \to B_{\sigma(i)})$ is 2 if $\sigma = 12$ and 6 if $\sigma = 21$. This is in agreement of the Lindström–Gessel–Viennot lemma:

$$\operatorname{sgn}(12) \cdot 2 + \operatorname{sgn}(21) \cdot 6 = 2 - 6 = -4.$$

Remark 7.2.8. What if $A_i = A_j$ or $B_i = B_j$ for some $i \neq j$? Then clearly there is no nonintersecting n-path from $\boldsymbol{A}$ to $\boldsymbol{B}$. Thus, $\det M = 0$. This can also be seen immediately from the definition of the matrix M. For example, if $A_i = A_j$ then row i and row j of M are identical so, $\det M = 0$.

Remark 7.2.9. What if $A_i = B_j$ for some $i \neq j$? Then every nonintersecting n-path $\boldsymbol{p} = (p_1, \dots, p_n)$ must satisfy $p_i \in P(A_i \to B_j)$. To see this suppose $p_i \in P(A_i \to B_k)$ for some $k \neq j$. Then there is a path $p_r \in P(A_r \to B_j)$ for some $r \neq i$. Then p_i and p_r have a common vertex $A_i = B_j$, a contradiction.

Note also that if $A_i = B_j$, then $p_i \in P(A_i \to B_j)$ means that $p_i = (A_i)$ is a path of length 0, i.e., a path with no edges. Hence, if $\boldsymbol{p} = (p_1, \dots, p_n)$ is nonintersecting, then every path other than p_i does not touch the vertex A_i. This can be used to find the number of paths in G from u to v avoiding a given list of vertices.

Example 7.2.10. Let G be a directed graph with vertices $A_1, \dots, A_n, B_1, \dots, B_n$ and edges (A_i, B_j) for $1 \leq i, j \leq n$ with edge weight w. Let $\boldsymbol{A} = (A_1, \dots, A_n)$ and $\boldsymbol{B} = (B_1, \dots, B_n)$. Then each $P(A_i \to B_j)$ has only one path (A_i, B_j) with only one edge and every path $\boldsymbol{p} \in P(\boldsymbol{A} \to \boldsymbol{B})$ is nonintersecting. Thus the matrix $M = (M_{i,j})$ has entries

$$M_{i,j} = \sum_{p \in P(A_i \to B_j)} w(p) = w(A_i, B_j).$$

Then the Lindström–Gessel–Viennot lemma says

$$\begin{aligned}\det M &= \sum_{\boldsymbol{p}\in \mathrm{NI}(\boldsymbol{A}\to\boldsymbol{B})} \mathrm{sgn}(\boldsymbol{p}) w(\boldsymbol{p}) = \sum_{\boldsymbol{p}\in P(\boldsymbol{A}\to\boldsymbol{B})} \mathrm{sgn}(\boldsymbol{p}) w(\boldsymbol{p}) \\ &= \sum_{\sigma\in\mathfrak{S}_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} w(A_i, B_{\sigma(i)}) = \sum_{\sigma\in\mathfrak{S}_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} M_{i,j},\end{aligned}$$

which is nothing but the definition of the determinant of a matrix.

There is an interesting application of the Lindström–Gessel–Viennot lemma to a useful determinant identity called the Cauchy–Binet formula.

Definition 7.2.11. Let $M = (M_{i,j})_{i\in[m],j\in[n]}$ be an $m \times n$ matrix. We denote by $\binom{[m]}{k}$ the set of all subsets of $[m]$ with cardinality k. For $I \subseteq \binom{[m]}{k}$ and $J \subseteq \binom{[n]}{k}$, the (I, J)-minor $[M]_{I,J}$ of M is defined by

$$[M]_{I,J} = \det(M_{i,j})_{i\in I,j\in J}.$$

Theorem 7.2.12 (The Cauchy–Binet formula). *Let M be an $n \times \ell$ matrix let N be an $\ell \times n$ matrix. Then*

$$\det(MN) = \sum_{I\in\binom{[\ell]}{n}} [M]_{[n],I}[N]_{I,[n]}.$$

Proof. Let G be the directed graph with vertex set V and edge set E given by

$$\begin{aligned}V &= \{A_1, \dots, A_n, B_1, \dots, B_\ell, C_1, \dots, C_n\},\\ E &= \{(A_i, B_j) : i \in [n], j \in [\ell]\} \cup \{(B_i, C_j) : i \in [\ell], j \in [n]\}.\end{aligned}$$

Define an edge weight w of G by

$$w(A_i, B_j) = M_{i,j}, \qquad w(B_i, C_j) = N_{i,j}.$$

Consider the $n \times n$ matrix $L = (L_{i,j})_{i,j\in[n]}$ whose (i, j)-entry is

$$L_{i,j} = \sum_{p\in P(A_i\to C_j)} w(p).$$

Since the above sum is equal to

$$\sum_{k=1}^{\ell} w(A_i, B_k)w(B_k, C_j) = \sum_{k=1}^{\ell} M_{i,k}N_{k,j} = (MN)_{i,j},$$

we have $L = MN$. Letting $\boldsymbol{A} = (A_1, \dots, A_n)$ and $\boldsymbol{C} = (C_1, \dots, C_n)$, the Lindström–Gessel–Viennot lemma implies

$$\det L = \sum_{\boldsymbol{p}\in\mathrm{NI}(\boldsymbol{A}\to\boldsymbol{C})} \mathrm{sgn}(\boldsymbol{p})w(\boldsymbol{p}).$$

Observe that every $\boldsymbol{p} = (p_1, \dots, p_n) \in \mathrm{NI}(\boldsymbol{A} \to \boldsymbol{C})$ is of the form $p_i = (A_i, B_{I_{\tau(i)}}, C_{\sigma(i)})$ for some $I = \{I_1 < \cdots < I_n\} \in \binom{[\ell]}{n}$ and $\tau, \sigma \in \mathfrak{S}$. In this case $\mathrm{sgn}(\boldsymbol{p}) = \mathrm{sgn}(\sigma)$ and we can write

$\sigma = \rho\tau$ for some $\rho \in \mathfrak{S}_n$, in fact, $\rho = \sigma\tau^{-1}$. Therefore

$$\begin{aligned}
\det L &= \sum_{I \in \binom{[\ell]}{n}} \sum_{\tau \in \mathfrak{S}_n} \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} w(A_i, B_{I_{\tau(i)}}, C_{\sigma(i)}) \\
&= \sum_{I \in \binom{[\ell]}{n}} \sum_{\tau \in \mathfrak{S}_n} \sum_{\rho \in \mathfrak{S}_n} \operatorname{sgn}(\rho\tau) \prod_{i=1}^{n} w(A_i, B_{I_{\tau(i)}}, C_{\rho\tau(i)}) \\
&= \sum_{I \in \binom{[\ell]}{n}} \sum_{\tau \in \mathfrak{S}_n} \operatorname{sgn}(\tau) \left(\prod_{i=1}^{n} w(A_i, B_{I_{\tau(i)}}) \right) \sum_{\rho \in \mathfrak{S}_n} \operatorname{sgn}(\rho) \left(\prod_{i=1}^{n} w(B_{I_{\tau(i)}}, C_{\rho\tau(i)}) \right) \\
&= \sum_{I \in \binom{[\ell]}{n}} \sum_{\tau \in \mathfrak{S}_n} \operatorname{sgn}(\tau) \left(\prod_{i=1}^{n} w(A_i, B_{I_{\tau(i)}}) \right) \sum_{\rho \in \mathfrak{S}_n} \operatorname{sgn}(\rho) \left(\prod_{i=1}^{n} w(B_{I_i}, C_{\rho(i)}) \right) \\
&= \sum_{I \in \binom{[\ell]}{n}} \sum_{\tau \in \mathfrak{S}_n} \operatorname{sgn}(\tau) \left(\prod_{i=1}^{n} M_{i, I_{\tau(i)}} \right) \sum_{\rho \in \mathfrak{S}_n} \operatorname{sgn}(\rho) \left(\prod_{i=1}^{n} N_{I_i, \rho(i)} \right) \\
&= \sum_{I \in \binom{[\ell]}{n}} [M]_{[n], I} [N]_{I, [n]}.
\end{aligned}$$

Since $L = MN$, the proof is completed. □

7.3 Hankel determinants of moments

In this section we compute the Hankel determinants of moments using the Lindström–Gessel–Viennot lemma.

As usual let $\{P_n(x)\}_{n\ge 0}$ be a monic OPS satisfying

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

and let μ_n be the nth moment.

Definition 7.3.1. The (infinite) **Hankel matrix** H of the sequence $\{\mu_n\}_{n\ge 0}$ is defined by

$$H = (\mu_{i+j})_{i,j=0}^{\infty}.$$

The **Hankel determinant** Δ_n is defined by

$$\Delta_n = [H]_{\{0,1,\dots,n\},\{0,1,\dots,n\}} = \det\left(\mu_{i+j}\right)_{i,j=0}^{n} = \det \begin{pmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{pmatrix}.$$

We will use the convention that the determinant of an empty matrix is 1. For example, $\Delta_n = 1$ for $n < 0$.

Note that μ_{i+j} is the generating function for Motzkin paths of length $i + j$ with starting and ending heights 0. Let $\boldsymbol{A} = (A_0, \dots, A_n)$ and $\boldsymbol{B} = (B_0, \dots, _n)$, where $A_i = (-i, 0)$ and $B_i = (i, 0)$. Then

$$\mu_{i+j} = \sum_{p \in \mathrm{Motz}(A_i \to B_j)} \mathrm{wt}(p).$$

Therefore, by Theorem 7.2.3,

$$\Delta_n = \det\left(\mu_{i+j}\right)_{i,j=0}^{n} = \sum_{\boldsymbol{p} \in \mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B})} \operatorname{sgn}(\boldsymbol{p}) \,\mathrm{wt}(\boldsymbol{p}).$$

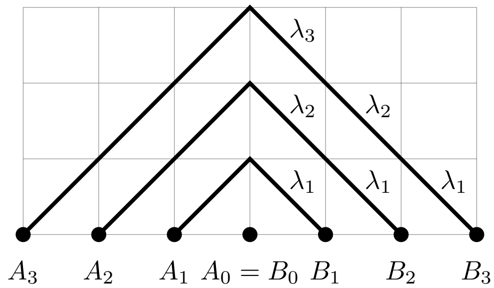

Figure 7.5: The unique nonintersecting $(n+1)$-path in $\mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B})$ for the case $n = 3$.

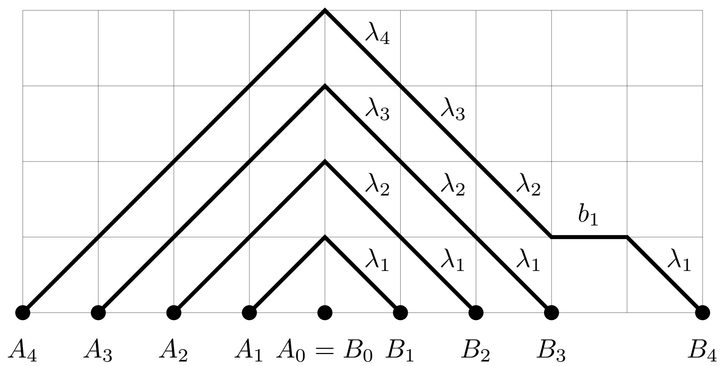

Figure 7.6: A nonintersecting $(n+1)$-path in $\mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B}')$ for the case $n = 4$.

But, $\mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B})$ has a unique $(n+1)$-path $\boldsymbol{p} = (p_0, \dots, p_n)$ as shown in Figure 7.5. Observe that $\mathrm{wt}(\boldsymbol{p}) = \lambda_1^n \lambda_2^{n-1} \cdots \lambda_n^1$. This shows the following theorem.

Theorem 7.3.2. *We have*

$$\Delta_n = \lambda_1^n \lambda_2^{n-1} \cdots \lambda_n^1.$$

Now let's consider the following minor of the Hankel matrix:

$$\Delta_n' := [H]_{\{0,1,\dots,n\},\{0,1,\dots,n-1,n+1\}} = \det(M_{i,j})_{i,j=0}^n,$$

where $M_{i,j} = \mu_{i+j}$ if $j < n$ and $M_{i,n} = \mu_{i+n+1}$. Let $\boldsymbol{A} = (A_0, \dots, A_n)$ and $\boldsymbol{B}' = (B_0, \dots, B_n)$, where $A_i = (-i, 0)$ and $B_i = (i + \delta_{i,n}, 0)$. Then by the Lindström–Gessel–Viennot lemma,

$$\Delta_n' = \sum_{\boldsymbol{p} \in \mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B}')} \mathrm{sgn}(\boldsymbol{p}) \,\mathrm{wt}(\boldsymbol{p}).$$

Suppose $\boldsymbol{p} = (p_0, \dots, p_n) \in \mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B}')$. Then $p_0, \dots, p_{n-1}$ are fixed as in the previous case. Moreover, the first n steps of p_n must be all up steps. Since p_n is a Motzkin path of length $2n+1$, the remaining steps are n down steps and one horizontal step. Thus $\mathrm{wt}(p_n) = \lambda_1 \cdots \lambda_n b_k$ for some $0 \le k \le n$ and p_n is determined by the number k. See Figure 7.6. This shows the following theorem.

Theorem 7.3.3. *We have*

$$\Delta_n' = \lambda_1^n \lambda_2^{n-1} \cdots \lambda_n^1 (b_0 + \cdots + b_n) = \Delta_n (b_0 + \cdots + b_n).$$

Using Theorem 7.3.2 and Theorem 7.3.3 we can express λ_n and b_n using μ_n's as follows.

Corollary 7.3.4. *For $n \geq 1$, we have*

$$\lambda_n = \frac{\Delta_n \Delta_{n-2}}{\Delta_{n-1}^2},$$
$$b_n = \frac{\Delta'_n}{\Delta_n} - \frac{\Delta'_{n-1}}{\Delta_{n-1}},$$

and $b_0 = \Delta'_0/\Delta_0$.

Proof. By Theorem 7.3.2, $\lambda_1 \cdots \lambda_n = \Delta_n/\Delta_{n-1}$. Thus

$$\lambda_n = \frac{\Delta_n}{\Delta_{n-1}} \cdot \left(\frac{\Delta_{n-1}}{\Delta_{n-2}}\right)^{-1} = \frac{\Delta_n \Delta_{n-2}}{\Delta_{n-1}^2}.$$

Similarly, by Theorem 7.3.3, $b_0 + \cdots + b_n = \Delta'_n/\Delta_n$. Thus

$$b_n = \frac{\Delta'_n}{\Delta_n} - \frac{\Delta'_{n-1}}{\Delta_{n-1}}.$$ □

Corollary 7.3.5. *Let $\{\mu_n\}_{n\geq 0}$ be a sequence of numbers. There is an OPS with moments μ_n if and only if $\Delta_n \neq 0$ for all $n \geq 0$.*

Proof. We know that $\{P_n(x)\}_{n\geq 0}$ is an OPS if and only if it satisfies a 3-term recurrence relation

$$P_{n+1}(x) = (x - b_n)P_n(x) - \lambda_n P_{n-1}(x) \tag{7.3.1}$$

with $\lambda_n \neq 0$. Thus, if $\{P_n(x)\}_{n\geq 0}$ is an OPS, then by Theorem 7.3.2, $\Delta_n \neq 0$.

Conversely, if $\Delta_n \neq 0$, then we can construct λ_n and b_n using Corollary 7.3.4. By the construction, λ_n and b_n are the sequences which give $\mu_n = \sum_{p\in \mathrm{Motz}_n} \mathrm{wt}(p)$. Hence, if we define $\{P_n(x)\}_{n\geq 0}$ by (7.3.1), then it is an OPS with moments μ_n. □

For the rest of this section, we consider the case $b_n = 0$. Recall that $\mu_{2n+1} = 0$ for all $n \geq 0$ if and only if $b_n = 0$ for all $n \geq 0$. In this case there is a correspondence between $\{\mu_{2n}\}_{n\geq 0}$ and $\{\lambda_n\}_{n\geq 0}$. For example,

$$\Delta_3 = \det \begin{pmatrix} \mu_0 & \mu_1 & \mu_2 & \mu_3 \\ \mu_1 & \mu_2 & \mu_3 & \mu_4 \\ \mu_2 & \mu_3 & \mu_4 & \mu_5 \\ \mu_3 & \mu_4 & \mu_5 & \mu_6 \end{pmatrix} = \det \begin{pmatrix} \mu_0 & 0 & \mu_2 & 0 \\ 0 & \mu_2 & 0 & \mu_4 \\ \mu_2 & 0 & \mu_4 & 0 \\ 0 & \mu_4 & 0 & \mu_6 \end{pmatrix} = \det \begin{pmatrix} \mu_0 & \mu_2 \\ \mu_2 & \mu_4 \end{pmatrix} \det \begin{pmatrix} \mu_2 & \mu_4 \\ \mu_4 & \mu_6 \end{pmatrix},$$

$$\Delta_4 = \det \begin{pmatrix} \mu_0 & \mu_1 & \mu_2 & \mu_3 & \mu_4 \\ \mu_1 & \mu_2 & \mu_3 & \mu_4 & \mu_5 \\ \mu_2 & \mu_3 & \mu_4 & \mu_5 & \mu_6 \\ \mu_3 & \mu_4 & \mu_5 & \mu_6 & \mu_7 \\ \mu_4 & \mu_5 & \mu_6 & \mu_7 & \mu_8 \end{pmatrix} = \det \begin{pmatrix} \mu_0 & 0 & \mu_2 & 0 & \mu_4 \\ 0 & \mu_2 & 0 & \mu_4 & 0 \\ \mu_2 & 0 & \mu_4 & 0 & \mu_6 \\ 0 & \mu_4 & 0 & \mu_6 & 0 \\ \mu_4 & 0 & \mu_6 & 0 & \mu_8 \end{pmatrix}$$
$$= \det \begin{pmatrix} \mu_0 & \mu_2 & \mu_4 \\ \mu_2 & \mu_4 & \mu_6 \\ \mu_4 & \mu_6 & \mu_8 \end{pmatrix} \det \begin{pmatrix} \mu_2 & \mu_4 \\ \mu_4 & \mu_6 \end{pmatrix}.$$

In general, we have

$$\Delta_{2n} = \Delta_n(2)\Delta^+_{n-1}(2), \tag{7.3.2}$$
$$\Delta_{2n+1} = \Delta_n(2)\Delta^+_n(2), \tag{7.3.3}$$

where

$$\Delta_n(2) = \det\left(\mu_{2i+2j}\right)_{i,j=0}^n,$$
$$\Delta^+_n(2) = \det\left(\mu_{2i+2j+2}\right)_{i,j=0}^n,$$

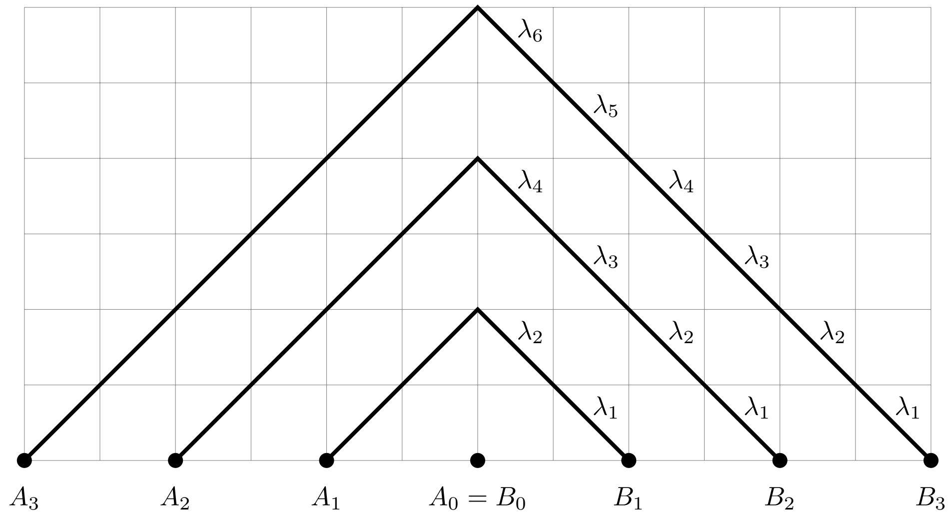

Figure 7.7: The unique nonintersecting $(n+1)$-path in $\mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B})$ for the case $n = 3$.

Since $b_n = 0$, we have

$$\mu_{2n} = \sum_{p \in \mathrm{Dyck}_{2n}} \mathrm{wt}(p).$$

Let $\boldsymbol{A} = (A_0, \ldots, A_n)$ and $\boldsymbol{B} = (B_0, \ldots, B_n)$, where $A_i = (-2i, 0)$ and $B_i = (2i, 0)$. Then

$$\mu_{2i+2j} = \sum_{p \in \mathrm{Dyck}(A_i \to B_j)} \mathrm{wt}(p).$$

Thus, by the Lindström–Gessel–Viennot lemma,

$$\Delta_n(2) = \sum_{\boldsymbol{p} \in \mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B})} \mathrm{sgn}(\boldsymbol{p}) \, \mathrm{wt}(\boldsymbol{p}).$$

There is only one element in $\mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B})$ as shown in Figure 7.7. This gives the following theorem.

Theorem 7.3.6. *If $b_n = 0$ for all $n \geq 0$, we have*

$$\Delta_n(2) = (\lambda_1\lambda_2)^n(\lambda_3\lambda_4)^{n-1} \cdots (\lambda_{2n-1}\lambda_{2n})^1.$$

Now let $\boldsymbol{A} = (A_0, \ldots, A_n)$ and $\boldsymbol{B}^+ = (B_0^+, \ldots, B_n^+)$, where $A_i = (-2i, 0)$ and $B_i^+ = (2i+2, 0)$. Then

$$\mu_{2i+2j} = \sum_{p \in \mathrm{Dyck}(A_i \to B_j)} \mathrm{wt}(p).$$

Thus, by the Lindström–Gessel–Viennot lemma,

$$\Delta_n(2) = \sum_{\boldsymbol{p} \in \mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B})} \mathrm{sgn}(\boldsymbol{p}) \, \mathrm{wt}(\boldsymbol{p}).$$

There is only one element in $\mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B})$ as shown in Figure 7.7. This gives the following theorem.

Theorem 7.3.7. *If $b_n = 0$ for all $n \geq 0$, we have*

$$\Delta_n^+(2) = \lambda_1^{n+1}(\lambda_2\lambda_3)^n(\lambda_4\lambda_5)^{n-1} \cdots (\lambda_{2n}\lambda_{2n+1})^1.$$

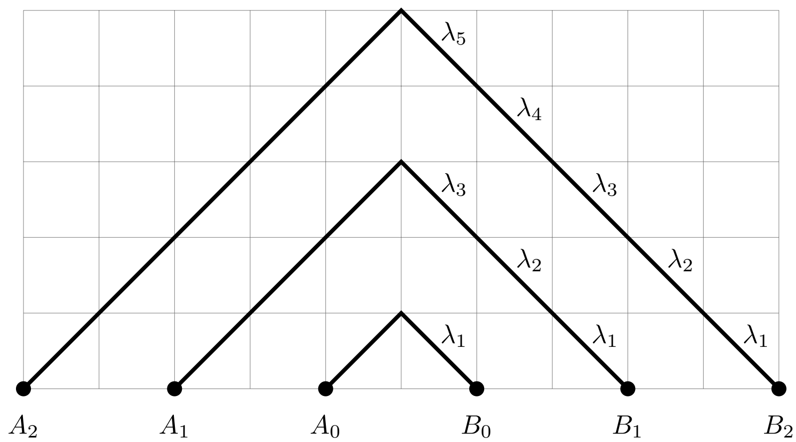

Figure 7.8: The unique nonintersecting $(n+1)$-path in $\mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B}^+)$ for the case $n = 2$.

Corollary 7.3.8. *If $b_n = 0$ for all $n \geq 0$, we have*

$$\lambda_{2n} = \frac{\Delta_n(2)\Delta^+_{n-2}(2)}{\Delta_{n-1}(2)\Delta^+_{n-1}(2)},$$

$$\lambda_{2n+1} = \frac{\Delta^+_n(2)\Delta_{n-1}(2)}{\Delta_n(2)\Delta^+_{n-1}(2)}.$$

Proof. By Theorem 7.3.6 and 7.3.7,

$$\frac{\Delta^+_n(2)}{\Delta_n(2)} = \lambda_1\lambda_3\cdots\lambda_{2n+1}, \qquad \frac{\Delta_n(2)}{\Delta^+_{n-1}(2)} = \lambda_2\lambda_4\cdots\lambda_{2n}.$$

Thus

$$\lambda_{2n+1} = \frac{\Delta^+_n(2)}{\Delta_n(2)}\left(\frac{\Delta^+_{n-1}(2)}{\Delta_{n-1}(2)}\right)^{-1} = \frac{\Delta^+_n(2)\Delta_{n-1}(2)}{\Delta_n(2)\Delta^+_{n-1}(2)},$$

$$\lambda_{2n} = \frac{\Delta_n(2)}{\Delta^+_{n-1}(2)}\left(\frac{\Delta_{n-1}(2)}{\Delta^+_{n-2}(2)}\right)^{-1} = \frac{\Delta_n(2)\Delta^+_{n-2}(2)}{\Delta_{n-1}(2)\Delta^+_{n-1}(2)}. \qquad \square$$

7.4 Another duality between moments and coefficients

Recall that we proved a duality between mixed moments $\mu_{n,k}$ and coefficients $\nu_{n,k}$. In this section we will prove the following theorem, which gives another type of duality between moments and coefficients.

Theorem 7.4.1. *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*

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

Let us write

$$P_n(x) = \sum_{k=0}^{n} \nu_{n,k}x^k.$$

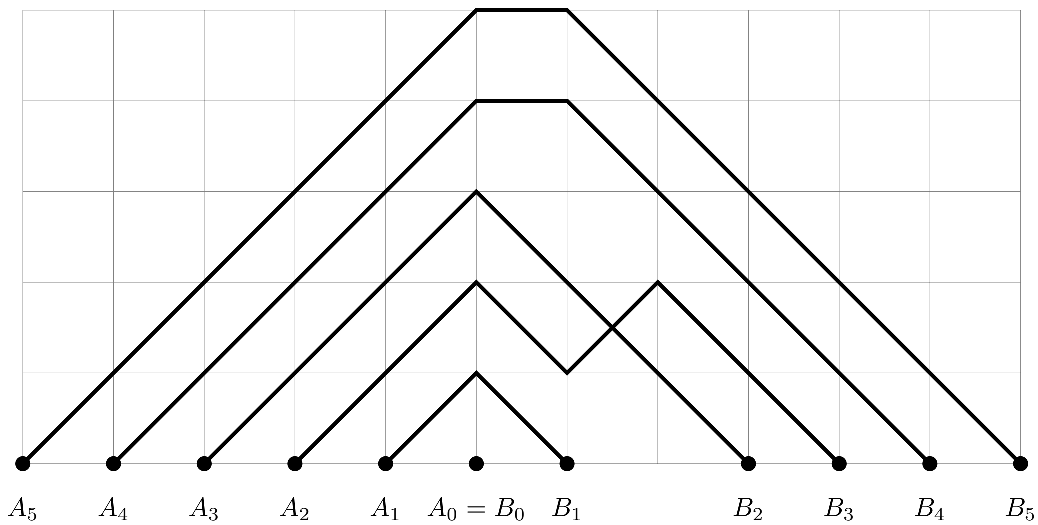

Figure 7.9: A nonintersecting $(n+1)$-path in $\mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B}^{(k)})$ for the case $n = 6$ and $k = 2$.

Then the theorem above is equivalent to

$$\nu_{n,k} = \frac{(-1)^{n-k}}{\Delta_{n-1}} [H]_{\{0,\dots,n-1\},\{0,\dots,k-1,k+1,\dots,n\}}. \tag{7.4.1}$$

Thus we need to compute

$$[H]_{\{0,\dots,n-1\},\{0,\dots,k-1,k+1,\dots,n\}} = \det(\mu_{i+j+\chi(j\ge k)})_{i,j=0}^{n-1},$$

where, for a statement P, $\chi(P) = 1$ if P is true and $\chi(P) = 0$ otherwise. Let $\boldsymbol{A} = (A_0, \dots, A_{n-1})$ and $\boldsymbol{B}^{(k)} = (B_0, \dots, B_{n-1})$, where

$$A_i = (-i, 0), \qquad B_i = (i + \chi(i \ge k), 0).$$

Then by the Lindström–Gessel–Viennot lemma,

$$[H]_{\{0,\dots,n-1\},\{0,\dots,k-1,k+1,\dots,n\}} = \sum_{\boldsymbol{p}\in \mathrm{NI}(\boldsymbol{A}\to\boldsymbol{B}^{(k)})} \mathrm{sgn}(\boldsymbol{p}) \,\mathrm{wt}(\boldsymbol{p}). \tag{7.4.2}$$

Now we investigate a nonintersecting $(n+1)$-path $\boldsymbol{p} = (p_0, \dots, p_{n-1}) \in \mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B}^{(k)})$, see Figure 7.9 for an example. For $1 \le i \le n-1$, let R_i be the region defined by

$$R_i = \{(x, y) \in \mathbb{R}^2 : x > 0, i-1 < y < i\}.$$

Then we have the following observations. See Figure 7.10.

(1) The first i steps of p_i are up steps. Hence p_i visits $(0, i)$.

(2) There are $n-i$ paths having at least one step in R_i, namely, $p_i, p_{i+1}, \dots, p_{n-1}$.

(3) If d and u are the number of down steps and up steps in R_i, respectively, then $d-u = n-i$. This is because each of $p_i, p_{i+1}, \dots, p_{n-1}$ must enter the line $y = i$ and exit the line $y = i-1$, contributing 1 to the number $d-u$.

(4) In R_i, every down step is of the form $(a, i) \to (a-1, i-1)$ for some $0 \le a \le n-i$. So there are $n-i+1$ possible down steps in R_i.

(5) In R_i, there is only one "missing" down step or there is only one figure "X", but not both.

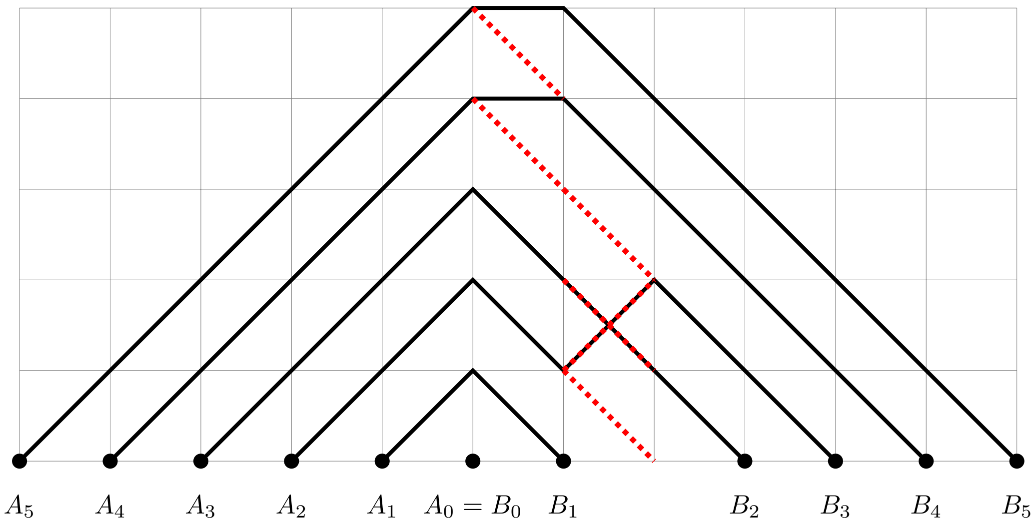

Figure 7.10: The missing down steps and the figure "X" in Figure 7.9 are drawn with red dotted lined.

(6) In R_1, we have exactly one of the following:

- $(k-1,1) \to (k,0)$ is missing;
- $(k,1) \to (k+1,0)$ is missing;
- $(k,1) \to (k+1,0)$ and $(k,0) \to (k+1,1)$ form a figure "X".

To see this, observe that if $(k-1,1) \to (k,0)$ presents, then the path containing this down step must have a horizontal step $(k,0) \to (k+1,0)$ or an up step $(k,0) \to (k+1,1)$ because $(k,0)$ cannot be an ending point. Therefore in this case either $(k,1) \to (k+1,0)$ is missing or $(k,1) \to (k+1,0)$ and $(k,0) \to (k+1,1)$ form a figure "X".

(7) If $(a,i) \to (a+1,i-1)$ is a missing down step, then exactly one of the following holds:

- $(a-1,i+1) \to (a,i)$ is missing;
- $(a,i+1) \to (a+1,i)$ is missing. In this case, there is a path p_j containing a horizontal step $(a,i) \to (a+1,i)$;
- $(a,i+1) \to (a+1,i)$ and $(a,i) \to (a+1,i+1)$ form a figure "X".

(8) If $(a,i) \to (a+1,i-1)$ forms a figure "X", both $(a,i+1) \to (a+1,i)$ and $(a,i-1) \to (a+1,i-2)$ are missing.

(9) Construct a path p' from $C := (k+1/2,-1/2)$ to $D := (1/2,n-1/2)$ by connecting the midpoints of all the missing down steps as shown in Figure 7.11. Then p' consists of northeast steps $(-1,1)$, north steps $(0,1)$, and double north steps $(0,2)$. This can be seen by the observations (6), (7), and (8).

(10) The map $\boldsymbol{p} \mapsto p'$ is a bijection from $\mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B}^{(k)})$ to all such paths $p' : C \to D$.

Lemma 7.4.2. *Let p' be the path obtained from $\boldsymbol{p} = (p_0, \ldots, p_{n-1}) \in \mathrm{NI}(\boldsymbol{A} \to \boldsymbol{B}^{(k)})$ as above. Define $w(p')$ to be the product of b_i for each north step intersecting $y = i$ and $-\lambda_i$ for each double north step intersecting $y = i$ and $y = i-1$. Then*

$$\mathrm{sgn}(\boldsymbol{p})\,\mathrm{wt}(\boldsymbol{p}) = w(p')\Delta_{n-1}.$$

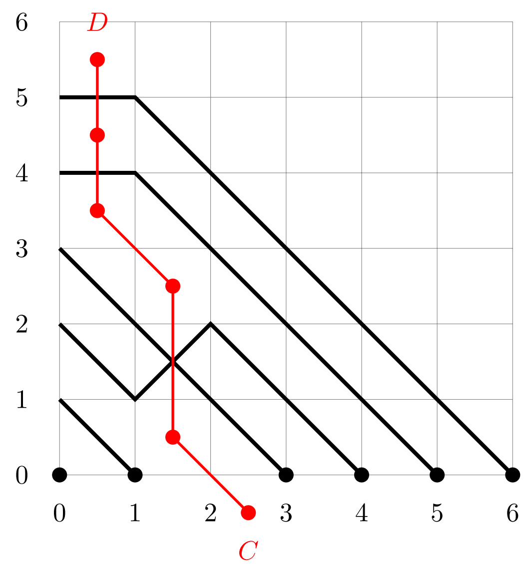

Figure 7.11: The path p' from $C = (k+1/2, -1/2)$ to $D = (1/2, n-1/2)$ obtained by connecting the midpoints of the missing down steps in Figure 7.10.

Proof. By observation (7), every north step intersecting $y = i$ corresponds to a horizontal step of height i, which has weight b_i. By observation (5), the product of the weights of all down steps in $\boldsymbol{p}$ is

$$\lambda_1^{n-1}\lambda_2^{n-2}\cdots\lambda_{n-1}^{1}X_1\cdots X_{n-1} = \Delta_{n-1}X_1\cdots X_{n-1},$$

where $X_i = \lambda_i$ if there is a figure "X" in R_i and $X_i = 1$ otherwise. Moreover, every figure "X" contributes -1 to the sign of $\boldsymbol{p}$. Combinining these facts, we obtain the lemma. $\square$

Lemma 7.4.3. *We have*

$$\sum_{p':C\to D} w(p') = (-1)^{n-k}\nu_{n,k}.$$

Proof. A path $p' : C \to D$ is essentially the same as a Favard tiling, say T. Since $p' : C \to D$ has $n-k$ northeast steps, T has k red monominos. We need to check their signs. Suppose p' has r north steps and s double north steps. Then $r + 2s + k = n$ since the height difference between C and D is n. The sign of p' is $(-1)^s$ while the sign of T is $(-1)^{r+s}$. Hence the sign difference of p' and T is $(-1)^r = (-1)^{n-2s-k} = (-1)^{n-k}$. Thus we have

$$\sum_{p':C\to D} w(p') = \sum_{T\in \mathrm{FT}_{n,k}} (-1)^{n-k}w(T) = (-1)^{n-k}\nu_{n,k},$$

as desired. $\square$

We are now ready to prove Theorem 7.4.1.

Proof of Theorem 7.4.1. By Lemma 7.4.2, Lemma 7.4.3, and (7.4.2),

$$\begin{aligned}\nu_{n,k} &= (-1)^{n-k}\sum_{p':C\to D} w(p') = (-1)^{n-k}\sum_{\boldsymbol{p}\in \mathrm{NI}(\boldsymbol{A}\to\boldsymbol{B}^{(k)})}\frac{\mathrm{sgn}(\boldsymbol{p})\,\mathrm{wt}(\boldsymbol{p})}{\Delta_{n-1}}\\ &= \frac{(-1)^{n-k}}{\Delta_{n-1}}[H]_{\{0,\dots,n-1\},\{0,\dots,k-1,k+1,\dots,n\}},\end{aligned}$$

which is equivalent to the statement of the theorem. $\square$

Theorem 7.4.1 has an interesting connection with the following well-known lemma. (In fact, we only need the case that $|I| = |J| = 1$, which is equivalent to the inverse matrix formula using cofactors.)

Lemma 7.4.4 (Inverse minor formula)**.** *Suppose that $A = (A_{i,j})_{i,j=0}^n$ is an invertible matrix. For subsets $I, J \subseteq \{0, \dots, n\}$ of the same cardinality, we have*

$$\left[A^{-1}\right]_{I,J} = (-1)^{\|I\|+\|J\|} \frac{[A]_{J',I'}}{\det A},$$

where $\|I\| = \sum_{i \in I} i$ and $I' = \{0, \dots, n\} \setminus I$.

Let $A = (\mu_{i+j})_{i,j=0}^n$. Note that

$$\det A = \Delta_n = \Delta_{n-1} \lambda_1 \cdots \lambda_n.$$

Thus Theorem 7.4.1 can be restated as

$$\nu_{n,k} = \frac{(-1)^{n-k} \lambda_1 \cdots \lambda_n}{\det A} [A]_{\{0,\dots,n\}\setminus\{n\},\{0,\dots,n\}\setminus\{k\}}.$$

By Lemma 7.4.4 with $I = \{k\}$ and $J = \{n\}$, we have

$$\nu_{n,k} = \lambda_1 \cdots \lambda_n \cdot (-1)^{\|I\|+\|J\|} \frac{[A]_{J',I'}}{\det A} = \lambda_1 \cdots \lambda_n [A^{-1}]_{I,J} = \lambda_1 \cdots \lambda_n (A^{-1})_{k,n}.$$

This implies that $\nu_{n,0}, \nu_{n,1}, \dots, \nu_{n,n}$ are the entries of the last column of the matrix A^{-1}. Equivalently, for $0 \le i \le n$, we have

$$\sum_{k=0}^n \mu_{i+k} \nu_{n,k} = \delta_{n,i} \lambda_1 \cdots \lambda_n.$$

Note that this is equivalent to the orthogonality relation

$$\mathcal{L}(x^i P_n(x)) = \delta_{n,i} \lambda_1 \cdots \lambda_n.$$

The above arguments show that Theorem 7.4.1 is equivalent to a formula for $(A^{-1})_{k,n}$ for $0 \le k \le n$. Therefore it is natural to ask whether there is a similar formula for $(A^{-1})_{r,s}$ for arbitrary $0 \le r, s \le n$.

Theorem 7.4.5. *Let $A = (\mu_{i+j})_{i,j=0}^n$. Then*

$$(A^{-1})_{r,s} = \sum_{k=0}^n \frac{\nu_{k,r} \nu_{k,s}}{\lambda_1 \cdots \lambda_k}.$$

Proof 1. Let $I = \{r\}$ and $J = \{s\}$. Then

$$(A^{-1})_{r,s} = (-1)^{r+s} \frac{[A]_{J',I'}}{\det A}. \tag{7.4.3}$$

Let $\boldsymbol{A}^{(s)} = (A_0, \dots, A_{n-1})$ and $\boldsymbol{B}^{(r)} = (B_0, \dots, B_{n-1})$, where

$$A_i = (-i - \chi(i \ge s), 0), \qquad B_i = (i + \chi(i \ge r), 0).$$

Then by the Lindström–Gessel–Viennot lemma,

$$[A]_{J',I'} = \sum_{\boldsymbol{p} \in \mathrm{NI}(\boldsymbol{A}^{(s)} \to \boldsymbol{B}^{(r)})} \operatorname{sgn}(\boldsymbol{p}) \operatorname{wt}(\boldsymbol{p}).$$

Consider $\boldsymbol{p} = (p_0, \dots, p_{n-1}) \in \mathrm{NI}(\boldsymbol{A}^{(s)} \to \boldsymbol{B}^{(r)})$. Then these paths must visit $(0, i)$ for all $0 \le i \le n$ with $i \ne k$ for some k such that $\max(r, s) \le k \le n$. By the same arguments above, we have

$$[A]_{J',I'} = \sum_{k=\max(r,s)}^n \frac{\Delta_n}{\lambda_1 \cdots \lambda_k} (-1)^{k-r} \nu_{k,r} (-1)^{k-s} \nu_{k,s}. \tag{7.4.4}$$

For example, see Figure 7.12. By (7.4.3) and (7.4.4) we obtain the theorem. □

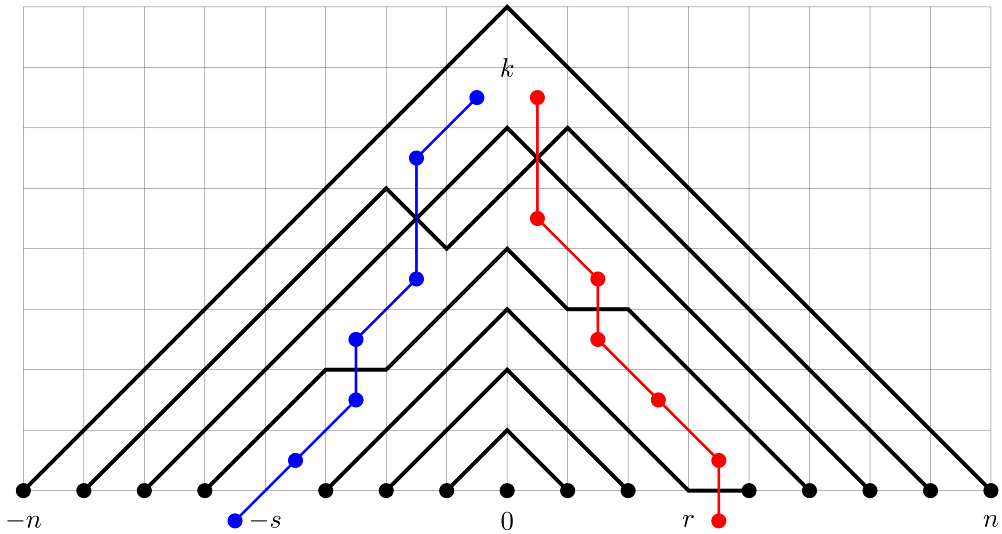

Figure 7.12: A nonintersecting $(n+1)$-path in $\mathrm{NI}(\boldsymbol{A}^{(s)} \to \boldsymbol{B}^{(r)})$. The red (resp. blue) path on the left contributes to $(-1)^{k-s}\nu_{k,s}$ (resp. $(-1)^{k-r}\nu_{k,r}$).

Proof 2. We can also prove this theorem using orthogonality. Given, r,s, it suffces to prove the validity of

$$\delta_{r,s} = \sum_{k=0}^{n} \mu_{r+k}(A^{-1})_{k,s} = \sum_{k=0}^{n} \mu_{r+k} \sum_{\ell=0}^{n} \frac{\nu_{\ell,k}\nu_{\ell,s}}{\lambda_1\cdots\lambda_\ell}.$$

We can write the right-hand side as

$$\begin{aligned}
\mathcal{L}\left(\sum_{k=0}^{n} x^{r+k} \sum_{\ell=0}^{n} \frac{\nu_{\ell,k}\nu_{\ell,s}}{\lambda_1\cdots\lambda_\ell}\right) &= \mathcal{L}\left(x^r \sum_{\ell=0}^{n} \frac{\nu_{\ell,s}}{\lambda_1\cdots\lambda_\ell} \sum_{k=0}^{n} \nu_{\ell,k}x^k\right) \\
&= \mathcal{L}\left(\sum_{a=0}^{r} \mu_{r,a}P_a(x) \sum_{\ell=0}^{n} \frac{\nu_{\ell,s}}{\lambda_1\cdots\lambda_\ell} P_\ell(x)\right) \\
&= \sum_{\ell=0}^{n} \mu_{r,\ell}\nu_{\ell,s} = \delta_{r,s}. \qquad \square
\end{aligned}$$

Note that Theorem 7.4.5 is equivalent to the following UDL decomposition of the inverse of a finite Hankel matrix:

$$\left((\mu_{i+j})_{i,j=0}^{n}\right)^{-1} = \left((\nu_{i,j})_{i,j=0}^{n}\right)^{T} D(\lambda_1^{-1}, \lambda_1^{-1}\lambda_2^{-1}, \dots, \lambda_1^{-1}\cdots\lambda_n^{-1})\,(\nu_{i,j})_{i,j=0}^{n}\,, \tag{7.4.5}$$

where $D(a_0,\dots,a_n)$ is the diagonal matrix with diagonal entries a_i's. Since $(\nu_{i,j})^{-1} = (\mu_{i,j})$, taking the inverses of both sides of (7.4.5), we get the following LDU decomposition of a finite Hankel matrix:

$$(\mu_{i+j})_{i,j=0}^{n} = (\mu_{i,j})_{i,j=0}^{n}\, D(\lambda_1, \lambda_1\lambda_2, \dots, \lambda_1\cdots\lambda_n) \left((\mu_{i,j})_{i,j=0}^{n}\right)^{T}. \tag{7.4.6}$$

This is equivalent to

$$\mu_{r+s} = \sum_{k=0}^{n} \mu_{r,k}\lambda_1\cdots\lambda_k\mu_{s,k}.$$

This can also be proved combinatorially using Theorem 4.3.3 as follows:

$$\sum_{k=0}^{n} \mu_{r,k}\lambda_1\cdots\lambda_k\mu_{s,k} = \sum_{k=0}^{n} \mu_{r,0,k}\lambda_1\cdots\lambda_k\mu_{s,0,k} = \sum_{k=0}^{n} \mu_{r,0,k}\mu_{s,k,0} = \mu_{r+s}.$$

Chapter 8

Continued fractions

Mathematicians introduced orthogonal polynomials in order to study continued fractions. In this chapter we first review basics of continued fractions. We then investigate the close connection with continued fractions and orthogonal polynomials.

8.1 Basics of continued fractions

Let's begin with the following example.

Example 8.1.1. What is the following value?

$$\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{\ddots}}}$$

To solve this, let x be this number. Then we must have

$$x = \frac{1}{1+x},$$

or equivalently $x^2 + x - 1 = 0$. Thus $x = \frac{-1\pm\sqrt{5}}{2}$. Since x must be positive, we obtain $x = \frac{\sqrt{5}-1}{2}$.

If you believe that the above solution is correct, let's look at another example.

Example 8.1.2. What is the following value?

$$\cfrac{1}{1-\cfrac{3}{1-\cfrac{3}{\ddots}}}$$

As before, let x be this number. Then we must have

$$x = \frac{1}{1-3x},$$

or equivalently $3x^2 - x + 1 = 0$. Thus $x = \frac{-1\pm\sqrt{11}i}{2}$. However, since x must be real, there is no solution.

What's the difference between the two examples above? To answer this question, we need to be precise on what a continued fraction actually means.

Definition 8.1.3. A **continued fraction** is an expression of the form

$$\beta_0 + \cfrac{\alpha_1}{\beta_1 + \cfrac{\alpha_2}{\beta_2 + \cfrac{\alpha_3}{\ddots}}},$$

for some sequences $\{\alpha_n\}_{n\ge 1}$ and $\{\beta_n\}_{n\ge 0}$. The **n-th convergent** is

$$C_n = \beta_0 + \cfrac{\alpha_1}{\beta_1 + \cfrac{\alpha_2}{\beta_2 + \ddots + \cfrac{\alpha_n}{\beta_n}}}$$

We write

$$\beta_0 + \cfrac{\alpha_1}{\beta_1 + \cfrac{\alpha_2}{\beta_2 + \cfrac{\alpha_3}{\ddots}}} = L$$

to mean

$$\lim_{n\to\infty} C_n = L.$$

Note that the argument in Example 8.1.1 is not complete because it did not check whether the sequence of nth convergents actually converges. Indeed, this sequence converges, so the answer was correct. However, the sequence of nth convergents does not converge in Example 8.1.2, hence we arrived at the strange situation that the continued fraction is a complex number.

8.2 Flajolet's combinatorial theory of continued fractions

In this section we study Flajolet's combinatorial theory of continued fractions. Throughout this section a "Motzkin path" means a Motzkin path whose starting and endig height are both 0.

Fix sequences $\boldsymbol{b} = (b_0, b_1, \dots)$ and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots)$. We define

$$\mu_n(\boldsymbol{b}, \boldsymbol{\lambda}) = \sum_{\pi\in \mathrm{Motz}_n} \mathrm{wt}(\pi; \boldsymbol{b}, \boldsymbol{\lambda}).$$

Here $\mathrm{wt}(\pi; \boldsymbol{b}, \boldsymbol{\lambda})$ is the usual weight of a Motzkin path π, i.e., the product of b_i for each horizontal step with starting height i and b_i for each down step with starting height i.

Let

$$F(x; \boldsymbol{b}, \boldsymbol{\lambda}) = \sum_{n\ge 0} \mu_n(\boldsymbol{b}, \boldsymbol{\lambda}) x^n.$$

Observe that every Motzkin path π is one of the following forms:

(1) π is an empty path.

(2) $\pi = H\pi'$ for some Motzkin path π',

(3) $\pi = U\pi' D\pi''$ for some Motzkin paths π' and π''.

Let δ be the operator that removes the first element of a sequence, namely, $\delta\boldsymbol{b} = (b_1, b_2, \dots)$ and $\delta\boldsymbol{\lambda} = (\lambda_2, \lambda_3, \dots)$. Then in the second case $\pi = H\pi'$, we have

$$\mathrm{wt}(\pi; \boldsymbol{b}, \boldsymbol{\lambda}) = b_0\, \mathrm{wt}(\pi'; \boldsymbol{b}, \boldsymbol{\lambda}),$$

and in the third case $\pi = U\pi' D\pi''$, we have

$$\mathrm{wt}(\pi; \boldsymbol{b}, \boldsymbol{\lambda}) = \lambda_1 \,\mathrm{wt}(\pi'; \delta\boldsymbol{b}, \delta\boldsymbol{\lambda})\,\mathrm{wt}(\pi''; \boldsymbol{b}, \boldsymbol{\lambda}).$$

Therefore, we have

$$F(x; \boldsymbol{b}, \boldsymbol{\lambda}) = 1 + b_0 x F(x; \boldsymbol{b}, \boldsymbol{\lambda}) + \lambda_1 x^2 F(x; \delta\boldsymbol{b}, \delta\boldsymbol{\lambda}) F(x; \boldsymbol{b}, \boldsymbol{\lambda}),$$

or equivalently,

$$F(x; \boldsymbol{b}, \boldsymbol{\lambda}) = \frac{1}{1 - b_0 x - \lambda_1 x^2 F(x; \delta\boldsymbol{b}, \delta\boldsymbol{\lambda})}.$$

Iterating this process, we obtain

$$\begin{aligned} F(x; \boldsymbol{b}, \boldsymbol{\lambda}) &= \cfrac{1}{1 - b_0 x - \cfrac{\lambda_1 x^2}{1 - b_1 x - \lambda_2 x^2 F(x; \delta\boldsymbol{b}, \delta\boldsymbol{\lambda})}} \\ &= \cdots = \cfrac{1}{1 - b_0 x - \cfrac{\lambda_1 x^2}{1 - b_1 x - \cfrac{\lambda_2 x^2}{\ddots}}}. \end{aligned} \tag{8.2.1}$$

However, we must make this argument rigorous by showing that the above continued fraction converges. Since we are working with formal power series we need to define convergence of a sequence of formal power series.

Definition 8.2.1. Let $\{F_n(x)\}_{n\ge 0}$ be a sequence of formal power series. We write

$$\lim_{n\to\infty} F_n(x) = F(x)$$

if for every $m \ge 0$, there is $N > 0$ such that for all $n > N$,

$$[x^m]F_n(x) = [x^m]F(x),$$

where $[x^m]F(x)$ is the coefficient of x^m in $F(x)$. In other words, for any $m \ge 0$, the coefficient of x^m in $F_n(x)$, $n \ge 0$, eventually becomes the coefficient of x^m in $F(x)$.

Example 8.2.2. Let $F(x) = 1+x+x^2+\cdots$ and $F_k(x) = 1+x+x^2+\cdots+x^k$. Then $\lim_{k\to\infty} F_k(x) = F(x)$.

To prove that Equation (8.2.1) converges we need to investigate its nth convergents. To do this we introduce the following definition. Let $\mathrm{Motz}_n^{\le k}$ denote the set of Motzkin paths in Motz_n that stays weakly below the height $y = k$. We also define

$$\mu_n^{\le k}(\boldsymbol{b}, \boldsymbol{\lambda}) = \sum_{\pi \in \mathrm{Motz}_n^{\le k}} \mathrm{wt}(\pi; \boldsymbol{b}, \boldsymbol{\lambda}).$$

Theorem 8.2.3. *For any $k \ge 0$, we have*

$$\sum_{n\ge 0} \mu_n^{\le k}(\boldsymbol{b}, \boldsymbol{\lambda}) x^n = \cfrac{1}{1 - b_0 x - \cfrac{\lambda_1 x^2}{1 - b_1 x - \cfrac{\lambda_2 x^2}{1 - b_2 x - \ddots - \cfrac{\lambda_k x^2}{1 - b_k x}}}}.$$

Proof. We use induction on k. If $k = 0$, then the identity we need to prove is

$$\sum_{n\geq 0} \mu_n^{\leq 0}(\boldsymbol{b}, \boldsymbol{\lambda})x^n = \frac{1}{1 - b_0 x}.$$

Since there is only one Motzkin path in $\mathrm{Motz}_n^{\leq 0}$, which consists of n horizontal steps, we have $\mu_n^{\leq 0}(\boldsymbol{b}, \boldsymbol{\lambda}) = b_0^n$.

Now let $k \geq 1$ and suppose the theorem holds for $k-1$. Then we can apply the decomposition of a Motzkin path π into $\emptyset$, $\pi = H\pi'$, or $\pi = U\pi' D\pi''$. Writing

$$F_k(x; \boldsymbol{b}, \boldsymbol{\lambda}) = \sum_{n\geq 0} \mu_n^{\leq k}(\boldsymbol{b}, \boldsymbol{\lambda})x^n,$$

the decomposition shows that

$$F_k(x; \boldsymbol{b}, \boldsymbol{\lambda}) = 1 + b_0 x F_k(x; \boldsymbol{b}, \boldsymbol{\lambda}) + \lambda_1 x^2 F_{k-1}(x; \delta\boldsymbol{b}, \delta\boldsymbol{\lambda}) F_k(x; \boldsymbol{b}, \boldsymbol{\lambda}).$$

Therefore

$$F_k(x; \boldsymbol{b}, \boldsymbol{\lambda}) = \frac{1}{1 - b_0 x - \lambda_1 x^2 F_{k-1}(x; \delta\boldsymbol{b}, \delta\boldsymbol{\lambda})}.$$

Then by the induction hypothesis, we obtain that the theorem holds for k. Hence, by induction, the proof is completed. □

Now we can give a rigorous proof of Equation (8.2.1).

Corollary 8.2.4. *We have*

$$\sum_{n\geq 0} \mu_n(\boldsymbol{b}, \boldsymbol{\lambda})x^n = \cfrac{1}{1 - b_0 x - \cfrac{\lambda_1 x^2}{1 - b_1 x - \cfrac{\lambda_2 x^2}{\ddots}}}. \tag{8.2.2}$$

Proof. By Theorem 8.2.3, the kth convergent of the continued fraction is $\sum_{n\geq 0} \mu_n^{\leq k}(\boldsymbol{b}, \boldsymbol{\lambda})x^n$. Therefore it suffices to show that

$$\lim_{k\to\infty} \mu_n^{\leq k}(\boldsymbol{b}, \boldsymbol{\lambda}) = \mu_n(\boldsymbol{b}, \boldsymbol{\lambda}). \tag{8.2.3}$$

Observe that any Motzkin path $\pi \in \mathrm{Motz}_n$ can have height at most $\lfloor n/2 \rfloor$. Hence, if $k > \lfloor n/2 \rfloor$, then $\mathrm{Motz}_n^{\leq k} = \mathrm{Motz}_n$. This shows (8.2.3) and we obtain the result. □

8.3 Continued fractions and orthogonal polynomials

In this section we study the connection between continued fractions and orthogonal polynomials. More precisely, we will show that the kth convergent of the continued fraction in (8.2.2) can be written in terms of orthogonal polynomials.

Let $\{P_n(x; \boldsymbol{b}, \boldsymbol{\lambda})\}_{n\geq 0}$ be the monic OPS given by

$$P_{n+1}(x; \boldsymbol{b}, \boldsymbol{\lambda}) = (x - b_n)P_n(x; \boldsymbol{b}, \boldsymbol{\lambda}) - \lambda_n P_{n-1}(x; \boldsymbol{b}, \boldsymbol{\lambda}). \tag{8.3.1}$$

For simplicity, we will write $\delta P_n(x; \boldsymbol{b}, \boldsymbol{\lambda}) = P_n(x; \delta\boldsymbol{b}, \delta\boldsymbol{\lambda})$. Note that $\mu_n(\boldsymbol{b}, \boldsymbol{\lambda})$ is the nth moment of this OPS.

The **inverted polynomial** $P_n^*(x; \boldsymbol{b}, \boldsymbol{\lambda})$ of $P_n(x; \boldsymbol{b}, \boldsymbol{\lambda})$ is defined by

$$P_n^*(x; \boldsymbol{b}, \boldsymbol{\lambda}) = x^n P_n(x^{-1}; \boldsymbol{b}, \boldsymbol{\lambda}).$$

Note that (8.3.1) implies

$$P_{n+1}^*(x; \boldsymbol{b}, \boldsymbol{\lambda}) = (1 - b_n x)P_n^*(x; \boldsymbol{b}, \boldsymbol{\lambda}) - \lambda_n x^2 P_{n-1}^*(x; \boldsymbol{b}, \boldsymbol{\lambda}). \tag{8.3.2}$$

The goal of this section is to prove the following theorem.

Theorem 8.3.1. *We have*

$$\sum_{n\geq 0} \mu_n^{\leq k}(\boldsymbol{b},\boldsymbol{\lambda})x^n = \frac{\delta P_k^*(x;\boldsymbol{b},\boldsymbol{\lambda})}{P_{k+1}^*(x;\boldsymbol{b},\boldsymbol{\lambda})}.$$

To prove this theorem we will use a continued fraction technique due to John Wallis (1616 – 1703). Let $\{\alpha_n\}_{n\geq 1}$ and $\{\beta_n\}_{n\geq 0}$ be fixed sequences and consider

$$C_n = \beta_0 + \cfrac{\alpha_1}{\beta_1 + \cfrac{\alpha_2}{\beta_2 + \ddots + \cfrac{\alpha_n}{\beta_n}}}$$

We can express C_n and a rational function

$$C_n = \frac{A_n}{B_n}. \tag{8.3.3}$$

For example,

$$C_0 = \beta_0 = \frac{\beta_0}{1} = \frac{A_0}{B_0}, \qquad C_1 = \beta_0 + \frac{\alpha_1}{\beta_1} = \frac{\beta_1\beta_0 + \alpha_1}{\beta_1} = \frac{A_1}{B_1}. \tag{8.3.4}$$

The following lemma gives a solution to (8.3.3).

Lemma 8.3.2 (Wallis' method)**.** *Suppose that $\{A_n\}_{n\geq 0}$ and $\{B_n\}_{n\geq 0}$ are given by $A_0 = \beta_0$, $B_0 = 1$, and for $n \geq 1$,*

$$\begin{aligned} A_n &= \beta_n A_{n-1} + \alpha_n A_{n-2}, \\ B_n &= \beta_n B_{n-1} + \alpha_n B_{n-2}, \end{aligned}$$

where $A_{-1} = 1$ and $B_{-1} = 0$. Then

$$\frac{A_n}{B_n} = \beta_0 + \cfrac{\alpha_1}{\beta_1 + \cfrac{\alpha_2}{\beta_2 + \ddots + \cfrac{\alpha_n}{\beta_n}}}. \tag{8.3.5}$$

Proof. We use induction on n. By (8.3.4), the equation (8.3.5) is true for $n = 0, 1$. Suppose that the lemma is true for $n \geq 1$ and consider the case $n+1$. Observe that the $(n+1)$st convergent can be written as

$$C_{n+1} = \beta_0 + \cfrac{\alpha_1}{\beta_1 + \cfrac{\alpha_2}{\beta_2 + \ddots + \cfrac{\alpha_n}{\beta_n + \cfrac{\alpha_{n+1}}{\beta_{n+1}}}}} = \beta_0 + \cfrac{\alpha_1}{\beta_1 + \cfrac{\alpha_2}{\beta_2 + \ddots + \cfrac{\alpha_n\beta_{n+1}}{\beta_n\beta_{n+1} + \alpha_{n+1}}}},$$

which is exactly the nth convergent with the following substitutions:

$$\alpha_n \mapsto \alpha_n\beta_{n+1}, \qquad \beta_n \mapsto \beta_n\beta_{n+1} + \alpha_{n+1}. \tag{8.3.6}$$

Therefore, we have

$$C_{n+1} = \frac{A_n^*}{B_n^*},$$

where A_n^* (resp. B_n^*) is the same as A_n (resp. B_n) with the substitutions in (8.3.6). Thus by the induction hypothesis,

$$\begin{aligned} A_n^* &= (\beta_n\beta_{n+1} + \alpha_{n+1})A_{n-1} + (\alpha_n\beta_{n+1})A_{n-2}, \\ &= \beta_{n+1}(\beta_n A_{n-1} + \alpha_n A_{n-2}) + \alpha_{n+1}A_{n-1}, \\ &= \beta_{n+1}A_n + \alpha_{n+1}A_{n-1} = A_{n+1}, \end{aligned}$$

and the same computation also shows that $B_n^* = B_{n+1}$. This shows that (8.3.5) also holds for $n+1$, completing the proof. □

Now we can prove Theorem 8.3.1.

Proof of Theorem 8.3.1. By Lemma 8.3.2 with $\alpha_n = -\lambda_n x^2$ and $\beta_n = 1 - b_n x$, we have

$$\frac{A_k}{B_k} = 1 - b_0 x - \cfrac{\lambda_1 x^2}{1 - b_1 x - \cfrac{\lambda_2 x^2}{1 - b_2 x - \ddots - \cfrac{\lambda_k x^2}{1 - b_k x}}},$$

where A_n and B_n are given by $A_{-1} = 1$, $A_0 = 1 - b_0 x$, $B_{-1} = 0$, $B_0 = 1$, and for $n \geq 1$,

$$\begin{aligned} A_n &= (1 - b_n x)A_{n-1} + \lambda_n x^2 A_{n-2}, \\ B_n &= (1 - b_n x)B_{n-1} + \lambda_n x^2 B_{n-2}. \end{aligned}$$

By (8.3.2), we obtain $A_n = P_{n+1}^*(x; \boldsymbol{b}, \boldsymbol{\lambda})$. Replacing n by $n-1$ and taking δ in (8.3.2) we obtain

$$\delta P_n^*(x; \boldsymbol{b}, \boldsymbol{\lambda}) = (1 - b_n x)\delta P_{n-1}^*(x; \boldsymbol{b}, \boldsymbol{\lambda}) - \lambda_n x^2 \delta P_{n-2}^*(x; \boldsymbol{b}, \boldsymbol{\lambda}).$$

This shows that $B_n = \delta P_n^*(x; \boldsymbol{b}, \boldsymbol{\lambda})$. Combining these results with Theorem 8.2.3, we obtain

$$\begin{aligned} \sum_{n \geq 0} \mu_n^{\leq k}(\boldsymbol{b}, \boldsymbol{\lambda}) x^n &= \cfrac{1}{1 - b_0 x - \cfrac{\lambda_1 x^2}{1 - b_1 x - \cfrac{\lambda_2 x^2}{1 - b_2 x - \ddots - \cfrac{\lambda_k x^2}{1 - b_k x}}}} \\ &= \left(\frac{A_k}{B_k}\right)^{-1} = \frac{\delta P_k^*(x; \boldsymbol{b}, \boldsymbol{\lambda})}{P_{k+1}^*(x; \boldsymbol{b}, \boldsymbol{\lambda})}. \end{aligned}$$

□

We end this section with a connection between Theorem 8.2.3 and Padé approximants.

Let $P(x)$ and $Q(x)$ be polynomials of degree p and q, respectively. We say that $\frac{P(x)}{Q(x)}$ is a **Padé approximant** of type (p, q) for $F(x)$ if

$$F(x) - \frac{P(x)}{Q(x)} = \sum_{n \geq p+q+1} a_n x^n.$$

Often times Padé approximants are better than Taylor polynomials since rational functions can handle poles. It is well known that if a power series has a Padé approximant of type (p, q), then it is unique.

Corollary 8.3.3. *Let*

$$F(x) = \sum_{n \geq 0} \mu_n(\boldsymbol{b}, \boldsymbol{\lambda}) x^n.$$

Then for a nonnegative integer k,

$$\frac{\delta P_k^*(x; \boldsymbol{b}, \boldsymbol{\lambda})}{P_{k+1}^*(x; \boldsymbol{b}, \boldsymbol{\lambda})}$$

is the Padé approximant of type $(k, k+1)$ for $F(x)$.

Proof. We have

$$\sum_{n\geq 0} \mu_n^{\leq k}(\boldsymbol{b},\boldsymbol{\lambda})x^n = \frac{\delta P_k^*(x;\boldsymbol{b},\boldsymbol{\lambda})}{P_{k+1}^*(x;\boldsymbol{b},\boldsymbol{\lambda})}.$$

Observe that if

$$\mu_n(\boldsymbol{b},\boldsymbol{\lambda}) \neq \mu_n^{\leq k}(\boldsymbol{b},\boldsymbol{\lambda}),$$

we must have $n \geq 2k+2$. Hence the lowest term of

$$F(x) - \frac{\delta P_k^*(x;\boldsymbol{b},\boldsymbol{\lambda})}{P_{k+1}^*(x;\boldsymbol{b},\boldsymbol{\lambda})} = \sum_{n\geq 0} \left(\mu_n(\boldsymbol{b},\boldsymbol{\lambda}) - \mu_n^{\leq k}(\boldsymbol{b},\boldsymbol{\lambda})\right) x^n$$

has degree at most $2k+2$. This implies the statement. □

8.4 Motzkin paths with fixed starting and ending heights

In this section we generalize Theorem 8.3.1 to Motzkin paths with bounded height and fixed starting and ending heights. We begin with some definitions.

Let $\mathrm{Motz}_{n,r,s}^{\leq k}$ denote the set of Motzkin paths in $\mathrm{Motz}_{n,r,s} = \mathrm{Motz}((0,r) \to (n,s))$ that stays weakly below the line $y = k$. We also define

$$\mu_{n,r,s}^{\leq k}(\boldsymbol{b},\boldsymbol{\lambda}) = \sum_{\pi\in \mathrm{Motz}_{n,r,s}^{\leq k}} \mathrm{wt}(\pi;\boldsymbol{b},\boldsymbol{\lambda}).$$

Theorem 8.4.1. *Let $r, s,$ and k be integers with $0 \leq r, s \leq k$. If $r \leq s$, then*

$$\sum_{n\geq 0} \mu_{n,r,s}^{\leq k}(\boldsymbol{b},\boldsymbol{\lambda})x^n = x^{s-r}\frac{P_r^*(x;\boldsymbol{b},\boldsymbol{\lambda})\delta^{s+1}P_{k-s}^*(x;\boldsymbol{b},\boldsymbol{\lambda})}{P_{k+1}^*(x;\boldsymbol{b},\boldsymbol{\lambda})}.$$

If $r > s$, then

$$\sum_{n\geq 0} \mu_{n,r,s}^{\leq k}(\boldsymbol{b},\boldsymbol{\lambda})x^n = \lambda_{s+1}\cdots\lambda_r\frac{P_s^*(x;\boldsymbol{b},\boldsymbol{\lambda})\delta^{r+1}P_{k-r}^*(x;\boldsymbol{b},\boldsymbol{\lambda})}{P_{k+1}^*(x;\boldsymbol{b},\boldsymbol{\lambda})}.$$

It is possible to prove Theorem 8.4.1 using sign-reversing involutions as before. However, we will not pursue in this direction. Instead we will give two alternative proofs: one using linear algebra techniques and the other using combinatorial arguments. We first need to introduce some definitions.

Let $G = (V, E)$ be a directed graph on the vertex set $V = [m]$. Let $w : E \to K$ be an edge-weight and let $A = (a_{i,j})_{i,j=1}^m$ be the adjacency matrix of this weighted graph G, that is, $a_{i,j} = w(i \to j)$. Here, if there is no edge $i \to j$ in G, we define $a_{i,j} = 0$.

A **path** of length n from u to v is a sequence $p = (u_0, \dots, u_n)$ of vertices such that $u_0 = u$, $u_n = v$, and $u_{i-1} \to u_i$ is an edge for all $i = 1, \dots, n$. The **weight** $w(p)$ of p is defined to be the product of the weight of each step:

$$w(p) = w(u_0 \to u_1)w(u_1 \to u_2)\cdots w(u_{n-1} \to u_n) = a_{u_0,u_1}a_{u_1,u_2}\cdots a_{u_{n-1},u_n}.$$

Let $P_n(u \to v)$ denote the set of all paths of length n from u to v. We also define $P(u \to v)$ to be the set of all paths from u to v.

Lemma 8.4.2. *For vertices $i, j \in [m]$ and an integer $n \geq 0$,*

$$\sum_{p\in P_n(i\to j)} w(p) = (A^n)_{i,j}.$$

Proof. This is immediate from the definition of the product of matrices. □

Proposition 8.4.3. *Suppose that $I-A$ is invertible. We have*

$$\sum_{p\in P(i\to j)} w(p) = \left((I-A)^{-1}\right)_{i,j} = \frac{(-1)^{r+s}[I-A]_{\{j\}^c,\{i\}^c}}{\det(I-A)},$$

where I is the $m\times m$ identity matrix.

Proof. By Lemma 8.4.2, the left-hand side is equal to

$$\sum_{n\ge 0}\left(\sum_{p\in P_n(i\to j)} w(p)\right) = \sum_{n\ge 0}(A^n)_{i,j} = \left(\sum_{n\ge 0} A^n\right)_{i,j} = \left((I-A)^{-1}\right)_{i,j},$$

which proves the first equality. The second equality follows from the minor inverse formula Lemma 7.4.4. □

Now let us return to the following generating function that we need to compute:

$$\sum_{n\ge 0}\mu^{\le k}_{n,r,s}(\boldsymbol{b},\boldsymbol{\lambda})x^n = \sum_{n\ge 0}\sum_{\pi\in \mathrm{Motz}^{\le k}_{n,r,s}} \mathrm{wt}(\pi;\boldsymbol{b},\boldsymbol{\lambda})x^n. \tag{8.4.1}$$

Define a graph G_k on $V=\{0,1,\dots,k\}$ with edges (i,j) for all possible $i,j\in V$ with $|i-j|\le 1$. Define an edge-weight w on G_k by $w(i,i+1)=x$, $w(i,i)=b_ix$, and $w(i,i-1)=\lambda_i x$. Then the adjacency matrix of G_k is the following tridiagonal matrix:

$$A = \begin{pmatrix} b_0x & x & & 0\\ \lambda_1 x & b_1 x & \ddots & \\ & \ddots & \ddots & x\\ 0 & & \lambda_k x & b_k x\end{pmatrix}.$$

Observe that every $\pi\in\mathrm{Motz}^{\le k}_{n,r,s}$ can be considered as a path p of length n from r to s for the graph G_k. In this case we have $\mathrm{wt}(\pi;\boldsymbol{b},\boldsymbol{\lambda})x^n = w(p)$. Therefore we can rewrite (8.4.1) as

$$\sum_{n\ge 0}\mu^{\le k}_{n,r,s}(\boldsymbol{b},\boldsymbol{\lambda})x^n = \sum_{n\ge 0}\sum_{p\in P_n(r\to s)} w(p) = \sum_{p\in P(r\to s)} w(p). \tag{8.4.2}$$

Then, by Proposition 8.4.3,

$$\sum_{n\ge 0}\mu^{\le k}_{n,r,s}(\boldsymbol{b},\boldsymbol{\lambda})x^n = \frac{(-1)^{r+s}[I-A]_{\{s\}^c,\{r\}^c}}{\det(I-A)}. \tag{8.4.3}$$

Now we need to compute the numerator and the denominator of the right-hand side of (8.4.3).

Lemma 8.4.4. *We have*

$$P^*_{k+1}(x;\boldsymbol{b},\boldsymbol{\lambda}) = \det(I-A).$$

Proof. Let

$$Q_{k+1}(x) = \det(I-A) = \det\begin{pmatrix} 1-xb_0 & -x & & 0\\ -x\lambda_1 & 1-xb_1 & \ddots & \\ & \ddots & \ddots & -x\\ 0 & & -x\lambda_k & 1-xb_k\end{pmatrix}.$$

Expanding the determinant along the last row, we have

$$Q_{k+1}(x) = (1-b_kx)Q_k(x) - \lambda_k x^2 Q_{k-1}(x).$$

This is the same recurrence for $P^*_n(x;\boldsymbol{b},\boldsymbol{\lambda})$ in (8.3.2). Since $Q_0(x)=1=P^*_0(x)$ and $Q_1(x)=1-xb_0=P^*_1(x)$, we obtain the lemma. □

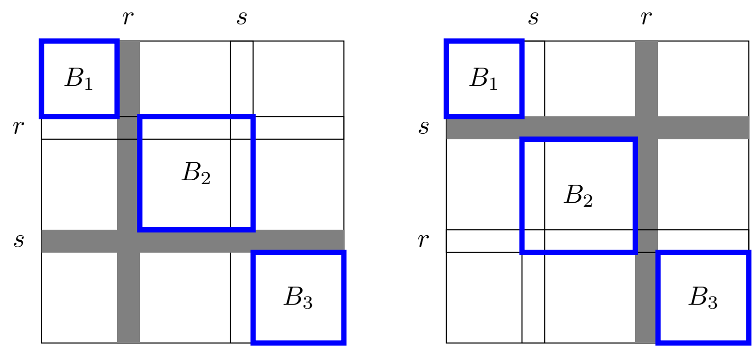

Figure 8.1: The minor $(-1)^{r+s}[I-A]_{\{s\}^c,\{r\}^c}$ is the product of the determinants of three block matrices B_1, B_2, B_3. The left diagram is for the case $r \le s$ and the right diagram is for the case $r > s$.

Lemma 8.4.5. *We have*

$$(-1)^{r+s}[I-A]_{\{s\}^c,\{r\}^c} = \begin{cases} x^{s-r} P_r^*(x;\boldsymbol{b},\boldsymbol{\lambda})\delta^{s+1}P_{k-s}^*(x;\boldsymbol{b},\boldsymbol{\lambda}) & \text{if } r \le s, \\ \lambda_{s+1}\cdots\lambda_r P_s^*(x;\boldsymbol{b},\boldsymbol{\lambda})\delta^{r+1}P_{k-r}^*(x;\boldsymbol{b},\boldsymbol{\lambda}) & \text{if } r > s. \end{cases}$$

Proof. Since $I-A$ is tridiagonal, we can decompose $(-1)^{r+s}[I-A]_{\{s\}^c,\{r\}^c} = \det B_1 \det B_2 \det B_3$ as shown in Figure 8.1. Suppose that $r \le s$. Then, by 8.4.4, we have $\det B_1 = P_r^*(x;\boldsymbol{b},\boldsymbol{\lambda})$ and $\det B_3 = \delta^{s+1}P_{k-s}^*(x;\boldsymbol{b},\boldsymbol{\lambda})$. Since B_2 is lower trianular with diagonal entries $x,\dots,x$, we have $\det B_2 = x^{s-r}$.

Similarly, if $r > s$, we have $\det B_1 = P_s^*(x;\boldsymbol{b},\boldsymbol{\lambda})$ and $\det B_3 = \delta^{r+1}P_{k-r}^*(x;\boldsymbol{b},\boldsymbol{\lambda})$. Since B_2 is upper trianular with diagonal entries $\lambda_{s+1},\dots,\lambda_r$, we have $\det B_2 = \lambda_{s+1},\dots,\lambda_r$. This proves the lemma. □

By (8.4.3), Lemma 8.4.4 and Lemma 8.4.5, we obtain Theorem 8.4.1.

8.5 A combinatorial proof using disjoint paths and cycles

In this section we give another proof of Theorem 8.4.1 using self-avoiding paths. As in the previous section let $G = (V, E)$ be a directed graph on the vertex set $V = [m]$ with edge-weight $w : E \to K$ and adjacency matrix $A = (a_{i,j})_{i,j=1}^m$.

A **cycle** is a path $p = (u_0,\dots,u_n)$ such that $u_n = u_0$. We identify a cycle $p = (u_0,\dots,u_n)$ with any of its cyclic shift $p = (u_j,\dots,u_n,u_0,\dots,u_j)$. We say that a path $p = (u_0,\dots,u_n)$ is **self-avoiding** if $u_i \ne u_j$ for all $i \ne j$ with a possible exception $u_0 = u_n$. A collection $\{p_1, p_2, \dots, p_t\}$ of paths is said to be **disjoint** if the following conditions hold:

- p_i is self-avoiding for all $1 \le i \le t$,
- p_i and p_j have no common vertices for all $1 \le i \ne j \le t$.

We define

$$\Gamma = \sum_{\{C_1,C_2,\dots,C_t\}\in\mathcal{C}} (-1)^t w(C_1)\cdots w(C_t),$$

where $\mathcal{C}$ is the set of all disjoint collections $\{C_1, C_2, \dots, C_t\}$ of cycles in G. For $1 \le i, j \le m$, we define

$$\Gamma_{i,j} = \sum_{(p,\{C_1,C_2,\dots,C_t\})\in\mathcal{C}_{i,j}} w(p)\cdot(-1)^t w(C_1)\cdots w(C_t), \tag{8.5.1}$$

where $\mathcal{C}_{i,j}$ is the set of all pairs $(p, \{C_1, C_2, \dots, C_t\})$ of a path $p \in P(i \to j)$ and a collection $\{C_1, C_2, \dots, C_t\}$ of cycles in G such that $\{p, C_1, C_2, \dots, C_t\}$ is disjoint.

Note that since

Proposition 8.5.1. *We have*

$$\sum_{p\in P(i\to j)} w(p) = \frac{\Gamma_{i,j}}{\Gamma}.$$

Proof. We will prove the equivalent statement

$$\Gamma \sum_{p\in P(i\to j)} w(p) = \Gamma_{i,j}.$$

The left-hand side can be written as

$$\sum_{(p,\{C_1,C_2,\dots,C_t\})\in X} w(p)\cdot(-1)^t w(C_1)\cdots w(C_t),$$

where X is the set of pairs $(p,\{C_1,C_2,\dots,C_t\})$ of any path $p\in P(i\to j)$ and any disjoint collection $\{C_1,C_2,\dots,C_t\}$ of cycles in G. By (8.5.1), it suffices to find a sign-reversing weight-preserving involution $\phi: X\to X$ with fixed point set $\mathcal{C}_{i,j}$.

Consider $(p,\{C_1,C_2,\dots,C_t\})\in X$. If $(p,\{C_1,C_2,\dots,C_t\})\in\mathcal{C}_{i,j}$, then define $\phi(p,\{C_1,C_2,\dots,C_t\}) = (p,\{C_1,C_2,\dots,C_t\})$. Now suppose that $(p,\{C_1,C_2,\dots,C_t\})\notin\mathcal{C}_{i,j}$ and let $p=(u_0,u_1,\dots,u_n)$. Then we can find the smallest integer s such that $u_s=u_r$ for some $0\le r<s$ or u_s is contained in C_r for some $1\le r\le t$.

Case 1: $u_s=u_r$ for some $0\le r<s$. Then we define

$$\phi(p,\{C_1,C_2,\dots,C_t\}) = (p',\{C_0,C_1,C_2,\dots,C_t\}),$$

where $p'=(u_0,\dots,u_r,u_{s+1},\dots,u_n)$ and $C_0=(u_r,\dots,u_s)$.

Case 2: u_s is contained in a cycle in $\{C_1,C_2,\dots,C_t\}$. Since the ordering of the cycles does not mattter, we may assume that $u_s\in C_1$. Let $C_1=(v_0,v_1,\dots,v_q)$, where $v_0=v_q=u_s$. Then we define

$$\phi(p,\{C_1,C_2,\dots,C_t\}) = (p',\{C_2,\dots,C_t\}),$$

where $p'=(u_0,\dots,u_s,v_1,\dots,v_q,u_{s+1},\dots,u_n)$.

Note that these two cases cannot happen at the same time because then $0\le r<s$ would have been chosen instead of s.

Since ϕ increases or decreases the number of cycles by 1 and does not change the set of edges, it is a desired sign-reversing weight-preserving involution. This completes the proof. $\square$

Now we are ready to give another proof of Theorem 8.4.1. We restate the theorem for the reader's convenience.

Theorem 8.5.2 (Theorem 8.4.1)**.** *Let $r,s,$ and k be integers with $0\le r,s\le k$. If $r\le s$, then*

$$\sum_{n\ge0}\mu^{\le k}_{n,r,s}(\boldsymbol{b},\boldsymbol{\lambda})x^n = x^{s-r}\frac{P^*_r(x;\boldsymbol{b},\boldsymbol{\lambda})\delta^{s+1}P^*_{k-s}(x;\boldsymbol{b},\boldsymbol{\lambda})}{P^*_{k+1}(x;\boldsymbol{b},\boldsymbol{\lambda})}.$$

If $r>s$, then

$$\sum_{n\ge0}\mu^{\le k}_{n,r,s}(\boldsymbol{b},\boldsymbol{\lambda})x^n = \lambda_{s+1}\cdots\lambda_r\frac{P^*_s(x;\boldsymbol{b},\boldsymbol{\lambda})\delta^{r+1}P^*_{k-r}(x;\boldsymbol{b},\boldsymbol{\lambda})}{P^*_{k+1}(x;\boldsymbol{b},\boldsymbol{\lambda})}.$$

Proof. Recall the weighted graph G_k defined in the previous section whose adjacency matrix is

$$A=\begin{pmatrix} b_0x & x & & 0\\ \lambda_1 x & b_1x & \ddots & \\ & \ddots & \ddots & x\\ 0 & & \lambda_k x & b_kx\end{pmatrix}.$$

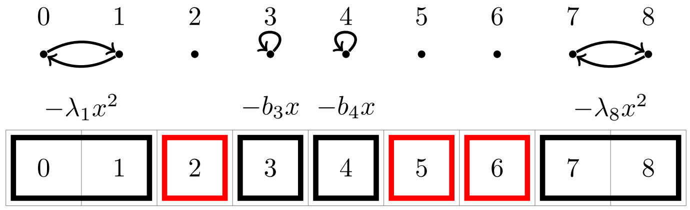

Figure 8.2: A disjoint collection of cycles in G_k (up) and the corresponding Favard tiling $T \in \mathrm{FT}_{k+1}$ with $\mathrm{wt}^*(T) = \lambda_1\lambda_8 b_3 b_4 x^6$ (bottom) for $k = 8$.

By (8.4.2) and Proposition 8.5.1,

$$\sum_{n\geq 0} \mu_{n,r,s}^{\leq k}(\boldsymbol{b},\boldsymbol{\lambda})x^n = \sum_{p\in P(r\to s)} w(p) = \frac{\Gamma_{r,s}}{\Gamma}, \tag{8.5.2}$$

where

$$\Gamma = \sum_{\{C_1,C_2,\dots,C_t\}\in\mathcal{C}} (-1)^t w(C_1)\cdots w(C_t),$$
$$\Gamma_{r,s} = \sum_{(p,\{C_1,C_2,\dots,C_t\})\in\mathcal{C}_{r,s}} w(p)\cdot(-1)^t w(C_1)\cdots w(C_t).$$

Since A is tridiagonal, every cycle with nonzero weight must be a cycle of the form (i) or $(i-1,i)$ whose weight is $-b_i x$ or $-\lambda_i x^2$. Then a collection $\{C_1, C_2, \dots, C_t\}$ of disjoint cycles is equivalent to a Favard tiling $T \in \mathrm{FT}_{k+1}$, where a cycle (i) corresponds to a black monomino labeled i and a cycle $(i-1,i)$ corresponds to a black domino labeled $i-1, i$. A red monomino labeled i in T means that there is no cycle containing i in $\{C_1, C_2, \dots, C_t\}$. In this correspondence we have

$$(-1)^t w(C_1)\cdots w(C_t) = \mathrm{wt}^*(T),$$

where $\mathrm{wt}^*(T)$ is the product of $-b_i x$ for each black monomino labeled i and $-\lambda_i x^2$ for each black domino labeled $i-1, i$. For example, see Figure 8.2. This implies that

$$\Gamma = \sum_{T\in\mathrm{FT}_{k+1}\,\mathrm{wt}^*(T)} = P_{k+1}^*(x), \tag{8.5.3}$$

where the second identity follows from the recurrence (8.3.2).

On the other hand, if $(p, \{C_1, C_2, \dots, C_t\}) \in \mathcal{C}_{r,s}$, then p must be the unique self-avoiding path from r to s, namely $p = (r, r+1, \dots, s)$. If $r \leq s$, then $w(p) = x^{s-r}$, and if $r > s$, then $w(p) = \lambda_{s+1}\cdots\lambda_r$. Moreover, we can divide $\{C_1, C_2, \dots, C_t\}$ into two sets $\{C_1', \dots, C_{t_1}'\}$ and $\{C_1'', \dots, C_{t_2}''\}$ such that C_i' is a cycle with elements less than $\min\{r, s\}$ and C_i'' is a cycle with elements great than $\max\{r, s\}$. This shows that

$$\Gamma_{r,s} = \begin{cases} P_r^*(x;\boldsymbol{b},\boldsymbol{\lambda})\delta^{s+1}P_{k-s}^*(x;\boldsymbol{b},\boldsymbol{\lambda}) & \text{if } r\leq s,\\ P_s^*(x;\boldsymbol{b},\boldsymbol{\lambda})\delta^{r+1}P_{k-r}^*(x;\boldsymbol{b},\boldsymbol{\lambda}) & \text{otherwise.}\end{cases} \tag{8.5.4}$$

By (8.5.2), (8.5.3), and (8.5.4), we obtain the theorem. □

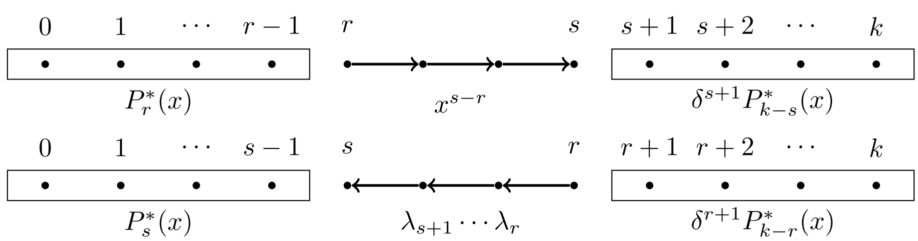

Figure 8.3: A pair $(p, \{C_1, C_2, \dots, C_t\}) \in \mathcal{C}_{r,s}$ is illustrated at the top for the case $r \leq s$ and at the bottom for the case $r > s$.

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)

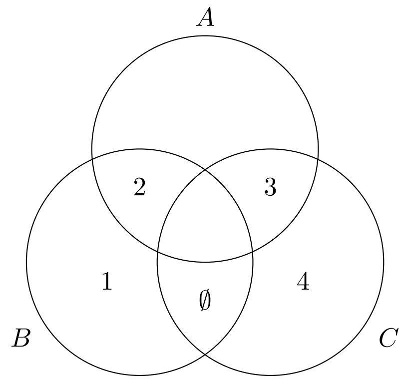

Figure A.1: The triple (A, B, C).

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 $\mathrm{sgn}(A, B, C) = (-1)^{|B\setminus A|+|C\setminus A|}$. 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] M. Aigner and G. M. Ziegler. *Proofs from The Book*. Springer, Berlin, sixth edition, 2018. See corrected reprint of the 1998 original [MR1723092], Including illustrations by Karl H. Hofmann.

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

[3] S. Corteel and A. Nunge. Combinatorics of the 2-species exclusion processes, marked Laguerre histories, and partially signed permutations. *Electron. J. Combin.*, 27(2):Paper No. 2.53, 27, 2020.

[4] D. Foata and D. Zeilberger. Denert's permutation statistic is indeed Euler-Mahonian. *Stud. Appl. Math.*, 83(1):31–59, 1990.

[5] J. Françon and G. Viennot. Permutations selon leurs pics, creux, doubles montées et double descentes, nombres d'Euler et nombres de Genocchi. *Discrete Math.*, 28(1):21–35, 1979.

[6] I. Gessel and G. Viennot. Binomial determinants, paths, and hook length formulae. *Adv. in Math.*, 58(3):300–321, 1985.

[7] B. Lindström. On the vector representations of induced matroids. *Bull. London Math. Soc.*, 5:85–90, 1973.

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

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

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