SYMMETRIC FUNCTIONS AND INTEGRABLE PROBABILITY

NIKOS ZYGOURAS

Abstract. We will present the basics of symmetric functions theory, starting from the representation theory of the symmetric group, then moving to related combinatorial structures eg Young tableaux and Robinson-Schensted-Knuth correspondences, their relations to Schur functions and characters of the symmetric group. Then we will move to probabilistic aspects and see how the algebraic and probabilistic structured interplay. We will also look at fundamental generalisations of Schur functions, called Macdonald functions. If time permits we will discuss how symmetric functions emerge as partition functions of vertex models and the relations to Yang-Baxter idead and R-matrices.

Contents

Date: April 19, 2024.

2020 *Mathematics Subject Classification.* Primary: 60Cxx, 05Exx, 82B23.

Key words and phrases. symmetric group, representation theory, KPZ universality, Robinson-Schensted-Knuth correspondence, Young tableaux, Gelfand-Tsetlin patterns, tropical combinatorics, growth models, directed polymer models, Whittaker functions, Schur functions, Pieri rule, Branching rule.

1. Introduction

For the moment, look at the introductions of [BG16, BP14, BP16a, Z22]. What we roughly plan to cover in this course is:

A. Basics of representation theory

- Basics of permutations
- Definitions of representations
- Modules
- Reducibility and Mascke's theorem
- G-hommomorphisms
- Characters and inner products

B. Combinatorial structures.

- Young tableaux and its relations to the representations of S_n.
- Gelfand-Tsetlin patterns
- Robinson-Schensted-Knuth correspondence

C. First symmetric functions

Basic symmetric functions: complete, elementary and power symmetric functions

Schur functions: combinatorial and representation formulations

D. Introduction to integrable probability

Schur measure

Integrable last passage percolation

Determinatal processes and Fredholm determinants

Role of algebraic structures in integrable probability: Cauchy identity, Pieri rule, Branching rule

Related Markovian dynamics

E. Macdonald functions and various specialisations

F. (?) Vertex models and Yang-Baxter relations

2. Review of representation theory

2.1. Introductory setting of permutations. Here we review some basics of representation theory with emphasis on the representation theory of the symmetric group. The main reference is the book [S13].

The group S_n of permutations of n elements $\{1, 2, ..., n\}$ is a mapping $\pi\colon \{1, 2, ..., n\} \to \{1, 2, ..., n\}$. We can represent a permutation π is a few ways. The first one is the two-row array

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(n) \end{pmatrix},$$

for example the permuation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 1 & 10 & 2 & 4 & 7 & 5 & 6 & 9 & 3 & 8 \end{pmatrix},$$

sends $1 \to 1$, $2 \to 10$, $3 \to 2$ etc. We also have the cycle representation of a permutation, say

$$\pi = (i_1, ..., i_{k_1})(i_{k_1+1}, ..., i_{k_2}) \cdots (i_{k_{\ell-1}} \cdots i_{k_\ell}),$$

which means that

$$i_2 = \pi(i_1), i_3 = \pi(i_2), \ldots, i_1 = \pi(i_{k_1}),$$

or that under π we have the sequence of mappings $i_1 \to i_2 \to \cdots \to i_k \to i_1$ and so on for the rest of the cycles.

Remark 2.1. Understanding the cycle structure of a (random) permutation is a very interesting problem with a large number of very interesting results, which have relations to probability, algebraic structures, number theory, quantum mechanics and beyond.

An important related notion, that we will meet often, is that of a **partition**. A partition, denoted by $\lambda = (\lambda_1, \lambda_2, ...)$ of an integer n is a non-increasing sequence of integers $\lambda_1 \geq \lambda_2 \geq \cdots$ such that $\lambda_1 + \lambda_2 + \cdots = n$. If λ is a partition of n, we write $\lambda \vdash n$. Numbers $\lambda_1, \lambda_2, ...$ are called the parts of the partition.

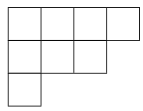
Another related notion is that of the **type** of a partition of a permutation or of a partition. In particular, for a permutation π we write

$$type(\pi) := (1^{m_1} 2^{m_2} \cdots n^{m_n}),$$

to mean that π has m_i cycles of length i, for $i = 1, ..., n$. Similarly, for a partition λ, $type(\lambda)$ denotes that λ has m_i parts equal to i.

Related to partitions is the notions of **Young diagrams**, which will also play an important role in our study. For a partition λ, a Young diagram is a left-aligned array of boxes, the first row of which has λ_1 boxes, the second row λ_2 boxes etc. For example:

$$\lambda = (4, 3, 1) \qquad \longleftrightarrow$$



Let us now expose some group theoretic terms.

Definition 2.2. *Let G be a group. We say that elements $g, h \in G$ are* **conjugates** *if $g = khk^{-1}$ for some element $k \in G$. For $g \in G$, we define the* **conjugacy class** *of g to be $K_g := \{h \in G \colon h = kgk^{-1}$ for some $k \in G\}$.*

Remark 2.3. Understanding the number of conjugacy classes is important as it will turn out that this is also the number of irreducible representations.

Definition 2.4. *The* **centraliser** *of an element $g \in G$ is the set*

$$Z_g : \{h \in G \colon g = hgh^{-1}\},$$

or, in other words, it is the set of elements of G that commute with g.

The following proposition determines the size of the centraliser of an element $\pi \in S_n$. This number will actually appear in all symmetric functions that we will expose, though via a different route...

Proposition 2.5. *Let $\lambda = (1^{m_1} 2^{m_2} \cdots)$ be a partition with the given type and let $\pi \in S_n$ with type λ. The size of the centraliser of π is*

$$z_\lambda := |Z_\pi| = 1^{m_1} m_1! 2^{m_2} m_2! \cdots n^{m_n} m_n! \tag{2.1}$$

Proof. We first need to determine a feature of the cycle structure of two permutations π, σ, which are conjugate to each other, ie $\pi = \sigma\pi\sigma^{-1}$. Let π have the cycle structure $\pi = (i_1 \cdots i_k) \cdots (i_m \cdots i_n)$. Then for every $\sigma \in S_n$ it holds that

$$\sigma\pi\sigma^{-1} = \big(\sigma(i_1) \cdots \sigma(i_k)\big) \cdots \big(\sigma(i_m) \cdots \sigma(i_n)\big).$$

This is because

$$\sigma\pi\sigma^{-1}\big(\sigma(i_1)\big) = \sigma\pi(i_1) = \sigma(i_2),$$

since $\pi(i_1) = i_2$, and similarly for the rest of the elements in the cycle representation. In other words, conjugation preserves the cycle structure.

Let us, now, suppose that $\pi = \sigma\pi\sigma^{-1}$. In cycle representation this translates as

$$(i_1 \cdots i_k) \cdots (i_m \cdots i_n) = \big(\sigma(i_1) \cdots \sigma(i_k)\big) \cdots \big(\sigma(i_m) \cdots \sigma(i_n)\big),$$

and this can hold if and only if σ results to a permutation of the cycle of same length (since there are m_i cycles of length i the number of such is $m_i!$) but also makes a cyclic permutation of elements within a cycle. In a cycle of length i this can happen if any element of the cycle mapped k positions clockwise with possible values for $k = 1, ..., i$. Since there are m_i such cycles the total number of such mappings in i^{m_i}. Combining the two possible scenaria we obtain that the total number is $1^{m_1} m_1! \cdots n^{m_n} m_n!$. $\square$

It is known that there is a bijection between conjugacy classes of an element $g \in G$ and left cosets of the centraliser of g, which is denoted by G/Z_g. Thus,

$$|K_g| = |G/Z_g| = \frac{|G|}{|Z_g|} \tag{2.2}$$

and in the case of S_n this translates to

$$|K_\pi| = \frac{n!}{1^{m_1} m_1! 2^{m_2} m_2! \cdots n^{m_n} m_n!}.$$

if permutation π has type $(1^{m_1} \cdots n^{m_n})$.

2.2. Matrix representations.

Definition 2.6. *Let G be a group. A matrix representation of G is a group homomorphism*

$$X \colon G \to \mathrm{GL}_d,$$

i.e. from G to the set of invertible $d \times d$ matrices (in which case, we say that the degree of the representation is d), if

- $X(\varepsilon) = \mathrm{I}_d$, *i.e. the unit element of ε of G is mapped to the $d \times d$ identity matrix I_d,*
- $X(gh) = X(g)X(h)$, *for any $g, h \in G$.*

Let us present some examples.

Example 1. *The first example is the* **trivial representation***, which maps*

$$G \ni g \xrightarrow{\mathbb{1}_G} 1 \in \mathbb{R}.$$

Example 2. Degree 1 representations of the **cyclic group**. Let C_n by the cyclic group $C_n := \{g, g^2, ..., g^n = \varepsilon\}$. Assume that $X(g) = c \in \mathbb{C}$. Then by the homomorphism property, we will have that

$$c^n = X(g^n) = X(\varepsilon) = 1,$$

which forces c to be a root of unity. This leads to the fact that there are n representations of degree 1 of the cyclic group.

We can, readily, also produce degree 2 representations in the form of diagonal matrices $X = diag(c_1, c_2)$ with c_1, c_2 two roots of unity. This can also be written in a *direct product* form $X = X^{(1)} \oplus X^{(2)}$ where $X^{(1)}, X^{(2)}$ are degree 1 representations, which will be called later on *irreducible components*.

It also turns out that every representation of C_n can be constructed using representations of degree 1 in a similar fashion.

Example 3. (Sign representations of S_n). A *transposition* is a cycle of length two which permutes two neighbouring elemensts, e.g. $(i, i+1)$. Every permutation π can be written as a composition of traspositions, e.g. $\pi = \tau_1\tau_2 \cdots \tau_k$. The sign of a permutation $\pi = \tau_1\tau_2 \cdots \tau_k$ is defined as $sgn(\pi) := (-1)^k$. The sign of a permutation plays an important role in determinantal considerations that we will see later on.

The mapping $X(\pi) := sgn(\pi)$, defines a degree 1 representation of S_n.

Exercise 1. *Show that the sign of a permutation defines a representation of S_n.*

Example 4. *The* **defining representation** *of S_n is a degree n representation, which is given by the matrices $X(\pi) = (x_{i,j})_{1\le i,j\le n}$ with*

$$x_{i,j} = \begin{cases} 1, & \text{if } \pi(j) = i, \\ 0, & \text{otherwise.} \end{cases}$$

For example,

$$X(\varepsilon) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad X\big((1,2)\big) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad X\big((3,2,1)\big) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

and so on.

The matrices involved in the above matrix representation will play an important role in probabilstic models later on.

Exercise 2. *Show that the defining representation of S_n is a representation.*

In a sense, representation theory tries to understand the invariant actions of a group. Towards this, it is helpful to recast the notion of representations in the framework of **G-modules**, i.e. the action of a group G on vector spaces.

Definition 2.7. (G-modules). *Let V be a vector space and G a group. We say that V is a G-module if there is a group homomorphism*

$$\varrho\colon G \to \mathrm{GL}(V),$$

where $\mathrm{GL}(V)$ is the space of invertible, linear forms on V (eg you can think of matrices). More precisely, there is a notion of multiplication $g\boldsymbol{v}$ of elements of V by elements of G which we can actually better think of as $g\boldsymbol{v} \equiv \varrho(g)\boldsymbol{v}$ and which, additionally, has the properties that for any $d, c \in \mathbb{C}, g, h \in G$ and $\boldsymbol{v} \in V$:

1. $g\boldsymbol{v} \in V$,
2. $g(c\boldsymbol{v} + d\boldsymbol{w}) = c(g\boldsymbol{v}) + d(g\boldsymbol{w})$,
3. $(gh)\boldsymbol{v} = g(h\boldsymbol{v})$,
4. $\varepsilon\boldsymbol{v} = \boldsymbol{v}$.

Exercise 3. *Show that the notions of G-modules and matrix representations are equivalent. More precisely, if X is a matrix representation and $\boldsymbol{v} \in V$, define $g\boldsymbol{v} := X(g)\boldsymbol{v}$ and check that the defining properties of the G-module coincide with the defining properties of the matrix representation $X(g)$. In the other direction, i.e. given a G-module on a vector space V, you can pick a basis $\mathcal{B}$ of V and let $X(g)$ be the matrix of linear transformation of g on V. Check the equivalence of the conditions.*

Let us now discuss **group actions**. Given a finite set $S = \{s_1, ..., s_n\}$ we can think of S as a vector space $\mathbb{C}S = span\{\boldsymbol{s}_1, ..., \boldsymbol{s}_n\}$, viewing elements $s_1, ..., s_n$ as linearly independent objects $\boldsymbol{s}_1, ..., \boldsymbol{s}_n$, thus, forming a vector space with the properties

- **(addition)** $\sum_i c_i\boldsymbol{s}_i + \sum_i d_i\boldsymbol{s}_i = \sum_i (c_i + d_i)\boldsymbol{s}_i$,
- **(scalar multiplication)** $c\sum_i c_i\boldsymbol{s}_i = \sum cc_i\boldsymbol{s}_i$,
- **(group action)** $g\sum_i c_i\boldsymbol{s}_i = \sum c_i(g\boldsymbol{s}_i)$.

The above turn S and hence $\mathbb{C}S$ to a G-module of dimension $|S|$.

In the case of permutations the group action on a set $S := \{1, 2, ..., n\}$ is as follows. We think of $1, 2, ..., n$ as linearly independent vectors $\boldsymbol{1}, \boldsymbol{2}, ..., \boldsymbol{n}$ and then $\mathbb{C}S := span\{\boldsymbol{1}, \boldsymbol{2}, ..., \boldsymbol{n}\}$ with $\{\boldsymbol{1}, \boldsymbol{2}, ..., \boldsymbol{n}\}$ being considered as the "standard" basis and with the group action

$$\pi(c_1\boldsymbol{1} + c_2\boldsymbol{2} + \cdots + c_n\boldsymbol{n}) = c_1\pi(\boldsymbol{1}) + c_2\pi(\boldsymbol{2}) = \cdots c_n\pi(\boldsymbol{n}),$$

with $\pi(\boldsymbol{i})$ being thought of as $\pi(i)$. In the basis $\{\mathbf{1}, \mathbf{2}, ..., \boldsymbol{n}\}$ we can determine the matrices $X(\pi)$ of the action of π. For example, let $\pi = (1,2)$, then

$$(1,2)\mathbf{1} = \mathbf{2}, \quad (1,2)\mathbf{2} = \mathbf{1}, \quad (1,2)\mathbf{3} = \mathbf{3},$$

and, thus, in this basis the matrix representation of $(1,2)$ is

$$X\big((1,2)\big) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and this agrees with the standard matrix representation.

Definition 2.8. (Left regular representation). *This is related to the group acting on itself. If G is a group with elements $g_1, ..., g_n$, then we consider the vector space (actually the algebra) $\mathbb{C}[G] = span\{\boldsymbol{g}_1, ..., \boldsymbol{g}_n\}$ generated by elements $g_1, ..., g_n$ viewed as linearly independent vectors $\boldsymbol{g}_1, ..., \boldsymbol{g}_n$. Then the action of G on the group algebra $\mathbb{C}[G]$ is expressed as*

$$g(c_1\boldsymbol{g}_1 + \cdots c_n\boldsymbol{g}_n) = c_1(\boldsymbol{g}\boldsymbol{g}_1) + \cdots + c_n(\boldsymbol{g}\boldsymbol{g}_n).$$

Example 5. (Regular representation of the cyclic group). In the case of C_4, we have t $\mathbb{C}[C_4] = span\{\boldsymbol{\varepsilon}, \boldsymbol{g}, \boldsymbol{g}^2, \boldsymbol{g}^3\}$. The matrix representation of g^2 in the (standard) basis $\boldsymbol{\varepsilon}, \boldsymbol{g}, \boldsymbol{g}^2, \boldsymbol{g}^3$ can be computed via

$$g^2\boldsymbol{\varepsilon} = \boldsymbol{g}^2, \quad g^2\boldsymbol{g} = \boldsymbol{g}^3, \quad g^2\boldsymbol{g}^2 = \boldsymbol{\varepsilon}, \quad g^2\boldsymbol{g}^3 = \boldsymbol{g},$$

thus

$$X(g^2) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Example 6. (Coset representation). Let H be a subgroup of G, which we will write as $H \leq G$. Let $g_1, ..., g_k$ be **transversal** for H, i.e. the set $\mathcal{H} := \{\boldsymbol{g}_1 H, ..., \boldsymbol{g}_k H\}$ is a complete set of *disjoint cosets* of H in G. We consider the module

$$\mathbb{C}\mathcal{H} := \{c_1\boldsymbol{g}_1\boldsymbol{H} + \cdots c_k\boldsymbol{g}_k\boldsymbol{H} : c_i \in \mathbb{C}\}$$

with the group action

$$g(c_1\boldsymbol{g}_1\boldsymbol{H} + \cdots c_k\boldsymbol{g}_k\boldsymbol{H}) = c_1(\boldsymbol{g}\boldsymbol{g}_1)\boldsymbol{H} + \cdots c_k(\beta\boldsymbol{g}_k)\boldsymbol{H}.$$

If $H = \{\varepsilon\}$, then all $g \in G$ are transversal and so $\mathcal{H} = G$ and the coset representation coincides with the left regular representation. If $H = G$, then $\mathcal{H} = \{\varepsilon\}$ and this gives rise to the trivial representation.

A less trivial coset representation is the following: Let $G = S_3$ and $H := \{\varepsilon, (2,3)\}$. Then $\mathcal{H} = \{H, (1,2)H, (1,3)H\}$ (why ?). Proceeding, we have that $\mathbb{C}\mathcal{H} = \{c_1\boldsymbol{H} + c_2(\mathbf{1},\mathbf{2})\boldsymbol{H} + c_3(\mathbf{1},\mathbf{3})\boldsymbol{H}\}$ and computing the matrix of $(1,2)$ on the basis $\boldsymbol{H}, (\mathbf{1},\mathbf{2})\boldsymbol{H}, (\mathbf{1},\mathbf{3})\boldsymbol{H}$, we obtain that

$$(1,2)\boldsymbol{H} = (\mathbf{1},\mathbf{2})\boldsymbol{H}, \quad (1,2)(\mathbf{1},\mathbf{2})\boldsymbol{H} = \boldsymbol{H}, \quad (1,2)(\mathbf{1},\mathbf{3})\boldsymbol{H} = (\mathbf{1},\mathbf{2},\mathbf{3})\boldsymbol{H} = (\mathbf{1},\mathbf{3})\boldsymbol{H}$$

(Exercise: check the last equality). So the matrix $X((1,2))$ is

$$X((1,2)) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which is the corresponding matrix in the defining representation.

2.3. Reducibility and Maschke's theorem.

One of the purposes of representation is to understand the decomposition to invariant spaces under the group action. This is related to the notion of *irreducible representations.* To formulate this, we start with the definition

Definition 2.9. (invariant subspaces). *Let V be a G-module. A* **submodule** *W of V is a subspace that is closed under the action of G, that is, if $\boldsymbol{w} \in W$ then for all $g \in G$ it holds that $g\boldsymbol{w} \in W$. W is then called a G-invariant subspace.*

Example 7. Let $G = S_n$ and $V = \mathbb{C}\{\mathbf{1}, \mathbf{2}, ..., \boldsymbol{n}\}$ and $w = \mathbb{C}\{\mathbf{1} + \mathbf{2} + \cdots + \boldsymbol{n}\}$. W is S_n invariant as for any $\pi \in S_n$:

$$\pi(\mathbf{1} + \mathbf{2} + \cdots + \boldsymbol{n}) = \pi(\mathbf{1}) + \pi(\mathbf{2}) + \cdots + \pi(\boldsymbol{n}) = \mathbf{1} + \mathbf{2} + \cdots + \boldsymbol{n}.$$

So W if invariant. Moreover, we can ask what is the representation of the restriction of S_n on W and since for any $\boldsymbol{w} \in W$ we have seen that $\pi(\boldsymbol{w}) = \boldsymbol{w}$, it follows that $X(\pi) = 1 \in \mathbb{R}$, thus, the trivial representation. *In this case (and in general in similar situations) we say that $G(= S_n)$ acts trivially on W.*

Exercise 4. *Show that the sign representation of S_n can be recovered by using the submodule*

$$W = \mathbb{C}\big[\sum_{\pi \in S_n} sgn(\pi)\boldsymbol{\pi}\big].$$

Definition 2.10. *A nonzero G-module V is called* **reducible** *if it contains a non-trivial submodule W. Otherwise, it is called* **irreducible**.

Proposition 2.11. *V is reducible iff it has a basis $\mathcal{B}$ in which every $g \in G$ has a block matrix representation of the form*

$$X(g) = \begin{pmatrix} A(g) & B(g) \\ O & C(g) \end{pmatrix}$$

where $A(g), C(g)$ are square matrices, of the same size for all g, and O is a zero matrix.

Proof. Assume that V has a nontrivial submodule W and let $d > f > 0$ be the dimensions of V, W, respectively. Let $\{\boldsymbol{w}_1, ..., \boldsymbol{w}_f$ be a basis of W and complement this to a basis of V as $\{\boldsymbol{w}_1, ..., \boldsymbol{w}_f, \boldsymbol{v}_{f+1}, ..., \boldsymbol{v}_d\}$. In this basis, compute the matrix representation $X(g)$ for arbitrary $g \in G$. Since W is an invariant submodule, it follows that $g\boldsymbol{w}_i \in W$ for any $i = 1, ..., f$. So $g\boldsymbol{w}_i$ will be written as a linear combination of $\{\boldsymbol{w}_1, ..., \boldsymbol{w}_f\}$, as identified but the square matrix $A(g)$, and will not involve the rest of the basis vectors, in other words, the coordinates of $g\boldsymbol{w}_i$ from $f + 1$ to d will all be zero, hence the presence of the 0 matrix O.

In the opposite direction, assume that $X(g)$ has the above block form. Then it suffices to consider the subpace spanned by vectors of the standard basis $\{\boldsymbol{e}_1, ..., \boldsymbol{e}_f\}$, i.e. $\boldsymbol{e}_i$ has a single 1 entry at position i. Then this is an invariant sub-module of V, which is nontrivial if O is also nontrivial. ☐

2.3.1. Maschke's theorem. We will next move towards stating Maschke's theorem, which is, in a sense, is a *global version* of the Jordan canonical decomposition from linear algebra. It says that any representation matrix $X(g)$ is a conjugate of a block-diagonal form

$$\begin{pmatrix} A_1(g) & & \\ & A_2(g) & \\ & & \ddots \end{pmatrix}$$

where the empty spaces correspond to zero matrices and $A_1, A_2, ...$ are square matrices. Equivalently, it says that

Theorem 2.12. (Maschke's theorem) *Let G be a finite group and V a nonzero G-module. Then*

$$V = W^{(1)} \oplus \cdots \oplus W^{(m)},$$

with $W^{(i)}$ irreducible G-submodules.

The equivalent matrix formulation is the following:

Theorem 2.13. *Let G be a finite group and let X be a matrix representation G of dimension $d > 0$. Then there is a matrix T such that for every $g \in G$, $X(g)$ satisfies*

$$TX(g)T^{-1} = \begin{pmatrix} X^{(1)}(g) & 0 & \cdots & 0 \\ 0 & X^{(2)}(g) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & X^{(m)}(g) \end{pmatrix}$$

where $X^{(i)}(g), i = 1, ..., m$ *are irreducible representations of* G.

Proof. Let $V = \mathbb{C}^d$ and consider the action

$$\boldsymbol{g}\boldsymbol{v} = X(g)\boldsymbol{v}, \qquad \text{for all } g \in G \text{ and } \boldsymbol{v} \in V.$$

By Maschke's theorem V decomposes into irreducibles as $V = W^{(1)} \oplus \cdots \oplus W^{(m)}$. Consider, now, a basis for V which consists of the basis vectors of each $W^{(i)}, i = 1, ..., m$. The matrix T which has as columns the above vectors transforms the standard basis of $\mathbb{C}^d$ to the above basis and by standard linear algebra the result follows. Because the submodule are G-invariant the representation is the same for any $g \in G$. $\square$

We now want to ask the question: *when two representations are the same ?*
For this we first need to notion of G-**homomorphisms**.

Definition 2.14. *Let* V, W *be* G*-modules.* $\vartheta \colon V \to W$ *is a* G*-homomorphism if it is a linear transformation such that*

$$\vartheta(g\boldsymbol{v}) = g\vartheta(\boldsymbol{v}), \qquad \textit{for all } g \in G, \boldsymbol{v} \in V.$$

In other words ϑ preserves the action of g. In the language of matrices, if T is the transfer matrix between the basis of V and W, then the G-homomorphism property translates to

$$TX(g) = Y(g)T, \qquad \text{for all } g \in G. \tag{2.3}$$

Two representations will be equivalent if, moreover, ϑ is a bijection or equivalently T is invertible, in which case

$$Y(g) = TX(g)T^{-1}, \qquad \text{for all } g \in G.$$

Remark 2.15. Relation (2.3) is often called **intertwining** and we will see version of this in the integrable probability and integrable Markovian dynamics parts of the course.

Associated to ϑ are two invariant sup-modules: the kernes

$$ker(\vartheta) := \{\boldsymbol{v} \in V \colon \vartheta(\boldsymbol{v}) = 0\},$$

and the image

$$im(\vartheta) := \{\boldsymbol{w} \in W \colon \boldsymbol{w} = \vartheta(\boldsymbol{v}), \text{ for some } \boldsymbol{v} \in V\}.$$

The fact that $ker(\vartheta)$ is a sub-module follows easily, since by the homomorphism property of ϑ we have that

$$\vartheta(g\boldsymbol{v}) = g\vartheta(\boldsymbol{v}) = g\mathbf{0} = \mathbf{0},$$

and so also $\vartheta(\boldsymbol{v}) \in ker(\vartheta)$. As an exercise check that $im(\vartheta)$ is also a sub-module. The invariance of the above spaces is used in the proof of Schur's lemma

Lemma 2.16. (Schur) *Let* V, W *be two irreducible* G*-modules. If* $\vartheta \colon V \to W$ *is a* G*-homomorphism, then either*

- ϑ *is a* G*-isomorphism or*
- $\vartheta \equiv 0$.

Proof. Since V is irreducble and $ker\vartheta$ is a submodule it must be that either $ker(\vartheta) = \{\mathbf{0}\}$ or $ker(\vartheta) = V$. Similarly, for $im(\vartheta) = \{\mathbf{0}\}$ or $im(\vartheta) = W$. If $ker(\vartheta) = V$ or $im(\vartheta) = \{\mathbf{0}\}$, then $\vartheta \equiv \mathbf{0}$ and if $ker(\vartheta) = \{\mathbf{0}\}$ and $im(\vartheta) = W$ then ϑ is a bijection. $\square$

Translated to matrices, Schur lemma says that if X, Y are two irreducible matrix representations, which are equivalent, that is, there exists a matrix T such that $X(g)T = TY(g)$ for every $g \in G$, then either $T = 0$ or T is invertible. In the latter case the two representations are related as $X(g) = TY(g)T^{-1}$.

The following corollary of Schur's lemma will be useful in the orthogonality property of characters that will follow.

Corollary 2.17. *Let X be an irreducible matrix representation of a group G such that $TX(g) = X(g)T$ for every $g \in G$. Then $T = cI$ for some $c \in \mathbb{C}$ and I the identity matrix.*

Proof. The commutation relations can be easily extended to $(T - cI)X(g) = X(g)(T - cI)$ for any $c \in \mathbb{C}$. Schur's lemma imply that either $T - cI = 0$ or that $T - cI$ is invertible, since X is irreducible. But we can take c to be an eigenvalue of T and then it must be that $T - cI = 0$. □

2.4. Group characters. In this section we will define the notion of group characters, an object that contains a lot of information about the representation of the group. Characters are just traces of the corresponding matrix representations. More precisely,

Definition 2.18. *Let $X(g), g \in G$ be a matrix representation of of G. Then the character of the representation is defined as*

$$\chi(g) := \mathrm{Tr} X(g)$$

where Tr *stands for the trace of matrix X.*

The notion of a character is well defined in the sense that if two representations X, Y are equivalent, then they must have the same trace. This is because in such a case $X(g) = TY(g)T^{-1}$ and then $\chi_X(g) := \mathrm{Tr}(TY(g)T^{-1}) = \mathrm{Tr}(Y(g)) = \chi_Y(g)$ by the cyclic property of the trace. We will later prove that the converse also holds, i.e. if two representations have the same character, then they much be equivalent.

Exercise 5. *Let χ^{def} be the character corresponding to the defining representation of S_n given in Example 4. Show that for every $\pi \in S_n$ it holds that $\chi^{def}(\pi)$ equals the number of fixed points of permutation π.*

The following proposition gives some more reason why characters contain a lot of information. The proof, which is omitted is, again, a consequence of the traces.

Proposition 2.19. *Let X be a matrix representation of a group G of degree d with character χ. Then $\chi(\varepsilon) = d$. Moreover, if K is a conjugacy class of G, then $\chi(g) = \chi(h)$ for every $g, h \in K$.*

We now define an inner product:

Definition 2.20. *Let χ, ψ be two functions on a group G. We define the inner product $\langle \cdot, \cdot \rangle$*

$$\langle \chi, \psi \rangle := \frac{1}{|G|} \sum_{g \in G} \chi(g) \overline{\psi(g)}.$$

It turns out (explain) that when χ, ψ are characters, then there is an equivalent formulation of the above inner product, taking the form

$$\langle \chi, \psi \rangle := \frac{1}{|G|} \sum_{g \in G} \chi(g) \psi(g^{-1}).$$

The main theorem is that irreducible characters are orthogonal with respect to this inner product. This orthogonality is closely related to certain fundamental identities that symmetric functions satisfy (called **Cauchy identities**) and which is one of the pillars of Integrable Probability.

Theorem 2.21. *Let χ, ψ irreducible characters of a group G. Then*

$$\langle \chi, \psi \rangle = \delta_{\chi, \psi}.$$

Proof. The proof makes use of Schur's lemma.

Consider an abritrary $d \times f$ matrix $X = (x_{i,j})_{d,f}$, where $x_{i,j}$ are viewed as arbitrary variables (indeterminates). If A, B are matrix representations corresponding to characters χ, ψ, then consider the matrix

$$Y = \frac{1}{|G|} \sum_{g \in G} A(g) X B(g^{-1}).$$

We will check that for every $h \in G$ it holds that $A(h)Y = YB(h)$. Indeed,

$$\begin{aligned} A(h)YB(h)^{-1} &= \frac{1}{|G|}\sum_{g\in G} A(h)A(g)XB(g^{-1})B(h^{-1}) \\ &= \frac{1}{|G|}\sum_{g\in G} A(hg)XB(g^{-1}h^{(-1)}) \\ &= \frac{1}{|G|}\sum_{\widetilde{g}=hg\in G} A(\widetilde{g})XB(\widetilde{g}^{-1}) \\ &= Y. \end{aligned}$$

Therefore, by Corrollary 2.17 we have the triviality of Y, i.e.

$$Y = \begin{cases} 0, & \text{if } A, B \text{ are inequivalent,} \\ cI_d, & \text{if } A, B \text{ are equivalent.} \end{cases} \tag{2.4}$$

Let us now translate what this means. First, consider the case where A, B are inequivalent, i.e. $\chi \neq \psi$. Then we can rewrite the first branch of the above equality as

$$Y_{i,j} = \frac{1}{|G|}\sum_{k,\ell}\sum_{g\in G} a_{i,k}(g)\, x_{k,\ell}\, b_{\ell,j}(g^{-1}) = 0,$$

and since $(x_{k,\ell})$ are arbitrary and the above equality holds for all indeterminates $x_{k,\ell}$, it means that all their coeffocients must be 0, which translates to

$$\frac{1}{|G|}\sum_{g\in G} a_{i,k}(g)b_{\ell,j}(g^{-1}) = 0.$$

Viewing $a_{i,k}(\cdot), b_{\ell,j}(\cdot)$ as functions on G the above is the expression of the inner product, thus,

$$\langle a_{i,k}, b_{\ell,j}\rangle = 0, \qquad \text{for all } i, k, \ell, j.$$

In particular,

$$0 = \sum_i\sum_j \langle a_{i,i}, b_{j,j}\rangle = \langle \sum_i a_{i,i}, \sum_j b_{j,j}\rangle = \langle \mathrm{Tr}A, \mathrm{Tr}B\rangle = \langle \chi, \psi\rangle,$$

which shows the orthogonality of inequivalent characters.

It remains to show that $\langle \chi, \chi\rangle = 1$. This follows by similar considerations: In the above formulation, let $\chi = \psi$, i.e. the corresponding matrix representations A, B are equivalent. Without loss of generality we can take $A = B$. By (2.4) we have that

$$Y := \frac{1}{|G|}\sum_{g\in G} A(g)XA(g^{-1}) = cI_d,$$

and taking traces it follows that

$$cd = \frac{1}{|G|}\sum_{g\in G} \mathrm{Tr}\Big(A(g)XA(g^{-1})\Big) = \frac{1}{|G|}\sum_{g\in G} \mathrm{Tr}X = \mathrm{Tr}X$$

since X does not depend on G. The two above relations give that $y_{i,i} = \frac{1}{d}\mathrm{Tr}(X)$ or that

$$\frac{1}{|G|}\sum_{k,\ell}\sum_{g\in G} a_{i,k}(g)x_{k,\ell}a_{\ell,i}(g^{-1}) = \frac{1}{d}(x_{1,1} + \cdots + x_{d,d}),$$

which, after equating the coefficient of the indeterminates $x_{i,j}$, gives

$$\langle a_{i,k}, a_{\ell,i}\rangle := \frac{1}{|G|}\sum_{g\in G} a_{i,k}(g)\, a_{\ell,i}(g^{-1}) = \frac{1}{d}\delta_{k,\ell},$$

and translating to a character relation:

$$\langle \chi, \chi \rangle = \langle \mathrm{Tr} A, \mathrm{Tr} A \rangle = \sum_{1 \le i,j \le d} \langle a_{ii}, a_{jj} \rangle = \sum_{i=1}^{d} \langle a_{ii}, a_{ii} \rangle = \sum_{i=1}^{d} \frac{1}{d} = 1.$$

This completes the proof. □

2.4.1. Prelude to symmetric functions. This is now a good point to stop the general theory of representations and move to symmetric functions. The link is that the characters of the symmetric group, which have a nice expression in terms of Young tableaux. One thing that would have been nice to do but, unfortunately, won't have time is to show that Young diagrams classify the irreducible representations of the symmetric group. Young diagrams are left justified array of boxes, indexed by a partition λ and filled with numbers, eg

$\lambda = (4, 3, 1) \quad \longleftrightarrow$

1	1	2	2	3
2	2	3		
3				

The number of times α_i that the integer i appears in the Young diagram is called the **type** of the Young diagram and it is related to the type of the associated permutation.

We will (probably) see that the character of a permutation λ with type α can be expressed in terms of a variation of the Young diagram, called **border-strip tabaleaux** defined by

- every row and column has integer numbers that are weakly increasing
- integer i appears α_i times
- the set of boxes filled with i forms a **border strip**, ie there is no 2×2 block of boxed occupied by the same number. For example:

1	1	1	1	5	5	5
1	2	2	4	5		
3	3	4	4	5		
3	4	4				

Then the characters of the symmetric group admit the representation:

$$\chi^{\lambda}(\alpha) = \sum_{T} (-1)^{\mathrm{height}(T)} \tag{2.5}$$

where the sum is over all border-strip tableaux and the height height(T) of such is defined as

$$\mathrm{height}(T) := \sum_{i} \mathrm{height}(B_i), \qquad B_i \text{ is a border strip of numbers } i,$$
$$\mathrm{height}(B_i) := \{\text{the vertical length of strip } B_i\} - 1.$$

The characters of the symmetric group are also related to Schur functions, eg they are linked via the relation

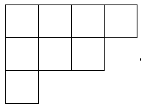
$$s_{\mu} \, p_{\alpha} = \sum_{\lambda} \chi^{\lambda/\mu}(\alpha) s_{\mu}$$

where p_{α} are the *power symmetric polynomials*, the sum is over partitions λ and λ/μ are *skew partitions*. But now it is time to move on to symmetric functions and explain all these terms...

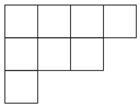
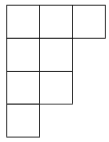
3. Symmetric functions

Symmetric functions have many incarnations: they are related to characters, they can represent solutions to PDE problems, can arise as generating functions of combinatorial problems and many more. An important point is to be able to describe bases in the space of symmetric functions and how you can you change in between bases.

3.1. Preparation. Let us start by reminding some things about partitions and also introduce some related quantities. $\lambda = (\lambda_1, \lambda_2, ...)$ is a sequence of nonnegative integers (called **"parts"**), ordered in decreasing order. We say that $\lambda \vdash$ is a partition of n is $n = \lambda_1 + \lambda_2 + \cdots$. Partitions can be represented by Young diagrams as

$\lambda = (4, 3, 1) \equiv$  .

Given a partition λ we also define the **conjugate** partition, denoted by λ' as the partition we get by the transpose Young diagram. For example if $\lambda = (4, 3, 1)$ then $\lambda' = (3, 2, 2, 1)$

if $\lambda = (4, 3, 1) \longleftrightarrow$  then $\lambda' = (3, 2, 2, 1) \longleftrightarrow$ 

Recall that we can also represent a partition in terms of its type $\lambda = (1^{m_1} 2^{m_2} \cdots)$, where m_i it number of parts of λ which are equal to i. Finally, we denote the number of nonzero parts of a partition λ by $\ell(\lambda)$ and we call it the **length of the partition**. The length of a partition coincides with the number of rows in a Young diagram.

Exercise 6. *Given a partition λ with type $(m_1, m_2, ...)$, express the type of the partition λ' in terms λ.*

We can put two **partial orderings** on partitions, which are useful when running induction arguments.

The first one is called the **containment ordering**: If μ, λ partitions, then we define the ordering $\mu \subset \lambda$, if $\mu_i \le \lambda_i$ for all i. In terms of Young diagrams, this is that the μ Young diagram is contained in the λ diagram.

The second ordering is the **dominance ordering**: We say that $\mu \le \lambda$ in *dominance order* if

$$\mu_1 + \cdots + \mu_i \le \lambda_1 + \cdots + \lambda_i, \qquad \text{for every } i \tag{3.1}$$

3.2. Basic symmetric functions - definitions. Let us start by introducing four families of basic symmetric functions: the monomial, elementary, homogeneous and power sums. All these form a basis of the ring of symmetric functions. We will discuss this fact as well as how to transfer between these basis and a number of fundamental identities involving these. The type of identities will be a preparation for the more fundamental identities (eg Cauchy identity) that will follow.

The most basic symmetric functions are the **monomial symmetric functions**. They are defined as follows: For a partition $\lambda = (\lambda_1, \lambda_2, ...)$, introduce the monomial $x^\lambda = x_1^{\lambda_1} x_2^{\lambda_2} \cdots$. Then the monomial symmetric function m_λ, indexed by λ, is defined as the sum of all distinct monomials obtained from x^λ by permuting the indeterminates $x_1, x_2, ...$ More precisely, when we restrict to n indeterminates, $x_1, x_2, ..., x_n$, we have

$$m_\lambda(x_1, ..., x_n) := \sum_{\sigma \in S_n} \sigma(x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_n^{\lambda_n}) = \sum_{\sigma \in S_n} x_{\sigma(1)}^{\lambda_1} x_{\sigma(2)}^{\lambda_2} \cdots x_{\sigma(n)}^{\lambda_n}$$

For example $m_{(2,1,1)}(x_1, x_2, ...) = \sum_{i \ne j \ne k} x_i^2 x_j x_k$. The definition can be extended to an infinite number of indeterminates as

$$m_\lambda(x_1, x_2 ...,) := \sum_{\sigma \in S_\infty} \sigma(x_1^{\lambda_1} x_2^{\lambda_2} \cdots = \sum_{\sigma \in S_\infty} x_{\sigma(1)}^{\lambda_1} x_{\sigma(2)}^{\lambda_2} \cdots$$

When $\lambda = (n)$, that is, the partition has only one part of length r, then $m_{(n)}$ becomes the **power symmetric function**

$$p_n(x_1, x_2, ...) = \sum_i x_i^n. \tag{3.2}$$

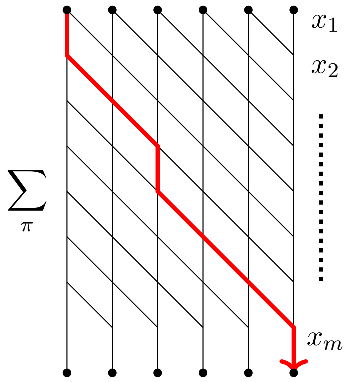
The definition can be extended to partitions λ as

$$p_\lambda(x_1, x_2, ...) = \prod_{i \ge 1} p_{\lambda_i}(x_1, x_2, ...).$$

The **elementary symmetric functions** can be deduced from the monomial ones, if $\lambda = (1^n)$, i.e. the partition has n parts of size 1. The corresponding elementary symmetric function is

$$e_n(x_1, x_2, ...) = \sum_{i_1<i_2<\cdots<i_n} x_{i_1} x_{i_2} \cdots x_{i_r}. \tag{3.3}$$

Path representations of symmetric functions will be very important. So let us start by giving the path representation of the elementary function of n variables e_n. This looks as



where $m \geq n$ and the sum is over all down and diagonally-right paths π from the upper left corner to the down right. Paths are given weights 1 at each diagonal step and x_i at each vertical step at level i (from the top). The total weight of a path is the product of the individual steps. We remark that we drew the paths going downwards but we could equivalently have drawn the picture flipped, in which case paths would be going upwards. Paths of the above type are often called "*strict-weak*" paths. Elementary symmetric functions indexed by more general partition $\lambda = (\lambda_1, \lambda_2, ...)$ is defined as

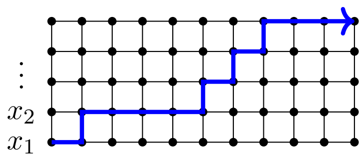
$$e_\lambda = e_{\lambda_1} e_{\lambda_2} \cdots = \prod_{i\geq 1} e_{\lambda_i}. \tag{3.4}$$

The **complete homogeneous symmetric functions** of degree n are the sum of all monomials of degree r, that is

$$h_n(x_1, x_2, ...) = \sum_{\lambda \vdash n} m_\lambda(x_1, x_2, ...) = \sum_{i_1 \leq i_2 \leq \cdots \leq i_n} x_{i_1} x_{i_2} \cdots x_{i_n}, \tag{3.5}$$

The graph representation is as follows:

$$h_n(x_1, x_2, ...) = \sum_\pi \tag{3.6}$$



where the sum is over all paths that go up and right and make n horizontal steps. Every vertical step is given weight 1 and every horizontal step at level i is given weight x_i. As before, for a parition $\lambda = (\lambda_1, \lambda_2, ...)$ we define

$$h_\lambda(x_1, x_2, ...) = \prod_{i\geq 1} h_{\lambda_i}(x_1, x_2, ...).$$

3.3. Basic symmetric functions - first properties and identities.

Proposition 3.1. *If λ is a partition of n, then*

$$e_\lambda = \sum_{\mu \vdash n} M_{\lambda\mu} m_\mu, \tag{3.7}$$

where $M_{\lambda\mu}$ is the number of infinite matrices $(a_{i,j})_{i,j\geq 1}$ with $0, 1$ entries, such that $\sum_{j\geq 1} a_{i,j} = \lambda_i$ and $\sum_{i\geq 1} a_{i,j} = \mu_j$

Proof. If $\lambda = \lambda_1 \geq \lambda_2 \geq \cdots$, we have that

$$\begin{aligned} e_\lambda = \prod_{i\geq 1} e_{\lambda_i} &= \prod_{i\geq 1} \sum_{S_i \subset \{1,2,..\}, |S_i|=\lambda_i} \prod_{j\geq 1} x_j^{\mathbb{1}_{j\in S_i}} \\ &= \sum_{\substack{S_i\subset\{1,2,..\},\\ |S_i|=\lambda_i\,, i\geq 1}} \prod_{i\geq 1}\prod_{j\geq 1} x_j^{\mathbb{1}_{j\in S_i}} \\ &= \sum_{\substack{S_i\subset\{1,2,..\},\\ |S_i|=\lambda_i\,, i\geq 1}} \prod_{j\geq 1} x_j^{\sum_{i\geq 1}\mathbb{1}_{j\in S_i}} \end{aligned}$$

The result follows by setting $a_{i,j} := \mathbb{1}_{j\in S_i}$ and the change of variables $\alpha_j := \sum_{i\geq 1}\mathbb{1}_{j\in S_i}$ and noting that

$$\sum_i a_{i,j} = \sum_j \mathbb{1}_{j\in S_i} = |S_i| = \lambda_i,$$

therefore

$$e_\lambda = \sum_{\alpha_1,\alpha_2,\ldots\in\mathbb{N}} \sum_{A=(a_{i,j})_{i,j\geq 1}} \mathbb{1}_{\{\mathrm{row}(A)=\lambda, \mathrm{col}(A)=\alpha\}} \prod_{i\geq 1} x_i^{\alpha_i}$$

where for a matrix A we define $\mathrm{row}(A)_i := \sum_j a_{i,j}$ and $\mathrm{col}(A)_j := \sum_i a_{i,j}$. Finally, rewrite the above sum in terms of partitions μ by rearranging the $\alpha_1, \alpha_2, \ldots$ in decreasing order and this will yield (3.7). ☐

Remark 3.2. Relation (3.7) gives the way to change between the basis (m_λ) and the basis (e_λ) (it will turn out that the latter is also a basis for symmetric functions) and the matrix $M_{\lambda\mu}$ is called the *transition matrix* between the bases.

Exercise 7. *Show that the transition matrix $(M_{\lambda\mu})_{\lambda,\mu}$ is symmetric.*

Exercise 8. *The transfer matrix $(M_{\lambda\mu})$ has the following combinatorial interpretation: Entry $M_{\lambda\mu}$ is the number of ways to place n balls with λ_i of them labeled i, for $i \geq 1$, into boxes $1, 2, 3, \ldots$ such that no box contains two balls with the same label and box i contains exactly μ_i balls.*

The next proposition is a baby Cauchy identity.

Proposition 3.3. *Let $(x_i)_{i\geq 1}$ and $(y_i)_{i\geq 1}$ be indeterminates. Then*

$$\prod_{i,j}(1 + x_i y_j) = \sum_\lambda m_\lambda(x) e_\lambda(y) \tag{3.8}$$

$$= \sum_{\lambda,\mu} M_{\lambda\mu} m_\lambda(x) m_\mu(y). \tag{3.9}$$

where the sum is over all partitions λ, μ.

Proof. Here we can take a first glimpse of the power of probabilistic thinking. Let us define the *independent* "random variables"

$$a_{i,j} = \begin{cases} 1, & \text{with weight } 1, \\ 0, & \text{with weight } 1, \end{cases}$$

note that to really talk about probabilities we should be normalising the weights to $\frac{1}{2}, \frac{1}{2}$ but because the product in the left-hand side of (3.8) is infinite we want to avoid normalisations of the sort 2^∞, so we just talk about weights instead of probabilities. In this formulation, in any way, we can define the "expectation" (total weight) of $\mathbb{E}[(x_i y_i)^{a_{i,j}}] := 1 + x_i y_j$ and then we can rewrite the left-hand side of (3.8) as

$$\prod_{i,j}(1 + x_i y_j) = \prod_{i,j} \mathbb{E}[(x_i y_j)^{a_{i,j}}] = \mathbb{E}\big[\prod_{i,j}(x_i y_j)^{a_{i,j}}\big],$$

where in the second equality we used the independence (or in other words Fubini). We can proceed by writing

$$\begin{aligned}\prod_{i,j}(1+x_iy_j) &= \mathbb{E}\Big[\prod_i x_i^{\sum_j a_{i,j}} \prod_j y_j^{\sum_i a_{i,j}}\Big] \\ &= \mathbb{E}\Big[\prod_i x_i^{\mathrm{row}(A)_i} \prod_j y_j^{\mathrm{col}(A)_j}\Big]\end{aligned}$$

□

where, again, for a matrix A we define $\mathrm{row}(A)_i := \sum_j a_{i,j}$ and $\mathrm{col}(A)_j := \sum_i a_{i,j}$. Now, writing out the "expectation" in terms of sums we have

$$\begin{aligned}\prod_{i,j}(1+x_iy_j) &= \sum_{\substack{\lambda_1,\lambda_2,\dots \\ \mu_1,\mu_2,\dots}} \sum_{\substack{A:\ \{0,1\}\text{ matrices} \\ \mathrm{row}(A)_i=\lambda_i,\,\mathrm{col}(A)_j:=\mu_j}} \prod_i x_i^{\lambda_i} \prod_j y_j^{\mu_j}. \\ &= \sum_{\substack{\lambda_1,\lambda_2,\dots \\ \mu_1,\mu_2,\dots}} M_{\lambda\mu} m_\lambda(x)\, m_\mu(y) \\ &= \sum_\lambda m_\lambda(x) e_\lambda(y),\end{aligned}$$

where in the penultimate equality we used the definition of $M_{\lambda\mu}$ and in the last we used Proposition 3.1.

Proposition 3.4. *Let $\lambda,\mu \vdash n$. Then $M_{\lambda,\mu} = 0$ unless $\mu \le \lambda'$, where λ' is the conjugate partition of λ and $\le$ is the dominance order as in* (3.1)*. Recall the definition of $M_{\lambda,\mu}$ from Proposition 3.1. In particular, this implies that the family of elementary symmetric functions is a basis.*

Proof. Recall from Proposition 3.1 that $M_{\lambda,\mu}$ is the number of $\{0,1\}$ matrices with row sums $\mathrm{row}(A) = \lambda$ and column sums $\mathrm{col}(A) = \mu$. Assume that $M_{\lambda,\mu} \neq 0$. Then for every $\{0,1\}$ matrix A with row sums $\mathrm{row}(A) = \lambda$ and column sums $\mathrm{col}(A) = \mu$, consider the matrix A', which has all $1's$ in every row aligned to the left and row i has λ_i entries equal to 1 for example

$$A = \begin{pmatrix} 0&1&1&0&1&0&1&0 \\ 1&0&1&1&0&0&0&0 \\ 1&1&0&0&0&1&0&0 \\ 1&0&0&0&0&0&1&0 \\ 0&0&0&0&1&0&0&0 \end{pmatrix} \quad \text{then } A' = \begin{pmatrix} 1&1&1&1&0&0&0&0 \\ 1&1&1&0&0&0&0&0 \\ 1&1&1&0&0&0&0&0 \\ 1&1&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0 \end{pmatrix}$$

For every i the numbers of $1's$ in the first i columns of A' is larger or equal to the number of $1's$ in the first i columns of A. This implies that

$$\sum_{j=1}^{i} \mathrm{col}(A')_j \ge \sum_{j=1}^{i} \mathrm{col}(A)_j. \tag{3.10}$$

Now, $\mathrm{col}(A)_j = \mu_j$ and $\mathrm{col}(A')_j = \lambda'_j$, The latter is because the $1's$ in A' form a Young diagram of shape $\lambda_1,\lambda_2,\dots$ and conjugating this Young diagram we get the Young diagram with shape $\lambda'_1,\lambda'_2,\dots$ formed by the $1's$ in the columns of A'. Relations (3.10) imply that $\mu \le \lambda'$ in the dominance order and so $M_{\lambda,\mu} = 0$ unless $\mu \le \lambda'$. This means that the matrix $(M_{\lambda,\mu})_{\lambda,\mu}$ is triangular. Moreover, $M_{\lambda,\lambda'} = 1$ since A' is the only matrix A with $\mathrm{row}(A) = \lambda$ and $\mathrm{col}(A') = \lambda'$. These two facts imply that $(e_\lambda)_\lambda$ form a basis since $(m_\lambda)\lambda$ is a basis and the transfer matrix $M_{\lambda,\mu}$ is invertible. □

The triangularity of the transfer matrix $M_{\lambda\mu}$ can be seen the following table:

$$\begin{aligned} e_1 &= m_1, \\ e_{11} &= m_2 + 2m_{11} \\ e_2 &= \qquad m_{11} \\ e_{111} &= m_3 + 3m_{21} + 6m_{111} \\ e_{21} &= \qquad m_{21} + 3m_{111} \\ e_3 &= \qquad\qquad m_{111} \end{aligned}$$

The complete homogeneous functions satisfy the analogous to Propositions 3.1 and 3.3 properties:

Proposition 3.5. *Let $\lambda \vdash n$. Then*

$$h_\lambda = \sum_{\mu \vdash n} N_{\lambda\mu} m_\mu$$

where $N_{\lambda\mu}$ is the number of $\{0, 1, 2, ...\}$ matrices with $\mathrm{row}(A) = \lambda$ *and* $\mathrm{col}(A) = \mu$. *The transfer matrix $N_{\lambda\mu}$ is symmetric. Moreover, the following identity holds*

$$\begin{aligned} \prod_{i,j} (1 - x_i y_j)^{-1} &= \sum_{\lambda,\mu} N_{\lambda\mu} m_\lambda(x) m_\mu(y) \\ &= \sum_\lambda m_\lambda(x) h_\lambda(y). \end{aligned}$$

Exercise 9. *Prove proposition 3.5.*

Proving that (h_λ) is a basis is a bit more tricky as the transfer matrix $N_{\lambda\mu}$ is not triangular. For example, we have

$$\begin{aligned} h_1 &= m_1, \\ h_{11} &= 2m_{11} + m_2 \\ h_2 &= m_{11} + m_2 \\ h_{111} &= 6m_{111} + 3m_{21} + m_3 \\ h_{21} &= 3m_{111} + 2m_{21} + m_3 \\ h_3 &= m_{111} + m_{21} + m_3. \end{aligned}$$

To prove the fact that (h_λ) form also a basis, we introduce a useful *involution*, which we denote by ω.

Definition 3.6. *Define the homomorphism $\omega \colon \Lambda \to \Lambda$ (with Λ the ring of symmetric functions, so in this case a homomorphism is also an endomorphism), so that for every n we have the mapping $\omega(e_n) = h_n$.*

By the fact that ω is a homomorphism, we have that $\omega(e_\lambda) = h_\lambda$. We will show that ω is also an involution, i.e. $\omega^2 = 1$, so that it is invertible and also $\omega(h_\lambda) = e_\lambda$.

Proposition 3.7. *The endomorphism ω is an involution.*

Proof. Consider the generating functions

$$H(t) := \sum_n h_n(x) t^n, \qquad \text{and} \qquad E(t) := \sum_n e_n(x) t^n, \tag{3.11}$$

where remember that $x = (x_1, x_2, ...)$. It turns out that

$$H(t) = \prod_n (1 - x_n t)^{-1}, \qquad \text{and} \qquad E(t) = \prod_n (1 + x_n t), \tag{3.12}$$

from which it follows that $H(t)E(-t) = 1$. From this and looking at the coefficients of t^n, we derive that

$$0 = \sum_{i=0}^{n} (-1)^i e_i h_{n-i}, \qquad \text{for} \qquad n \geq 1.$$

Apply ω on this equation to obtain

$$0 = \sum_{i=0}^{n} (-1)^i \omega(e_i\, h_{n-i}) = \sum_{i=0}^{n} (-1)^i \omega(e_i)\, \omega(h_{n-i}) = \sum_{i=0}^{n} (-1)^i h_i\, \omega(h_{n-i}) = (-1)^n \sum_{i=0}^{n} (-1)^i \omega(h_i)\, h_{n-i}$$

where the last equality is a simple change of variables in the summation.

An easy argument (explain why !) is that u_i is a sequence of symmetric functions such that $\sum_{i=0}^{n} (-1)^i u_i h_{n-i} = 0$ for all $n \geq 1$ and with $u_0 = 1$, then it must be that $u_i = e_i$. From this and the last sequence of equalities, it follows that $\omega(h_i) = e_i$. Therefore $\omega^2(e_i) = \omega(h_i) = e_i$ and the involution follows from the homomorphism property and the fact that e_λ form a basis. □

We can now show

Proposition 3.8. (h_λ) *forms a basis of symmetric functions.*

Proof. This follows from the fact that (e_λ) is a basis and also that $\omega(e_\lambda) = h_\lambda$ and that ω is invertible. □

We will close this section with a few identities of the power symmetric functions. The first one, whose proof we omit is

Proposition 3.9. *The following identity between m_λ and p_λ holds:*

$$p_\lambda = \sum_{\mu \vdash n} R_{\lambda\mu} m_\mu,$$

where $R_{\lambda\mu}$ is equal to the number of ordered partitions $\pi = (B_1, ..., B_{\ell(\mu)})$ of the set $\{1, ..., \ell(\mu)\}$, where $\ell(\mu)$ is the number of (nonzero) parts of partition μ† *with*

$$\mu_j = \sum_{i \in B_j} \lambda_i, \qquad \textit{for} \quad 1 \leq j \leq k.$$

The next proposition describes Cauchy-type identities for the power symmetric polynomials. The number

$$z_\lambda = \prod_i i^{m_i} m_i!, \qquad \text{for a partition } \lambda = (1^{m_1} 2^{m_2} \cdots n^{m_n}),$$

which we met at the beginning of the notes in the decomposition of a permutation cycle, will appear. We remind that m_i is the number of parts of λ of length i. The parameter

$$\varepsilon_\lambda := (-1)^{m_2 + m_4 + \cdots}, \tag{3.13}$$

will also appear.

Proposition 3.10. *The following identities hold:*

$$\prod_{i,j} (1 - x_i y_j)^{-1} = \exp\Big\{ \sum_{n \geq 1} \tfrac{1}{n} p_n(x) p_n(y) \Big\} = \sum_\lambda z_\lambda^{-1} p_\lambda(x) p_\lambda(y), \tag{3.14}$$

and

$$\prod_{i,j} (1 + x_i y_j) = \exp\Big\{ \sum_{n \geq 1} \tfrac{(-1)^{n-1}}{n} p_n(x) p_n(y) \Big\} = \sum_\lambda z_\lambda^{-1} \varepsilon_\lambda\, p_\lambda(x) p_\lambda(y), \tag{3.15}$$

---

†note that π is a partition of a set and its "parts" are subsets of the set and not to be confused with the notion of partition λ, μ

Proof. We prove (3.14). The proof of the second equality will be an exercise. Let us first prove the first equality in (3.14). This is essentially Taylor expansion:

$$\begin{aligned}\log\prod_{i,j}(1-x_iy_j)^{-1} &= \sum_{i,j}\log(1-x_iy_j)^{-1}\\ &= \sum_{i,j}\sum_{n\geq 1}\frac{1}{n}x_i^n y_j^n\\ &= \sum_{n\geq 1}\frac{1}{n}\Big(\sum_i x_i^n\Big)\Big(\sum_j y_j^n\Big)\\ &= \sum_{n\geq 1}\frac{1}{n}p_n(x)p_n(y).\end{aligned}$$

The proof of the second identity in (3.14) will take us on a detour on generating functions calculus, which we present in the following sequence of propositions. The proof of (3.14) will finally be presented after Proposition 3.14. ☐

Exercise 10. *Prove* (3.15). *Prove also that* $\varepsilon_\lambda = (-1)^{n-\ell(\lambda)}$*, where* $\ell(\lambda)$ *is the number of nonzero parts of partition* λ.

The following discussion summarises the discussion of **exponential generating function** in Chapter 5 of [S23]. First, we need the definition of exponential generating function:

Definition 3.11. *For* $f\colon \mathbb{N}\to\mathbb{R}$*, we define*

$$E_f(z) = \sum_{n\geq 0} f(n)\frac{z^n}{n!}.$$

Proposition 3.12. *Let* $f_1, f_2, \ldots\colon \mathbb{N}\to\mathbb{R}$ *and define the function* $h_k\colon\mathbb{N}\to\mathbb{R}$ *by*

$$h_k(n) := \sum_{(T_1,\ldots T_k)\in\Pi_{\mathrm{ord}}(n)} f_1(|T_1|)\cdots f_k(|T_k|),$$

where the sum is over all **ordered** *partitions* $\Pi_{\mathrm{ord}}(n)$ *of* $\{1,\ldots,n\}$ *into* k *ordered* k*-tuples of sets* $(T_1,\ldots,T_k)$ *with* $T_1\cup\cdots\cup T_k=\{1,\ldots,n\}$ *and* $T_i\cap T_j=\emptyset$*. Then*

$$E_{h_k}(z) = \prod_{i=1}^k E_{f_i}(z).$$

Proof. The number of parit? partitions: The number of paritions $T_1,\ldots T_k$ of $\{1,\ldots,n\}$ with $|T_j|=t_j$, for $j=1,\ldots,k$, with $t_1+\cdots+t_k=n$ is $\frac{n!}{t_1!\cdots t_k!}$, thus

$$h_k(n) = \sum_{\substack{t_1,\ldots,t_k\\ t_1+\cdot+t_k=n}} \frac{n!}{t_1!\cdots t_k!} f_1(t_1)\cdots f_k(t_k),$$

and

$$\begin{aligned}
E_{h_k}(z) &= \sum_n h_k(n)\frac{z^n}{n!} \\
&= \sum_n \frac{z^n}{n!} \sum_{\substack{t_1,\dots,t_k \\ t_1+\cdot+t_k=n}} \frac{n!}{t_1!\cdots t_k!} f_1(t_1)\cdots f_k(t_k) \\
&= \sum_n \sum_{\substack{t_1,\dots,t_k \\ t_1+\cdot+t_k=n}} \frac{1}{t_1!\cdots t_k!} \prod_{i=1}^k z^{t_i} f_i(t_i) \\
&= \sum_{t_1,\dots,t_k} \sum_n \mathbb{1}_{\{t_1+\cdot+t_k=n\}} \prod_{i=1}^k \frac{z^{t_i}}{t_i!} f_i(t_i) \\
&= \sum_{t_1,\dots,t_k} \prod_{i=1}^k \frac{z^{t_i}}{t_i!} f_i(t_i) \\
&= \prod_{i=1}^k \Big(\sum_{t_i} f_i(t_i) \frac{z^{t_i}}{t_i!} \Big) \\
&= \prod_{i=1}^k E_{f_i}(z).
\end{aligned}$$

□

Proposition 3.13. *If $f, g\colon \mathbb{N} \to \mathbb{R}$ and $g(0) = 1$, define the function*

$$h(n) := \sum_k \sum_{\{B_1,\dots,B_k\}\in\Pi(n)} f(|B_1|)\cdots f(|B_k|)\, g(k), \qquad \text{if } n > 0 \text{ and } h(0) := 1,$$

where the sum is over all partitions $\Pi(n)$ of $\{1,\dots,n\}$, in k disjoints sets $B_1,\dots,B_k$ with $B_1\cup\cdots\cup B_k = n$ and $B_i \cap B_j = \emptyset$. Then

$$E_h(z) = E_g(E_f(z)).$$

Proof. The difference between the statement of this proposition and that of proposition 3.12 is that in this case we considered unordered k-tuples of partitions $\{B_1,\dots,B_k\}$, while in (3.12) all $k!$ permutations of $B_1,\dots,B_k$ are considered different. So

$$\sum_{\{B_1,\dots,B_k\}} f(|B_1|)\cdots f(|B_k|) = \frac{1}{k!} \sum_{(B_1,\dots,B_k)} f(|B_1|)\cdots f(|B_k|),$$

and hence

$$h(n) = \sum_k g(k) \sum_{\{B_1,\dots,B_k\}} f(|B_1|)\cdots f(|B_k|) = \sum_k \frac{g(k)}{k!} \sum_{(B_1,\dots,B_k)} f(|B_1|)\cdots f(|B_k|),$$

Then

$$\begin{aligned}
E_h(z) &= \sum_n h(n)\frac{z^n}{n!} \\
&= \sum_n h(n)\frac{z^n}{n!} \sum_k \frac{g(k)}{k!} \sum_{(B_1,\dots,B_k)\in\Pi_{\mathrm{ord}}(n)} f(|B_1|)\cdots f(|B_k|) \\
&= \sum_k \frac{g(k)}{k!} \sum_n h(n)\frac{z^n}{n!} \sum_{(B_1,\dots,B_k)\in\Pi_{\mathrm{ord}}(n)} f(|B_1|)\cdots f(|B_k|),
\end{aligned}$$

but by Proposition 3.12 the inner sum is just $E_x(x)^k$ and so

$$E_h(z) = \sum_k \frac{g(k)}{k!}\big(E_f(z)\big)^k = E_g\big(E_f(z)\big),$$

which completes the proof. □

Proposition 3.14. *Let $f\colon \mathbb{N} \to \mathbb{R}$ and define the function*

$$h(n) := \sum_k \sum_{\pi\in S_n\,:\,\pi=(c_1,\dots,c_k)} f(|c_1|)\cdots f(|c_k|)\,g(k), \qquad \textit{for } n\geq 1 \textit{ and } h(0)=1,$$

where $c_1,\dots,c_k$ are the cycles in permutation π and $|c|$ denotes the length of cycle c. Then

$$E_h(z) = E_g\Big(\sum_{n\geq 1} f(n)\frac{z^n}{n}\Big).$$

If $g=1$, then the above takes the form $\exp\Big(\sum_{n\geq 1} f(n)\frac{z^n}{n}\Big)$.

Proof. A set $\{c_{i_1},\dots,c_{i_k}\}$ gives rise to $(k-1)!$ cycles (because there are $k!$ permutations but each one of the k numbers can be the starting point). For a cycle c_i, let us denote by $\{c_i\}$ the unordered set of the elements of the cycle. Then we have

$$h(n) = \sum_k \sum_{\{c_1\},\dots,\{c_k\}} \big(|c_1|-1\big)!\,f(|c_1|)\cdots(|c_k|-1)!\,f(|c_k|)\,g(k),$$

and setting $\widetilde{f}(r) := (r-1)!f(r)$, the above formula is in the setting of Proposition 3.13 and, thus,

$$E_h(z) = E_g\Big(\sum_{n\geq 1}(n-1)!f(n)\frac{z^n}{n!}\Big) = E_g\Big(\sum_{n\geq 1} f(n)\frac{z^n}{n}\Big)$$

□

We are now ready to complete the proof of the second identity in (3.14):

Proof of (3.14). Consider

$$h(n) := \sum_k \sum_{\pi\in S_n\,:\,\pi=(c_1,\dots,c_k)} f(|c_1|)\cdots f(|c_k|) \qquad \text{for } n\geq 1 \text{ and } h(0)=1,$$

for a function f to be chosen. We can change variables in the sum on the right-hand side and instead of summing over permutations $\pi\in S_n$ and then over all possible cycles in this π, sum over all partitions $\lambda\vdash n$ and then over permutations $\pi\in S_n$ with cycle type λ. In other words,

$$\sum_k \sum_{\pi\in S_n\,:\,\pi=(c_1,\dots,c_k)} f(|c_1|)\cdots f(|c_k|) = \sum_{\lambda\vdash n} \sum_{\substack{\pi\in S_n \\ \pi \text{ has cycle type } \lambda=(\lambda_1,\lambda_2,\dots)}} \prod_i f(\lambda_i)$$

but by Proposition 2.5 and equation (2.2) we have that the number of permutations $\pi\in S_n$ with a given cycle type λ is the number of conjugacy classes of S_n, thus

$$\sum_{\substack{\pi\in S_n \\ \pi \text{ has cycle type } \lambda=(\lambda_1,\lambda_2,\dots)}} 1 = |K_\lambda| = \frac{n!}{z_\lambda},$$

with z_λ defined in (2.1) and so

$$\sum_{\lambda\vdash n} \sum_{\substack{\pi\in S_n \\ \pi \text{ has cycle type } \lambda=(\lambda_1,\lambda_2,\dots)}} \prod_i f(\lambda_i) = \sum_{\lambda\vdash n}\frac{n!}{z_\lambda}\prod_{i\geq 1} f(\lambda_i).$$

Given this, we can proceed by

$$\begin{aligned}\sum_{n\geq 1}\frac{h(n)}{n!} &= \sum_{n\geq 1}\frac{1}{n!}\sum_{\lambda\vdash n}\frac{n!}{z_\lambda}\prod_{i\geq 1} f(\lambda_i)\\ &= \sum_{n\geq 1}\sum_{\lambda_1+\lambda_2+\cdots=n}\frac{1}{z_\lambda}\prod_{i\geq 1} f(\lambda_i)\\ &= \sum_{\lambda}\frac{1}{z_\lambda}\prod_{i\geq 1} f(\lambda_i).\end{aligned}$$

If we now set $f(n) := p_n(x)p_n(y)$, then the above writes as

$$\sum_{\lambda}\frac{1}{z_\lambda}\prod_{i\geq 1} p_{\lambda_i}(x)p_{\lambda_i}(y) = \sum_{\lambda}\frac{1}{z_\lambda}p_\lambda(x)p_\lambda(y),$$

where the last equality comes from the definition of $p_\lambda(x) := \prod_{i\geq 1} p_{\lambda_i}(x)$. This completes the proof of (3.14). □

Exercise 11. *Complete the following alternative proof of the second relation of* (3.14)*: First, write*

$$\exp\Big(\sum_{n\geq 1}\frac{1}{n}p_n(x)p_n(y)\Big) = \prod_{n\geq 1}\exp\Big(\frac{1}{n}p_n(x)p_n(y)\Big),$$

and then use the Taylor expansion of the exponential: $e^x = \sum_{n\geq 0}\frac{x^n}{n!}$.

3.4. Scalar products and dualities to Cauchy type identities. The goal of this section is to show that Cauchy-type identities like that in (3.14) are *dual* to orthogonality of the symmetric functions involved in the Cauchy identity. This fact has been at the heart of the introduction of a fundamental class of symmetric functions, which are called *Macdonald functions* and which interpolate between several other important symmetric functions.

Let us start by defining an inner product on symmetric functions. An inner product is a bilinear form $\langle\cdot,\cdot\rangle$ and so to define an inner product it suffices to define it on the basis elements. So we define

Definition 3.15. *Define the inner product* $\langle\cdot,\cdot\rangle\colon \Lambda\times\Lambda\to\mathbb{R}$ *by the requirement that*

$$\langle m_\lambda, h_\mu\rangle = \delta_{\lambda\mu}, \qquad \text{for all partitions } \lambda,\mu,$$

where δ *is the Kronecker delta.*

If (u_i) *and* (v_j) *are two sets of bases for the ring of symmetric functions, we say that they are* **dual** *if* $\langle u_i, v_j\rangle = \delta_{ij}$.

Exercise 12. *Show that if* f,g *are homogenous symmetric functions then* $\langle f,g\rangle = 0$ *unless* f,g *have the same degree.*

Proposition 3.16. *The inner product defined above is symmetric, i.e.* $\langle f,g\rangle = \langle g,f\rangle$.

Proof. It suffices to check the symmetry when $f = h_\lambda$ and $g = h_\mu$ for λ,μ partitions (why ?). Towards this we will use Proposition 3.5:

$$\langle h_\lambda, h_\mu\rangle = \Big\langle\sum_\nu N_{\lambda\nu}m_\nu, h_\mu\Big\rangle = \sum_\nu N_{\lambda\nu}\big\langle m_\nu, h_\mu\big\rangle = N_{\lambda\mu},$$

by the definition of the inner product. The same computation shows that $\langle h_\mu, h_\lambda\rangle = N_{\mu\lambda}$ and the result follows from the symmetry of $N_{\lambda\mu}$ from Proposition 3.5. □

The main result of this section is the following:

Proposition 3.17. *Let $\{u_\lambda\}$ and $\{v_\lambda\}$ be bases of Λ such that for $\lambda \vdash n$, we have that $u_\lambda, v_\lambda \in \Lambda^n$, the space of symmetric functions in n variables. $\{u_\lambda\}, \{v_\lambda\}$ are dual bases, i.e. $\langle u_\lambda, v_\mu\rangle = \delta_{\lambda\mu}$ if and only if*

$$\sum_\lambda u_\lambda(x)v_\lambda(y) = \prod_{i,j}(1 - x_iy_j)^{-1} \tag{3.16}$$

Proof. $\{u_\lambda\}$ and $\{v_\lambda\}$ being bases means that we can decompose $m_\lambda = \sum_\varrho \zeta_{\lambda\varrho}u_\varrho$ and $h_\mu = \sum_\nu \eta_{\mu\nu}v_\nu$. Recalling the definition of the inner product we have

$$\delta_{\lambda\mu} = \langle m_\lambda, h_\mu\rangle = \sum_{\varrho,\nu} \zeta_{\lambda\varrho} \langle u_\varrho, v_\nu\rangle \eta_{\mu\nu}.$$

Considering the matrices $\zeta := (\zeta_{\lambda\varrho})$ and $\eta := (\eta_{\mu\nu})$ and $A := \big(\langle u_\varrho, v_\nu\rangle\big)_{\varrho\nu}$ we can write the above as

$$\mathrm{Id} = \zeta A\eta^{\mathrm{T}}.$$

$\{u_\lambda\}$ and $\{v_\lambda\}$ being dual bases means is equivalent to $A = \mathrm{Id}$, which combined with the identity $\mathrm{Id} = \zeta A\eta^{\mathrm{T}}$ implies that

$$\mathrm{Id} = \zeta\eta^{\mathrm{T}} \iff \mathrm{Id} = \zeta^{\mathrm{T}}\eta \iff \mathrm{Id} = \sum_\lambda \zeta_{\lambda\varrho}\eta_{\lambda\nu}. \tag{3.17}$$

Let's keep this in mind for the moment and make now use of the identity

$$\begin{aligned}
\prod_{i,j}(1 - x_iy_j)^{-1} &= \sum_\lambda m_\lambda(x)h_\mu(y)\\
&= \sum_\lambda \Big(\sum_\varrho \zeta_{\lambda\varrho}u_\varrho(x)\Big)\Big(\sum_\nu \eta_{\lambda\nu}v_\mu(y)\Big)\\
&= \sum_{\varrho,\nu} \Big(\sum_\lambda \zeta_{\lambda\varrho}\eta_{\lambda\nu}\Big)u_\varrho(x)v_\nu(y)\\
&= \sum_{\varrho,\nu} \delta_{\varrho\nu}\, u_\varrho(x)v_\nu(y)\\
&= \sum_\varrho u_\varrho(x)v_\varrho(y).
\end{aligned}$$

On the other hand, assume that

$$\prod_{i,j}(1 - x_iy_j)^{-1} = \sum_\varrho u_\varrho(x)v_\varrho(y),$$

and compare it with

$$\begin{aligned}
\prod_{i,j}(1 - x_iy_j)^{-1} &= \sum_\lambda m_\lambda(x)h_\mu(y)\\
&= \sum_\lambda \Big(\sum_\varrho \zeta_{\lambda\varrho}u_\varrho(x)\Big)\Big(\sum_\nu \eta_{\lambda\nu}v_\mu(y)\Big)\\
&= \sum_{\varrho,\nu} \Big(\sum_\lambda \zeta_{\lambda\varrho}\eta_{\lambda\nu}\Big)u_\varrho(x)v_\nu(y)
\end{aligned}$$

which imply by the fact that that $\{u_\lambda\}$ and $\{v_\lambda\}$ are bases that

$$\sum_\lambda \zeta_{\lambda\varrho}\eta_{\lambda\nu} = \delta_{\varrho\nu} \iff \zeta^{\mathrm{T}}\eta = \mathrm{Id} \iff \zeta\eta^{\mathrm{T}} = \mathrm{Id}$$

but since also $\mathrm{Id} = \zeta A\eta^{\mathrm{T}}$ and ζ, η are change of bases matrices, i.e. invertible, we have that $A = \mathrm{Id} \iff \langle u_\varrho, v_\nu\rangle = \delta_{\varrho\nu}$. □

3.5. Schur functions. We will mostly restrict attention when Schur functions involve only a finite number of indeterminates, in which case we talk about **Schur polynomials**. We present three different ways to define Schur polynomials, each one of which has its own benefits and uses in this text.

The first one, which was the original definition given by Schur, is via a determinant. In particular, for a partition $\lambda = (\lambda_1, ..., \lambda_n)$, with n parts,

$$s_\lambda(x_1, ..., x_n) := \frac{\det\left(x_i^{\lambda_j+n-j}\right)_{1\le i,j\le n}}{\det\left(x_i^{n-j}\right)_{1\le i,j\le n}}. \tag{3.18}$$

This determinantal expression is crucial in expressing the law of observables of integrable models in terms of determinants and enabling the asymptotic analysis.

The second definition is combinatorial. In this fashion, Schur functions appear as generating functions of **seminstandard Young tableaux**. Let us first define the notion of a seminstandard Young tableaux. These are Young diagrams filled with integer numbers $\{1, 2, 3, ...\}$ so the numbers in each row are *weakly* increasing along the rows and *strictly* increasing along columns. For example:

1	1	2	2	3
2	2	3		
3				

A Young tableau is called **standard** if it is filled with numbers $\{1, 2, 3, ...\}$ so that the entries in both rows *and* columns are *strictly* increasing. For example:

1	2	5	6	7
3	4	6		
4	5	7		
6				
7				

The partition λ that corresponds to the Young tableau (either standard or seminstandard) will be called **the shape of the tableau**. For example the shape of the semistandard Young tableau above is $\lambda = (5, 3, 1)$ and the shape of the above standard Young tableau is $\lambda = (5, 3, 3, 1, 1)$.

We are now ready to give the combinatorial definition of the Schur functions as:

$$s_\lambda(x_1, ..., x_n) := \sum_{\mathsf{T}:\, sh(\mathsf{T})=\lambda} x_1^{\sharp 1's} x_2^{\sharp 2's} \cdots x_n^{\sharp n's}, \tag{3.19}$$

where the sum is over all semistandard Young tableaux with shape λ and $\sharp 1's$ denotes the number of boxes in the tableau filled in with 1, $\sharp 2's$ denotes the number of boxes in the tableau filled in with 2 and so on.

The third definition is via an orthogonalisation procedure, which is of Gram-Schmidt type. This approach was generalised by Macdonald in [M88], in his definition of the *Macdonald polynomials* and it does not restrict to polynomials. To define the Schur polynomials in this way, uses the inner product we introduced in (3.15) Given this inner product, it can be shown, see [M98], that Schur functions are uniquely determined by their expansion in terms of the monomial symmetric functions:

$$\text{(A)} \qquad s_\lambda = m_\lambda + \sum_{\mu<\lambda} K_{\lambda\mu} m_\mu,$$

and their orthogonality with respect to the inner product as

$$\text{(B)} \qquad \langle s_\lambda, s_\mu \rangle = 0, \qquad \text{if} \qquad \lambda \neq \mu.$$

In (A) the sum is over all partitions μ such that $\mu < \lambda$, where the dominance ordering $<$ on partitions according to which $\mu < \lambda$, if $\mu_1 + \cdots + \mu_i < \lambda_1 + \cdots + \lambda_i$ for all $i \ge 1$.

The **Cauchy identity** for Schur polynomials (actually also true for Schur funcrions) reads as

$$\sum_\lambda s_\lambda(x_1, x_2, ...)\, s_\lambda(y_1, y_2, ...) = \prod_{i,j} \frac{1}{1 - x_i y_j}.$$

This identity, as well as identities of this type for other special functions, will play a very important role is the basis of defining probability measures on partitions and have played a central role in integrable probability. Unlike the previous Cauchy identities, this one is deeper and we will prove it via the Robinson-Schensted-Knuth correspondence.

Before going into the RSK correspondence let us make some remarks on Schur functions.

Remark 3.18. For a partition λ, we have that $s_\lambda(1,1,...,1) =: f^\lambda$ is the number of all *standard* (i.e. all entries are *strictly increasing* over rows and colums) Young tableaux of *shape* λ, i.e. the first row has length λ_1, the second row length λ_2 etc.

Exercise 13. *Let λ be a partition with length $\ell := \ell(\lambda)$, i.e. there are ℓ nonzero parts. Provide an expression for the number of n-step lattice paths from $0 \in \mathbb{Z}^\ell$ to $(\lambda_1, \lambda_2, ..., \lambda_n) \in \mathbb{Z}^\ell$ such that*

- *every step has unit length (and stays on the lattice)*
- *the walks stays in the domain $x_1 \geq x_2 \geq \cdots \geq x_\ell \geq 0$.*

Exercise 14. *Show that*

$$s_\lambda(1^m) := s_\lambda(\underbrace{1,1,...,1}_{m \text{ times}}) = \prod_{1\leq i<j\leq m} \frac{\lambda_i - \lambda_j + j - i}{j-i}.$$

Exercise 15. *Show the identity*

$$\sum_\lambda q^{\lambda_1+\lambda_2+\cdots} = \prod_{n\geq 1} \frac{1}{1-q^n}.$$

3.6. The Robinson-Schensted-Knuth (RSK) correspondence. The Robinson-Schensted (RS) correspondence is a bijection between matrices with nonnegative entries and a pair of semi-standard Young tableaux with the same shape. RSK is an extension of the Robinson-Schensted (RS) correspondence which is a bijection between permutations (or permutation matrices) and a pair of standard Young tableaux of the same shape.

Let us describe the algorithm, first in the permutation case.

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & N \\ x_1 & x_2 & \cdots & x_N \end{pmatrix},$$

where we denote $x_i := \sigma(i)$. Then,

- Starting from a pair of empty tableaux $(P_0, Q_0) = (\emptyset, \emptyset)$, assume that we have inserted the first i **biletters** $\binom{j}{x_j}$, for $1 \leq j \leq i \leq N$, of the permutation $\sigma \in S_N$ and we have obtained a pair of Young tableaux (P_i, Q_i).
- Next, we **(row) insert** the biletter $\binom{i+1}{x_{i+1}}$ as follows: If the number x_{i+1} is larger or equal † than all the numbers of the first row of P_i, then a box is appended at the end of the first row of P_i and its content is set to be x_{i+1}. This is then the tableau P_{i+1}. Also a box is appended at the end of the first row of Q_i and its content is set to be $i+1$, giving the tableau Q_{i+1}. If, on the other hand, there is a box in the first row of P_i with content strictly larger than x_{i+1}, then the content of the first such box becomes x_{i+1} and the replaced content, call it b, drops down and is *row inserted* in the second row of P_i following the same rules and creating (possibly) a cascade of dropdowns (called **bumps**). Eventually a box will be appended at the end of a row in P_i or below its last row, in which case it creates a new row, and the content of this box will be the last bumped letter. At the same, corresponding, location a box will be added at Q_i and its content will be set to be $i+1$.
- We repeat the above steps until all biletters have been row inserted.

†in the case of a permutation the "or equal" condition is void but it becomes relevant in the Robsinson-Schensted-Knuth generalisation.

Let us see how this algorithm works via an example. Consider the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 5 & 1 & 6 & 2 & 4 & 7 \end{pmatrix}$$

The sequence is as follows:

$$(\emptyset,\emptyset) \xrightarrow{3} \begin{array}{|c|}\hline 3\\ \hline\end{array} \quad \begin{array}{|c|}\hline 1\\ \hline\end{array} \xrightarrow{5} \begin{array}{|c|c|}\hline 3&5\\ \hline\end{array} \quad \begin{array}{|c|c|}\hline 1&2\\ \hline\end{array} \xrightarrow{1} \begin{array}{|c|c|}\hline 1&5\\ \hline 3\\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 1&2\\ \hline 3\\ \cline{1-1}\end{array} \xrightarrow{6} \begin{array}{|c|c|c|}\hline 1&5&6\\ \hline 3\\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|c|}\hline 1&2&4\\ \hline 3\\ \cline{1-1}\end{array} \xrightarrow{2}$$

$$\xrightarrow{2} \begin{array}{|c|c|c|}\hline 1&2&6\\ \hline 3&5\\ \cline{1-2}\end{array} \quad \begin{array}{|c|c|c|}\hline 1&2&4\\ \hline 3&5\\ \cline{1-2}\end{array} \xrightarrow{4} \begin{array}{|c|c|c|}\hline 1&2&4\\ \hline 3&5&6\\ \hline\end{array} \quad \begin{array}{|c|c|c|}\hline 1&2&4\\ \hline 3&5&6\\ \hline\end{array} \xrightarrow{7} \begin{array}{|c|c|c|c|}\hline 1&2&4&7\\ \hline 3&5&6\\ \cline{1-3}\end{array} \quad \begin{array}{|c|c|c|c|}\hline 1&2&4&7\\ \hline 3&5&6\\ \cline{1-3}\end{array}.$$

In words, we have that we start by row inserting '3' and creating a box with content '3', identified with tableau P_1, and a box with content '1' constructing tableau Q_1. Then '5' is row inserted in P_1 and since it is larger than '3' it bypasses the box with content '3' and sits in a new box in the right of '3', creating P_2. A box with content '2' is also created in the right of the box with content '1' in Q_1 creating tableau Q_2. Then '1' is row inserted to P_2 and being smaller than '3' it bumps '3' and sits in the first box of P_2. '3' is then row inserted in the second row and since this is empty, it creates a new box whose content becomes '3'. At the same time a new box in the second row of Q_2 is created whose content is '3', giving Q_3. The procedure continues in this way.

There are a few observations to be made from this example.

- Tableaux P and Q have the same shape, i.e. the lengths of the successive rows in each tableau are equal.

This is a general fact. The tableaux that RS produces have the same shape. This can be easily seen as at any stage of the algorithm a box is created at the same location in both the P and Q tableau.

- Tableaux P and Q are actually equal.

This is not a general fact but a consequence of the fact that the permutation matrix associated to the above permutation, is symmetric, or that $\sigma = \sigma^{-1}$. In general, as we will state below, if (P, Q) is the output of a permutation σ, then the output of permutation σ^{-1} is (Q, P). Thus, if $\sigma = \sigma^{-1}$, then $P = Q$.

- A third observation that we make is that the length of the first row of either output tableau P and Q (which in this case is 4) equals the length of the longest increasing subsequence in the permutation $(3, 5, 1, 6, 2, 4, 7)$, which, for example (as there are more than one such), is the sequence $(3, 5, 6, 7)$. Moreover, the length of the second longest increasing subsequence $(1, 2, 4)$ equals the length of the second row of the output tableaux.

This is also not a coincidence and goes by the name of Greene's theorem (see Theorem 3.20 below), an extension of Schensted's theorem (see Theorem 3.19 below):

Theorem 3.19 (Schensted). *The* RS *correspondence is a bijection between permutations and pairs of standard Young tableaux* (P, Q) *of the same shape. If* $\sigma \in S_N$ *and* $(P, Q) = \mathsf{RS}(\sigma)$ *is the image of* σ *under RS-correspondence, then* $(Q, P) = \mathsf{RS}(\sigma^{-1})$*, where* σ^{-1} *is the inverse of permutation* σ*. In particular, if* $\sigma = \sigma^{-1}$*, then* $P = Q$*.*

Theorem 3.20 (Greene). *Let* $\sigma \in S_N$ *and* $(P, Q) = \mathsf{RS}(\sigma)$*. Then, the length* λ_1 *of the first row of the output tableaux* P *or* Q *equals the length of the longest increasing subsequence in* σ*. Moreover, the sum* $\lambda_1 + \lambda_2 + \cdots + \lambda_r$ *of the lengths of the first* r *rows equals the maximum possible length of dijoint unions of* r *increasing subsequences in* σ*.*

RSK **correspondence.** Bearing in mind that permutations are identified with matrices whose entries are either 0 or 1 and no two 1's are in the same row or column (permutation matrices), one can see the RS correspondence as a bijection between permutation matrices and pairs of standard Young tableaux. Knuth's generalisation constituted in extending RS as a bijection between matrices with nonnegative integer entries and pairs of *seminstandard* Young tableaux. We can think of a matrix $W = (w^i_j)_{\substack{1\le i\le n\\ 1\le j\le N}}$, where i indicates rows and j columns, as a sequence of n words

$$w^i := 1^{w_1^i} \cdots N^{w_N^i} := \underbrace{1 \cdots 1}_{w_1^i} \ \underbrace{2 \cdots 2}_{w_2^i} \cdots \underbrace{N \cdots N}_{w_N^i} \tag{3.20}$$

with letters $1, 2, ..., N$, such that w_j^i symbolises the number of letters j in word i. Knuth's extension of RS correspondence, named Robinson-Schensted-Knuth (RSK), consists of inserting, via the RS row-insertion, the letters of words $w^1, w^2, ..., w^n$ (in this order) with the letters of each word w^i as in (3.20) being read from left to right.

For a matrix W we will denote by $\mathsf{RSK}(W) = (P, Q)$ the output of the RSK correspondence. Later on we will transcribe the RSK correspondence in a matrix formulation inspired by studies in integrable systems and cluster algebras [NY04].

But let us finally move on to probabilistic aspects...

3.7. Gelfand-Tsetlin parametrisation. Gelfand-Tsetlin (GT) patterns are triangular arrays of numbers $(z_j^i)_{1 \le j \le i \le N}$ which interlace, meaning that

$$z_{j+1}^{i+1} \le z_j^i \le z_j^{i+1}, \tag{3.21}$$

and for this reason they are depicted as

$$\begin{array}{ccccccccc} & & & & z_1^1 & & & & \\ & & & z_2^2 & & z_1^2 & & & \\ & & z_3^3 & & z_2^3 & & z_1^3 & & \\ & \iddots & \iddots & & & & \ddots & \ddots & \\ z_N^N & z_{N-1}^N & & & \cdots & & & z_2^N & z_1^N \end{array} . \tag{3.22}$$

They provide a particularly useful parametrisation of Young tableaux: given a Young tableau consisting of letters $1, 2, ..., N$ (not all of which have to appear in the tableau) the Gelfand-Tsetlin variables z_j^i are defined as

$$z_j^i := \sum_{k=j}^{i} \sharp\{ \ k\text{'s in the } j^{th} \text{ row}\} \tag{3.23}$$

Given this definition, the right inequality in (6.4) is immediate, while the left one is a consequence of the fact that entries along columns in a Young tableau are strictly increasing.

The bottom row of a GT pattern is called the **shape**, since z_i^N equals the length of the i-th row of the corresponding tableau and thus the collection of $z_1^N, z_2^N, ...$ determines the shape of the tableau. We will denote the shape of a GT pattern Z by $sh(\mathsf{Z})$ and similarly the shape of a tableau P by $sh(P)$. We will also often identify a GT pattern Z with the corresponding Young tableau P.

We will see they provide a structure, which couples models in the KPZ class, e.g. longest increasing subsequence or last passage percolation, with Random Matrices. To give a preliminary idea of how this comes about, we point out that, as a consequence of Schensted's Theorem 3.19, entry z_1^N of a Gelfand-Tsetlin pattern is equal to the length of the first row of a Young Tableau. On the other hand, it turns out that in certain situations *random* Gelfand-Tsetlin patterns have a bottom row with law identical to the law of the eigenvalues of certain random matrices. Therefore, the element z_1^N has a dual nature: it is an observable of models within the KPZ class and at the same time its distribution is identical (or closely related) to the distribution of the largest eigenvalue of certain random matrices. This coupling has played a central role in formulating the integrable structure of models in the KPZ universality and we will explore this in the coming sections.

An important *invariant* of RSK is the **type** of a tableau P, denoted by $type(P)$. In GT parametrisation, this is defined to be the vector

$$\big(|z^i| - |z^{i-1}| : \ i = 1, ..., N\big), \quad \text{with} \quad |z^i| := \sum_{j=1}^{i} z_j^i,$$

and the convention that $|z^0| = 0$. Considering a pair of Gelfand-Tsetlin patterns $(\mathsf{Z}, \mathsf{Z}')$ as the output of RSK with input matrix $W = (w^i_j \colon 1 \le i \le n,\ 1 \le j \le N)$, that is $(\mathsf{Z}, \mathsf{Z}') = \mathsf{RSK}(W)$, with Z corresponding to the P tableau and Z' to the Q tableau in the RSK correspondence, then it holds that

$$|z^k| - |z^{k-1}| = \sum_{i=1}^{n} w^i_k. \tag{3.24}$$

This is due to the fact that both sides represent the number of letters k inserted from W via RSK. This is clear for the right-hand side, since (by definition) w^i_j is considered as the number of letters j in word i, while for the left-hand side this follows from

$$\begin{aligned} |z^k| - |z^{k-1}| &= \sum_{j=1}^{k} z^k_j - \sum_{j=1}^{k-1} z^{k-1}_j = z^k_k + \sum_{j=1}^{k-1} (z^k_j - z^{k-1}_j) \\ &= \sharp\{k\text{'s in word } k\} + \sum_{j=1}^{k-1} \sharp\{k\text{'s in word } j\}. \end{aligned}$$

4. A solvable Last Passage Percolation model

Let us skip to the study of a solvable probability model. This is the last passage percolation with geometric variables In this section we will show how RSK allows to write explicitly the distribution of the last passage percolation time in terms of Schur functions. The determinantal expression of the Schur functions will then set the stage for asymptotic analysis, which we will present in following sections.

To set things up, we consider a matrix $\mathsf{W} = (\boldsymbol{w}^i_j)_{\substack{1 \le i \le m,\\ 1 \le j \le n}}$, where we assume that the entries are independent random variables with geometric distribution

$$\mathbb{P}(\boldsymbol{w}^i_j = w^i_j) = (1 - p_i q_j)(p_i q_j)^{w^i_j} \mathbb{1}_{w^i_j \in \{0,1,2,\dots\}}, \qquad 1 \le i \le m\,,\ 1 \le j \le n, \tag{4.1}$$

where p_i, q_j are parameters in $(0, 1)$. The first question we want to ask is whether we can compute the law of

$$\tau_{m,n} := \max_{\pi \in \boldsymbol{\Pi}_{m,n}} \sum_{(i,j) \in \pi} \boldsymbol{w}^i_j, \tag{4.2}$$

where $\boldsymbol{\Pi}_{m,n}$ is the set of down-right paths going from site $(1, 1)$ to site $(m, n) \in \mathbb{N}^2$ (using the matrix rather than the cartesian index notation). For simplicity, let us assume that $m = n = N$, although the general case can also be treated following similar reasoning. For conciseness we will also denote $\tau_{N,N}$ by τ_N. The answer to this question is affirmative and the reason is that the geometric distribution fits the framework and the properties of RSK. In particular, we can answer the posed question by carrying out the following steps:

Step 1. Push forward law. The law of the random weight matrix W from (4.1) can be written explicitly in terms of GT variables as

$$\begin{aligned} \mathbb{P}(\mathsf{W} = \{w^i_j\}) &= \prod_{i,j} (1 - p_i q_j) \prod_i p_i^{\sum_j w^i_j} \prod_j q_j^{\sum_i w^i_j} \\ &= \prod_{i,j} (1 - p_i q_j) \prod_i p_i^{|(z^i)'| - |(z^{i-1})'|} \prod_j q_j^{|z^j| - |z^{j-1}|}. \end{aligned} \tag{4.3}$$

Here we related $\sum_i w^i_j$ to the *type* of the P-tableau, $\big(|z^j| - |z^{j-1}|\big)_{j=1,\dots,N}$ as $\sum_i w^i_j = |z^j| - |z^{j-1}|$ and the sum $\sum_j w^i_j$ to the *type* of the Q-tableau $|(z^i)'| - |(z^{i-1})'|$ as $\sum_j w^i_j = |(z^i)'| - |(z^{i-1})'|$. The former is just (3.24), while the latter follows from the fact that

$$\text{if } \mathsf{RSK}(\mathsf{W}) = (P, Q) = (\mathsf{Z}, \mathsf{Z}'), \text{ then } \mathsf{RSK}(\mathsf{W}^{\mathsf{t}}) = (Q, P) = (\mathsf{Z}', \mathsf{Z}).$$

Step 2. Marginalisation and determinantal measures. We are now ready to compute $\mathbb{P}(\tau_N \le u)$. Using the previous two steps we have that

$$\mathbb{P}(\tau_N \le u) = \sum_{\lambda:\, \lambda_1 \le u} \;\; \sum_{\substack{(Z,Z')\text{ pair of GT patterns}\\ \text{with shape } \lambda}} \mathbb{P}\big(\mathsf{RSK}(\mathsf{W}) = (Z, Z')\big)$$

and by (4.3) this equals

$$\begin{aligned}
&\prod_{i,j}(1-p_iq_j) \sum_{\lambda:\, \lambda_1 \le u} \;\; \sum_{\substack{(Z,Z')\text{ pair of GT patterns}\\ \text{with shape } \lambda}} \prod_i p_i^{|(z^i)'|-|(z^{i-1})'|} \prod_j q_j^{|z^j|-|z^{j-1}|} \\
&= \prod_{i,j}(1-p_iq_j) \sum_{\lambda:\, \lambda_1 \le u} \;\; \sum_{\substack{Z:\text{ GT pattern}\\ \text{with shape } \lambda}} \prod_j q_j^{|z^j|-|z^{j-1}|} \sum_{\substack{Z':\text{ GT pattern}\\ \text{with shape } \lambda}} \prod_i p_i^{|(z^i)'|-|(z^{i-1})'|},
\end{aligned}$$

and now each of the two rightmost summands are recognised to be the Schur functions, whose expression as generating series of Young tableaux (3.19) may be rewritten in the Gelfand-Tsetlin notation as

$$s_\lambda(q) := \sum_{\substack{Z:\text{ GT pattern}\\ \text{with shape } \lambda}} \prod_j q_j^{|z^j|-|z^{j-1}|}. \tag{4.4}$$

Thus, the above induces that (this step is really a change of notation)

$$\mathbb{P}(\tau_N \le u) = \prod_{i,j}(1-p_iq_j) \sum_{\lambda:\, \lambda_1 \le u} s_\lambda(q)\, s_\lambda(p). \tag{4.5}$$

We have, thus, computed the law of last passage percolation in terms of special functions, which furthermore possess many nice properties. In particular, they can be written in terms of determinants and in fact there are more than one such formulae. For example, if $\lambda = (\lambda_1, \lambda_2, ...)$ is a partition and $p_1, p_2, ..., p_N$ are nonnegative parameters (or variables), then

$$s_\lambda(p) = \frac{\det\big(p_i^{\lambda_j+N-j}\big)_{1\le i,j\le N}}{\det\big(p_i^{N-j}\big)_{1\le i,j\le N}}, \tag{4.6}$$

where in the denominator one recognises the Vandermonde determinant, which can be computed as $\Delta_N(p) := \prod_{1\le i<j\le N}(p_i - p_j)$.

The next question we want to ask is whether we can perform asymptotic analysis. For this, we have

Step 3. Fredholm determinants. Relation (4.6) allows to express (4.5) as a *Fredholm determinant*, in a form that is suitable to take the asymptotic limit and prove convergence to Tracy-Widom GUE distribution. We will introduce the notion of a Fredholm determinant and some of its properties in the next section. The significance of expressing (4.5) and other such probabilities in terms of Fredholm determinants is that doing so facilitates taking the limit of N tending to infinity. In (4.5), N is the number of varables $\lambda_1, ..., \lambda_N$, over which the sum in (4.5) is taken. Thus, taking the limit $N \to \infty$ corresponds to the number of summations to infinity and the meaning of such limit is not clear at all at this stage. The key to resolving this difficulty is the notion of Fredholm determinants, which re-expresses such sums and integrals in a way that the limit in N becomes unambiguous and tractable. We will see how this is done in the next section.

We are now going to introduce the notion of determinantal measures and Fredholm determinants. We will demonstrate how two basic tools from determinantal calculus, the Cauchy-Binet or Andreief's identity and the Sylvester's identity, can be used to turn a determinantal measure into a Fredholm determinant.

In many statistical models we encounter probability measures of the form

$$\mu_N(f) := \frac{Z_N(f)}{Z_N}, \tag{4.7}$$

where $\mu_N(f)$ denotes expectation of a functional $f \in L^2(\mathcal{X}, \mu)$ on a measure space $(\mathcal{X}, \mu)$ and

$$Z_N(f) := \int_{\mathcal{X}^N} \det\big(\phi_i(x_j)\big)_{1\le i,j\le N} \det\big(\psi_i(x_j)\big)_{1\le i,j\le N}\; f(x_1)\cdots f(x_N)\; \mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_N) \tag{4.8}$$

the quantity $Z_N = Z_N(1)$ is typically known as the *partition function*. Measures with determinants in this form in the right-hand side are known as *determinantal measures*. We note that the functions ϕ_i, ψ_i can be either non-negative, in which case, we have a genuine probability measure, or they may also be allowed to take negative values, in which case we deal with signed measures. For more regarding determinantal measures and processes we refer to [B11, J05].

Due to (4.6) we see that the Schur measure

$$\mathbb{P}(\lambda) := \prod_{i,j}(1 - p_i q_j)\, s_\lambda(q)\, s_\lambda(p). \tag{4.9}$$

(see (4.5)) on partitions λ is a determinantal measure with $\phi_i(\lambda_j) := p_i^{\lambda_j+N-j}$ and $\psi_i(\lambda_j) := q_i^{\lambda_j+N-j}$ This measure was introduced by Okounkov [O01].

Determinantal probabilites such as (4.5) can be written in terms of objects called *Fredholm determinants* and this is crucial in obtaining asymptotics. We will exhibit this in the example of the Schur measure and last passage percolation with geometric weights.

Let us first define the notion of **Fredholm determinant**. Given an integral operator K acting on $L^2(\mathcal{X}, \mu)$ of a general measure space $(\mathcal{X}, \mu)$ by

$$Kf(x) = \int_{\mathcal{X}} K(x,y) f(y)\, \mu(\mathrm{d}y),$$

we define the Fredholm determinant associated to K by

$$\det(I + K)_{L^2(\mathcal{X},\mu)} := 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{\mathcal{X}^n} \det\big(K(x_i, x_j)\big)_{n\times n}\, \mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_n). \tag{4.10}$$

Here I is the identity map. Of course, one has to make sure that this infinite series is convergent. This is usually guaranteed by requiring that K is a *trace class operator*. We refer to [S79] for more details, but let us go through a quick sketch: For a compact operator K on a Hilbert space, say $L^2(\mathcal{X}, \mu)$, we define its trace class norm as $\|K\|_1 := \mathrm{Tr}\sqrt{K^*K}$, where K^* is the adjoint of K and the square root can be defined via operator calculus, since K^*K is self-adjoint. In the case of a trace class norm operator one can obtain that (4.10) is well defined and the Fredholm determinant is bounded by

$$|\det(I + K)_{L^2(\mathcal{X},\mu)}| \le e^{\|K\|_1}. \tag{4.11}$$

Exercise 16. *Prove inequality* (4.11).

Moreover, one has the following continuity result

Exercise 17.

$$|\det(I + K_1)_{L^2(\mathcal{X},\mu)} - \det(I + K_2)_{L^2(\mathcal{X},\mu)}| \le \|K_1 - K_2\|_1\, e^{\|K\|_1 + \|K\|_2 + 1}. \tag{4.12}$$

A consequence of this inequality is that if we would like to establish convergence of certain Fredholm determinants, it is enough to establish the convergence of the corresponding operators in the trace class norm.

A way to get a feeling about definition (4.10) is to consider the case where K is an $N \times N$ matrix and let $\lambda_1, ..., \lambda_N$ denote its eigenvalues. Then

$$\det(I + K) = \prod_{i=1}^{N}(1 + \lambda_i) = 1 + \sum_{m=1}^{N} \sum_{1\le i_1 < \cdots < i_m \le N} \lambda_{i_1}\cdots\lambda_{i_m}. \tag{4.13}$$

From the standard property of trace, $\sum_{i=1}^N \lambda_i = \mathrm{Tr} K = \sum_x K(x,x)$, one sees immediately the identification of the first non-trivial terms in (4.10) and (4.13). The rest of the terms have similar interpretation as traces of tensor products of K, see [S79] for details. Without getting into details, we mention that Fredholm determinants formalise in some sense the inclusion-exclusion principle.

We will now state two important tools that will allow to relate determinantal measures to Fredholm determinants.

Proposition 4.1 (Cauchy-Binet or Andreief identity). *Consider a collection of functions* $\big(\phi_i(\cdot)\big)_{1\le i\le N}$ *and* $\big(\psi_i(\cdot)\big)_{1\le i\le N}$*, which belong to* $L^2(\mathcal{X},\mu)$ *of a measure space* $(\mathcal{X},\mu)$*. Then*

$$\begin{aligned}\frac{1}{N!}\int_{\mathcal{X}^N} &\det\big(\phi_i(x_j)\big)_{1\le i,j\le N}\det\big(\psi_i(x_j)\big)_{1\le i,j\le N}\;\mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_N)\\ &=\det\Big(\int_{\mathcal{X}}\phi_i(x)\psi_j(x)\,\mu(\mathrm{d}x)\Big)_{1\le i,j\le N}.\end{aligned}$$

Proof. Expand the determinants and swap integration and sums to get:

$$\begin{aligned}&\frac{1}{N!}\int_{\mathcal{X}^N}\det\big(\phi_i(x_j)\big)_{1\le i,j\le N}\det\big(\psi_i(x_j)\big)_{1\le i,j\le N}\;\mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_N)\\ &=\frac{1}{N!}\int_{\mathcal{X}^N}\Big(\sum_{\sigma\in S_n}sgn(\sigma)\prod_{i=1}^N\phi_i(x_{\sigma(i)})\Big)\Big(\sum_{\tau\in S_n}sgn(\tau)\prod_{j=1}^N\psi_j(x_{\tau(j)})\Big)\prod_{i=1}^N\mu(\mathrm{d}x_i)\\ &=\frac{1}{N!}\sum_{\sigma,\tau\in S_n}sgn(\sigma)sgn(\tau)\int_{\mathcal{X}^N}\prod_{i=1}^N\phi_i(x_{\sigma(i)})\prod_{j=1}^N\psi_j(x_{\tau(j)})\prod_{i=1}^N\mu(\mathrm{d}x_i).\end{aligned}$$

We now want to group the terms correspoding to the same spatial variable x_i. For this, we order the terms in the second product as

$$\prod_{i=1}^N\psi_{\tau^{-1}\sigma(i)}(x_{\sigma(i)}),$$

by the change of variables $j:=\tau^{-1}\sigma(i)$. Inserting this into the integral and grouping the $\mathrm{d}x_{\sigma(i)}$ integration we have that the above integral is

$$\begin{aligned}&\frac{1}{N!}\sum_{\sigma,\tau\in S_n}sgn(\sigma)sgn(\tau)\prod_{i=1}^N\int_{\mathcal{X}}\phi_i(x)\psi_{\tau^{-1}\sigma(i)}(x)\,\mu(\mathrm{d}x)\\ &=\frac{1}{N!}\sum_{\sigma,\tau\in S_n}sgn(\tau^{-1}\sigma)\prod_{i=1}^N\int_{\mathcal{X}}\phi_i(x)\psi_{\tau^{-1}\sigma(i)}(x)\,\mu(\mathrm{d}x)\end{aligned}$$

where in the second equality we used the character property of the sign of a permutation. We now change summation variables: $(\tau,\tau^{-1}\sigma)=(\tau,\eta)$ and write the above as

$$\begin{aligned}&\frac{1}{N!}\sum_{\eta,\tau\in S_n}sgn(\eta)\prod_{i=1}^N\int_{\mathcal{X}}\phi_i(x)\psi_{\eta(i)}(x)\,\mu(\mathrm{d}x)\\ &=\frac{1}{N!}\sum_{\tau\in S_n}\Big(\sum_{\eta}sgn(\eta)\prod_{i=1}^N\int_{\mathcal{X}}\phi_i(x)\psi_{\eta(i)}(x)\,\mu(\mathrm{d}x)\Big)\\ &=\det\Big(\int_{\mathcal{X}}\phi_i(x)\psi_j(x)\,\mu(\mathrm{d}x)\Big)_{1\le i,j\le N}\end{aligned}$$

□

Proposition 4.2. *Consider general measure spaces* $(\mathcal{X},\mu),(\mathcal{Y},\nu)$ *and trace class operators* $A\colon L^2(\mathcal{Y},\nu)\to L^2(\mathcal{X},\mu)$ *and* $B\colon L^2(\mathcal{X},\mu)\to L^2(\mathcal{Y},\nu)$*. Then*

$$\det\big(I+AB\big)_{L^2(\mathcal{X},\mu)}=\det\big(I+BA\big)_{L^2(\mathcal{Y},\nu)}.$$

Proof. Use the definition of the Fredholm determinant we have

$$\begin{aligned}
\det\big(I+AB\big)_{L^2(\mathcal{X},\mu)} &= 1+\sum_{n=1}^{\infty}\frac{1}{n!}\int_{\mathcal{X}^n}\det\big(AB(x_i,x_j)\big)_{n\times n}\mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_n)\\
&= 1+\sum_{n=1}^{\infty}\frac{1}{n!}\int_{\mathcal{X}^n}\sum_{\sigma\in S_n} sgn(\sigma)\prod_{i=1}^n AB(x_i,x_{\sigma(i)})\prod_{i=1}^n\mu(\mathrm{d}x_i)\\
&= 1+\sum_{n=1}^{\infty}\frac{1}{n!}\int_{\mathcal{X}^n}\sum_{\sigma\in S_n} sgn(\sigma)\prod_{i=1}^n\Big(\int_{\mathcal{Y}}A(x_i,y)B(y,x_{\sigma(i)})\nu(\mathrm{d}y)\Big)\prod_{i=1}^n\mu(\mathrm{d}x_i)\\
&= 1+\sum_{n=1}^{\infty}\frac{1}{n!}\int_{\mathcal{X}^n}\sum_{\sigma\in S_n} sgn(\sigma)\int_{\mathcal{Y}^n}\prod_{i=1}^n A(x_i,y_i)B(y_i,x_{\sigma(i)})\prod_{i=1}^n\nu(\mathrm{d}y_i)\prod_{i=1}^n\mu(\mathrm{d}x_i)\\
&= 1+\sum_{n=1}^{\infty}\frac{1}{n!}\int_{\mathcal{X}^n}\sum_{\sigma\in S_n} sgn(\sigma)\int_{\mathcal{Y}^n}\prod_{i=1}^n A(x_i,y_i)\prod_{i=1}^n B(y_i,x_{\sigma(i)})\prod_{i=1}^n\nu(\mathrm{d}y_i)\prod_{i=1}^n\mu(\mathrm{d}x_i)\\
&= 1+\sum_{n=1}^{\infty}\frac{1}{n!}\int_{\mathcal{X}^n}\sum_{\sigma\in S_n} sgn(\sigma)\int_{\mathcal{Y}^n}\prod_{i=1}^n A(x_i,y_i)\prod_{i=1}^n B(y_{\sigma^{-1}(i)},x_i)\prod_{i=1}^n\nu(\mathrm{d}y_i)\prod_{i=1}^n\mu(\mathrm{d}x_i)
\end{aligned}$$

where in the last equality we just rearranged the product $\prod_{i=1}^n B(y_i,x_{\sigma(i)})$ to $\prod_{i=1}^n B(y_{\sigma^{-1}(i)},x_i)$. Grouping now the terms with x_i together and interchanging the ν and μ integrations, we have that the above is equal to

$$\begin{aligned}
&1+\sum_{n=1}^{\infty}\frac{1}{n!}\int_{\mathcal{X}^n}\sum_{\sigma\in S_n} sgn(\sigma)\int_{\mathcal{Y}^n}\prod_{i=1}^n B(y_{\sigma^{-1}(i)},x_i)A(x_i,y_i)\prod_{i=1}^n\nu(\mathrm{d}y_i)\prod_{i=1}^n\mu(\mathrm{d}x_i)\\
&= 1+\sum_{n=1}^{\infty}\frac{1}{n!}\int_{\mathcal{Y}^n}\sum_{\sigma\in S_n} sgn(\sigma)\prod_{i=1}^n\Big(\int_{\mathcal{X}}B(y_{\sigma^{-1}(i)},x)A(x,y_i)\prod_{i=1}^n\mu(\mathrm{d}x)\Big)\prod_{i=1}^n\nu(\mathrm{d}y_i)\\
&= 1+\sum_{n=1}^{\infty}\frac{1}{n!}\int_{\mathcal{Y}^n}\sum_{\sigma^{-1}\in S_n} sgn(\sigma^{-1})\prod_{i=1}^n\Big(\int_{\mathcal{X}}B(y_{\sigma^{-1}(i)},x)A(x,y_i)\prod_{i=1}^n\mu(\mathrm{d}x)\Big)\prod_{i=1}^n\nu(\mathrm{d}y_i)\\
&= 1+\sum_{n=1}^{\infty}\frac{1}{n!}\int_{\mathcal{Y}^n}\det\Big(BA(y_i,y_j)\Big)_{1\le i,j\le n}\prod_{i=1}^n\nu(\mathrm{d}y_i)\\
&= \det\big(I+BA\big)_{L^2(\mathcal{Y},\nu)}
\end{aligned}$$

□

The next proposition shows how to express marginals of a determinantal measure to a Fredholm determinant. This tranformation is crucial in order to do large N asymptotics by transforming them into asymtotics of integral operators rather than computing asymptotics of determinants whose size grows to infinity.

Proposition 4.3. *Consider a collection of functions* $\big(\phi_i(\cdot)\big)_{1\le i\le N}$ *and* $\big(\psi_i(\cdot)\big)_{1\le i\le N}$, *which belong to* $L^2(\mathcal{X},\mu)$ *of a measure space* $(\mathcal{X},\mu)$. *Define the matrix*

$$\mathsf{G}_{ij} := \int_{\mathcal{X}}\phi_i(x)\psi_j(x)\,\mu(\mathrm{d}x) \tag{4.14}$$

and assume that it is invertible. Define also the operator K *with kernel*

$$K(x,y) := \sum_{i,j}\psi_i(x)\big(\mathsf{G}^{-1}\big)_{ij}\,\phi_j(y). \tag{4.15}$$

Then, following notation (4.8), *it holds that, for a general bounded function* g *on* $\mathcal{X}$,

$$\frac{Z_N(1+g)}{Z_N} = \det\big(I+gK\big)_{L^2(\mathcal{X},\mu)}.$$

Proof. The proof is a consequence of Propositions 4.1 and 4.2. Let $f = 1 + g$. By the Cauchy-Binet identity we have that

$$\begin{aligned} Z_N(f) &= \det\Big(\int_{\mathcal{X}} \phi_i(x)\psi_j(x)\, f(x)\,\mathrm{d}\mu\Big)_{1\le i,j\le N} \\ &= \det\Big(\int_{\mathcal{X}} \phi_i(x)\psi_j(x) + \int_{\mathcal{X}} \phi_i(x)\psi_j(x)g(x)\,\mathrm{d}\mu\Big)_{1\le i,j\le N} \\ &= \det\Big(\mathsf{G}_{ij} + \int_{\mathcal{X}} \phi_i(x)\psi_j(x)g(x)\,\mathrm{d}\mu\Big)_{1\le i,j\le N}, \end{aligned}$$

where by setting $g = 0$ or equivalently $f = 1$ we see that $Z_N = \det \mathsf{G}$. Using the multiplicativity of determinants, ie $\det(AB) = \det(A)\det(B)$ and denoting by $(\mathsf{G}^{-1})_{ij}$ the (i, j) entry of matrix G^{-1}, we see that

$$\begin{aligned} \frac{Z_N(1+g)}{Z_N} &= \det\Big(\delta_{ij} + \sum_k (\mathsf{G}^{-1})_{ik} \int_{\mathcal{X}} \phi_k(x)\psi_j(x)g(x)\,\mathrm{d}\mu\Big)_{1\le i,j\le N} \\ &= \det\Big(\delta_{ij} + \int_{\mathcal{X}} g(x) \sum_k (\mathsf{G}^{-1})_{ik}\, \phi_k(x)\psi_j(x)\,\mathrm{d}\mu\Big)_{1\le i,j\le N}. \end{aligned} \tag{4.16}$$

Now, we will use Proposition 4.2 with

$$A\colon L^2(\mathcal{X}, \mu) \to \ell^2(\{1, ..., N\}) \quad \text{with} \quad A(i; x) := \sum_k (\mathsf{G}^{-1})_{ik}\, \phi_k(x)$$

acting on functions in $f \in L^2(\mathcal{X}, \mu)$ as $(Af)(i) := \int_{\mathcal{X}} A(i; x)f(x)\mu(\mathrm{d}x)$, with the result being an element of $\ell^2(\{1, ..., N\}$, and

$$B\colon \ell^2(\{1, ..., N\}) \to L^2(\mathcal{X}, \mu) \quad \text{with} \quad B(x; i) := g(x)\psi_i(x),$$

acting on elements $h \in \ell^2(\{1, ..., N\}$ as $(Bh)(x) := g(x) \sum_{i=1}^N \psi_i(x)\, h(i)$, with the result being an element of $L^2(\mathcal{X}, \mu)$. Then, by Proposition 4.2, (4.16) may be written as

$$\frac{Z_N(1+g)}{Z_N} = \det\Big(I + AB\Big)_{\ell^2(\{1,...,N\})} = \det\Big(I + BA\Big)_{L^2(\mathcal{X},\mu)}$$

where the kernel of operator BA may be written explicitly as

$$BA(x, y) = \sum_{\ell=1}^N B(x; \ell)A(\ell, y) = g(x) \sum_{1\le \ell,k\le N} \psi_\ell(x)\ (\mathsf{G}^{-1})_{\ell,k}\ \phi_k(y)$$

for $x, y \in \mathcal{X}$, completing the proof in view of (4.15). □

The difficulty that arises when one would like to apply concretely Proposition 4.3 is to actually invert the matrix G. In some situations this can be done, as we will now see by applying this to last passage percolation with geometric weights. We have

Proposition 4.4. *Consider a matrix* $\mathsf{W} = (\boldsymbol{w}^i_j)_{1\le i,j\le N}$ *with distribution* $\mathbb{P}$ *as in* (4.1) *and* $\tau_N = \max_{\pi\in\boldsymbol{\Pi}_{N,N}} \sum_{(i,j)\in\pi} \boldsymbol{w}^i_j$. *Then*

$$\mathbb{P}\big(\tau_N \le x\big) = \det\big(I + g_N K^{LPP}_N\big)_{\ell^2(\mathbb{N})}, \tag{4.17}$$

with $g_N = \mathbb{1}_{[x+N,\infty)}$ *and for* $t, s \in \mathbb{N}$ *the kernel of operator* K^{LPP}_N *is given by*

$$K^{LPP}_N(t, s) = \frac{1}{(2\pi\iota)^2} \int_{\gamma_1} \mathrm{d}\zeta \int_{\gamma_2} \mathrm{d}\eta\, \frac{\eta^s\zeta^t}{1 - \zeta\eta} \prod_{j=1}^N \Big(\frac{1 - \eta q_j}{\eta - p_j}\Big) \prod_{i=1}^N \Big(\frac{1 - p_i\zeta}{\zeta - q_i}\Big) \tag{4.18}$$

where γ_2 *is the circle in the complex plane with counter-clockwise orientation, centred at zero and radius one and* γ_1 *is the circle with counter-clockwise orientation, centred at zero of radius* $r < 1$. *Without loss of generality, we assume that all* $p_i, q_i, 1 \le i \le N$ *are small enough so that they are contained within contour* γ_1. *We also note that* $\iota = \sqrt{-1}$.

Proof. Using (4.5) and (4.6) we can write

$$\mathbb{P}\big(\tau_N \leq x\big) = \frac{\prod_{1\leq i,j\leq N}(1-p_iq_j)}{\Delta_N(p)\Delta_N(q)} \sum_{x\geq\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_N\geq 0} \det\big(p_i^{\lambda_j+N-j}\big)_{1\leq i,j\leq N} \det\big(q_i^{\lambda_j+N-j}\big)_{1\leq i,j\leq N}, \quad (4.19)$$

where $\Delta_N(p)$ and $\Delta_N(q)$ are Vandermonde determinants. To bring the sum in (4.19) into form (4.7), (4.8), we make the change of variables $t_j := \lambda_j + N - j$ and write it as

$$\sum_{x+N-1\geq t_1>t_2>\cdots>t_N\geq 0} \det\big(p_i^{t_j}\big)_{1\leq i,j\leq N} \det\big(q_i^{t_j}\big)_{1\leq i,j\leq N}$$

Noticing that this sum is symmetric in variables $t_1, ..., t_N$ and that the summand vanishes if two of these are equal to each other (because the determinants do so in this case), we can extend the sum via symmetrization and write it as

$$\frac{1}{N!}\sum_{t_1,\ldots,t_N\in\mathbb{N}} \det\big(p_i^{t_j}\big)_{1\leq i,j\leq N} \det\big(q_i^{t_j}\big)_{1\leq i,j\leq N}\ \mathbb{1}_{[0,x+N-1]}(t_1)\cdots\mathbb{1}_{[0,x+N-1]}(t_N).$$

Thus, we can write (4.19) in the form (4.7) and (4.8) with $\phi_i(t) = p_i^t$, $\psi_j(t) = q_j^t$, $t \in \{0, 1, ...\}$ and μ being the counting measure. We will now use Proposition 4.3 and in this setting we compute

$$\mathsf{G}_{ij} = \sum_{t\geq 0}(p_iq_j)^t = \frac{1}{1-p_iq_j}. \quad (4.20)$$

To invert this matrix we will use Cramer's formula, which states that

$$\big(\mathsf{G}^{-1}\big)_{ij} = \frac{(-1)^{i+j}\det\mathsf{G}^{ji}}{\det\mathsf{G}}, \quad (4.21)$$

where G^{ji} denotes the minor matrix derived from G by deleting row j and column i. One of the computable determinants goes under the name of Cauchy determinant and is

$$\det\Big(\frac{1}{a_i-b_j}\Big)_{1\leq i,j\leq N} = (-1)^{\frac{n(n-1)}{2}}\frac{\Delta_N(a)\Delta_N(b)}{\prod_{1\leq i,j\leq N}(a_i-b_j)}$$

and so

$$\det\mathsf{G} = \frac{\prod_{1\leq k<\ell\leq N}(p_k-p_\ell)\ \prod_{1\leq k<\ell\leq N}(q_k-q_\ell)}{\prod_{k,\ell=1}^N(1-p_kq_\ell)} \quad (4.22)$$

one can also compute $\det\mathsf{G}^{ji}$ observing that this is also a Cauchy determinant of the same type and so the same formula as in (4.22) will be valid, just without terms which contain variables p_i and q_j. Thus, we obtain that

$$\big(\mathsf{G}^{-1}\big)_{ji} = \frac{\prod_{1\leq\ell\leq N}(1-p_jq_\ell)\ \prod_{1\leq k\leq N}(1-p_kq_i)}{(1-p_jq_i)\ \prod_{\ell\neq j}(p_j-p_\ell)\ \prod_{k\neq i}(q_i-q_k)}.$$

Inserting this into (4.15) with the choice $\psi_i(t) = q_i^t$ and $\phi_j(s) = p_j^s$, we obtain that (4.19) can be written as a Fredholm determinant

$$\mathbb{P}\big(\tau_N \leq x\big) = \det\big(I + g_N K_N^{LPP}\big)_{L^2(\mathbb{N})}$$

with

$$K_N^{LPP}(t,s) = \sum_{1\leq i,j\leq N} q_i^t p_j^s\ \frac{\prod_{1\leq\ell\leq N}(1-p_jq_\ell)\ \prod_{1\leq k\leq N}(1-p_kq_i)}{(1-p_jq_i)\ \prod_{\ell\neq j}(p_j-p_\ell)\ \prod_{k\neq i}(q_i-q_k)}.$$

Using the Residue Theorem we can write this kernel in an integral form

$$K_N^{LPP}(t,s) = \frac{1}{(2\pi\iota)^2}\int_{\gamma_1}\mathrm{d}\zeta\int_{\gamma_2}\mathrm{d}\eta\ \frac{\eta^s\zeta^t}{1-\zeta\eta}\prod_{j=1}^N\Big(\frac{1-\eta q_j}{\eta-p_j}\Big)\prod_{i=1}^N\Big(\frac{1-p_i\zeta}{\zeta-q_i}\Big)$$

finishing the proof. With regards to the application of the residue theorem we note that the function $\prod_{j=1}^N(\frac{1-\eta q_j}{\eta-p_j})$ has poles (p_i) included in the contout γ_2 and the function $\prod_{i=1}^N(\frac{1-p_i\zeta}{\zeta-q_i})$ has poles (q_i) included inside γ_1, while the choice of the contours excludes a pole from the case $\zeta\eta = 1$. □

Let us finish off this section by the central theorem:

Theorem 4.5. *Let the parameters of the geometric random variables to be all equal i.e.* $p_i = q_i = \sqrt{\alpha}$*. Then there are explicit constants* $\mathsf{f} = \mathsf{f}(\alpha)$ *and* $\sigma = \sigma(\alpha)$ *(*σ *here is not to be confused with a permutation !) such that*

$$\lim_{N\to\infty} \mathbb{P}\big(\tau_N \le \mathsf{f}N + \sigma N^{1/3}x\big) = \det\big(I + K_{\mathrm{Airy}_2}\big)_{L^2(x,\infty)}.$$

where K_{Airy_2} *is the Tracy-Widom GUE (or Airy-two) kernel given explicitly by*

$$K_{\mathrm{Airy}_2}(t,s) = \int_0^\infty Ai(\lambda+t)\,Ai(\lambda+s)\,\mathrm{d}\lambda, \qquad t,s\in\mathbb{R},$$

and $Ai(t)$ *is this Airy function, which is the solution of the second order ODE* $u'' = xu$ *subject to the condition that* $u(x)\to 0$ *when* $x\to\infty$ *and admits the contour integral representation*

$$Ai(t) = \frac{1}{2\pi\iota}\int_{\gamma_1} e^{z^3/3 - t\,z}\,\mathrm{d}z, \tag{4.23}$$

with the contour γ_1 *given by*

$$\gamma_1 = \{re^{-\iota\pi/3}\colon r\in(-\infty,0)\}\cup\{re^{\iota\pi/3}\colon r\in(0,\infty)\}, \qquad \text{traced upwards}$$

The proof of this theorem amounts to computing the asymptotics of the Fredholm determinant (4.17) and (4.18) where now g_N will be $\mathbb{1}_{[\mathsf{f}N+\sigma N^{1/3}+N}$ and

$$K_N^{LPP}(t,s) = \frac{1}{(2\pi\iota)^2}\int_{\gamma_1}\mathrm{d}\zeta\int_{\gamma_2}\mathrm{d}\eta\;\frac{\eta^s\,\zeta^t}{1-\zeta\eta}\,\Big(\frac{1-\eta\alpha}{\eta-\alpha}\Big)^N\Big(\frac{1-\alpha\zeta}{\zeta-\alpha}\Big)^N$$

The asymptotics of this integral can be computed with the method of *steepest descent.* We refer to [Z22], Section 7.3 for details.

5. Non-intersecting paths

In this section we will study how determinantal formulas arise via weights of ensembles of non-intersecting paths. Before going back to RSK and proving some of its properties, we will revisit Schur functions and we will prove some determinantal formulas known by the name Jacobi-Trudi identities. Besides the interest of these identities on their own, the will show us how to compute the total weight (or probabilities) of non-intersecting paths in a determinantal form.

5.1. Jacobi-Trudi identities and the Lindström-Gessel-Viennot theorem. The Jacobi-Trudi identity reads as

Theorem 5.1. *For a partition* $\lambda = (\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n)$ *and indeterminates* $x = (x_1, x_2, ..., x_n)$ *we have that*

$$s_\lambda(x) = \det\big(h_{\lambda_i - i + j}(x)\big)_{1\le i,j\le n}, \tag{5.1}$$

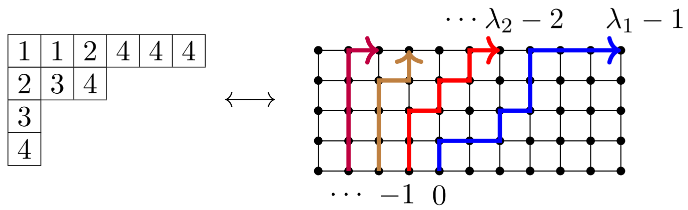
where $h_n(x) = \sum_{1\le i_1\le\cdots\le i_n\le n} x_{i_1}\cdots x_{i_n}$ *are the complete homogeneous symmetric functions.*

The proof of this theorem has two components. The first is to express the Schur function as a total weight of an ensemble of non-intersecting paths. The second is the Lindström-Gessel-Viennot Theorem, which expresses the total weight of an ensemble of non-intersecting paths in a determinantal form.

Let us start with the first component, which requires to represent a Young tableau as a path ensemble. The mapping goes as follows:

1. each row of a Young tableau is represented by a lattice. In particular the i^{th} row will be represented by a path on $\mathbb{Z}\times\mathbb{N}$ starting from $(-i,0)$ and moving up-right to $(\lambda_i - i)$
2. the horizontal steps of the path are encoded by the numbers in the boxes of row i. In particular, if $a_1, a_2, ...$ are the numbers in the i^{th} row, read from left to right, then the first horizontal step of the i^{th} path is at level a_1, the second at level a_2 etc.

For example, we have the mapping:



(5.2)

Exercise 18. *Show that this mapping maps a Young tableau to an ensemble of non-intersecting paths.*

We will now give a weight to each path in a similar way we did in the representation (3.6). In other words, vertical steps will have weight 1 and a horizontal step at level i will have weight x_i. The weight of a single path π will be

$$\mathrm{wt}(\pi) := \prod_{i=1}^{n} x_i^{\sharp\,\{\text{horizontal steps of } \pi \text{ at level } i\}},$$

and the total weight of a collection of paths $\pi_1, \pi_2, \ldots$ will be

$$\mathrm{wt}(\pi_1 \sqcup \pi_2 \sqcup \cdots) = \prod_{j \geq 1} \mathrm{wt}(\pi_j).$$

Given now the definition of Schur functions as

$$s_\lambda(x) = \sum_{T \in \mathrm{SSYT}(\lambda)} \prod_{i=1}^{n} x_i^{\sharp\,\{i's \text{ in } T\}}$$

and the above mapping, we easily get that

$$s_\lambda(x) = \mathrm{wt}\Big(\Pi_{(-i,0):\, i=1,\ldots,n}^{(\lambda_i - i):\, i=1,\ldots,n}\Big),$$

where for a collection of paths Π, the total weight $\mathrm{wt}(\Pi) = \sum_{\pi \in \Pi} \mathrm{wt}(\pi)$ and $\Pi_{(-i,0):\, i=1,\ldots,n}^{(\lambda_i - i, n):\, i=1,\ldots,n}$ is the collection of non-intersecting up-right paths from $(-1,0), (-2,0), \ldots, (-n,0)$ to $(\lambda_1 - 1, n), (\lambda_2 - 2, n), \ldots, (\lambda_n - n, n)$ as in (5.2).

We also note by (3.5) and (3.6) that

$$\mathrm{wt}\Big(\Pi_{(-1,0)}^{(k-1,n)}\Big) = h_k(x_1, \ldots, x_n).$$

Given the above consideration, the proof of the Jacobi-Trudi identity (5.1) will follow from the following theorem:

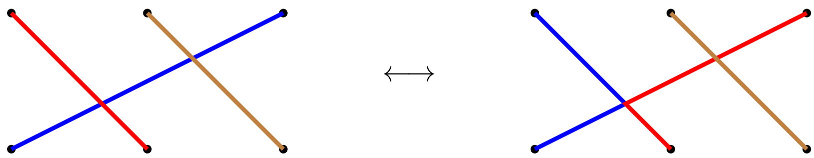
Theorem 5.2 (Lindström-Gessel-Viennot). *Let $G = (V, E)$ be a directed, acyclic graph with no multiple edges, with each edge $\boldsymbol{e}$ being assigned a weigth $\mathrm{wt}(\boldsymbol{e})$. A path π on G is assigned a weight $\mathrm{wt}(\pi) = \prod_{\boldsymbol{e} \in \pi} \mathrm{wt}(\boldsymbol{e})$. We say that two paths on G are non-intersecting if they do not share any vertex. Consider, now, $(u_1, \ldots, u_r)$ and $(v_1, \ldots, v_r)$ two disjoint subsets of V and denote by $\Pi_{v_1,\ldots,v_r}^{u_1,\ldots,u_r}$ the set of all r-tuples of non-intersecting paths $\pi_1, \ldots, \pi_r$ that start from $u_1, \ldots, u_r$ and end at $v_1, \ldots, v_r$, respectively. We assume that $\{u_1, \ldots, u_r\}$ and $\{v_1, \ldots, v_r\}$ have the property that for $i < j$ and $i' > j'$, any two paths $\pi \in \Pi_{v_j}^{u_i}$ and $\pi' \in \Pi_{v_{j'}}^{u_{i'}}$, which start at $u_i, u_{i'}$ and end at $v_j, v_{j'}$, necessarily intersect. Then*

$$\det\Big(\mathrm{wt}\big(\Pi_{v_j}^{u_i}\big)\Big) = \det\Big(\sum_{\pi \in \Pi_{v_j}^{u_i}} \mathrm{wt}(\pi)\Big)_{1 \leq i,j \leq r} = \sum_{\pi_1,\ldots,\pi_r \,\in\, \Pi_{v_1,\ldots,v_r}^{u_1,\ldots,u_r}} \mathsf{wt}(\pi_1) \cdots \mathsf{wt}(\pi_r).$$

Proof. We start by expanding the determinant in the left-hand side:

$$\begin{aligned}
\det\Big(\mathrm{wt}\big(\Pi^{u_i}_{v_j}\big)\Big) &= \sum_{\sigma\in S_n} sgn(\sigma)\prod_{i=1}^n \mathrm{wt}\big(\Pi^{v_{\sigma(i)}}_{u_i}\big)\\
&= \sum_{\sigma\in S_n} sgn(\sigma)\prod_{i=1}^n \sum_{\pi_i\in\Pi^{v_{\sigma(i)}}_{u_i}} \mathrm{wt}\big(\pi_i\big)\\
&= \sum_{\sigma\in S_n} sgn(\sigma) \sum_{\pi_i\in\Pi^{v_{\sigma(i)}}_{u_i}\,:\, i=1,\dots,n} \prod_{i=1}^n \mathrm{wt}\big(\pi_i\big)\\
&= \sum_{\substack{\pi_i\in\Pi^{v_i}_{u_i}\,:\, i=1,\dots,n\\ \text{non-intersecting}}} \prod_{i=1}^n \mathrm{wt}\big(\pi_i\big) + \sum_{\substack{\pi_i\in\Pi^{v_i}_{u_i}\,:\, i=1,\dots,n\\ \text{intersecting}}} \prod_{i=1}^n \mathrm{wt}\big(\pi_i\big)\\
&\quad + \sum_{\sigma\in S_n\,,\,\sigma\neq Id} sgn(\sigma)\prod_{i=1}^n \mathrm{wt}\big(\Pi^{v_{\sigma(i)}}_{u_i}\big)
\end{aligned}$$

We will show that the total sum of the second and third terms is equal to zero. Notice that the last two terms contain all paths that intersect. We construct a mapping (actually an involution) between intersecting paths that maintains the weight but reverses the sign of the permutation σ. To see how this works look at the following picture:



If paths meet find the first intersection point (going upwards, eg in the above picture the intersecting point of the blue and red path) and then map this set of paths to the set where the blue and red path are flipped. This new path is still in the class of intersecting paths. They have the same weight as the original set as the edges traced do not change but the sign of the corresponding $\sigma' s$ has changed, eg in the above example

$$\sigma = \begin{pmatrix} 1 & 2 & 3\\ 3 & 1 & 2\end{pmatrix} \qquad \text{has mapped to} \qquad \sigma' = \begin{pmatrix} 1 & 2 & 3\\ 1 & 3 & 2\end{pmatrix}$$

and note that $sgn(\sigma)=1$ while $sgn(\sigma')=-1$. Note also that the above mapping is an involution since if you apply it again, then the upper parts of the blue and red paths will exchange colours again, thus going back to the original set of paths. In other words the mapping is a bijection within the set on intersecting paths. Since the mapping maintains the weight of the paths but changes the sign the total (signed) weight will cancel, thus

$$\sum_{\substack{\pi_i\in\Pi^{v_i}_{u_i}\,:\, i=1,\dots,n\\ \text{intersecting}}} \prod_{i=1}^n \mathrm{wt}\big(\pi_i\big) + \sum_{\sigma\in S_n\,,\,\sigma\neq Id} sgn(\sigma)\prod_{i=1}^n \mathrm{wt}\big(\Pi^{v_{\sigma(i)}}_{u_i}\big) = 0$$

□

We can now complete the proof of the Jacobi-Trudi identity:

Prof of Theorem 5.1. By the mapping (5.2) we have that

$$s_\lambda(x) = \mathrm{wt}\Big(\Pi^{(\lambda_i-i,n)\,:\, i=1,\dots,n}_{(-i,0)\,:\, i=1,\dots,n}\Big),$$

and by the Lindström-Gessel-Viennot theorem we have that this equals

$$\det\Big(\Pi^{(\lambda_i-i,n)}_{(-j,0)}\Big) = \det\Big(h_{\lambda_i-i+j}(x)\Big).$$

□

Exercise 19. *Prove the following alternative Jacobi-Trudi identity. Recall that if $\lambda = (\lambda_1 \geq \cdots \geq \lambda_n)$ is a partition, then λ' is the conjugate partition. Recall also the elementary symmetric polynomials $e_k(x) := e_k(x_1, .., x_n) = \sum_{1 \leq i_1 < \cdots < i_k} x_{i_1} \cdots x_{i_k}$. We then have that*

$$s_\lambda(x) = s_\lambda(x_1, ..., x_n) = \det\left(e_{\lambda'_i - i + j}(x)\right)_{1 \leq i,j \leq n}.$$

5.2. Back to RSK.

Before getting to prove the properties of RSK developing a bit heavier methodology, let us warm up with a couple of fun applications.

Theorem 5.3 (Erdős-Szekeres theorem). *If $\sigma \in S_n$ is a permutation and $n > rs$ for some integers r, s, then either the longest increasing subsequence is $> r$ or the longest decreasing subsequence is $< s$.*

Proof. Map via RSK

$$\sigma \longleftrightarrow (P, Q).$$

Since $\sigma \in \S_n$ the P and Q tableau will have $n > rs$ boxes. The length of the longest increasing subsequence is the length of the first row and the length of the longest decreasing subsequence is the length of the first column but if the first is $\leq r$ and the second $\leq s$ the Young diagram P will be embedded in a rectangle of side lengths r, s, thus the total number of boxes will be $\leq rs$, which is a contradiction. □

The next proposition should be considered more as an exercise on RSK similar to that of the Cauchy identity.

Proposition 5.4. *Let h_μ be the complete homogeneous function corresponding to partition μ and $K_{\lambda\mu}$ be the Kostka number, i.e. the number of $SSYT(\lambda, \mu)$ of shape λ and type μ (the type of a SSYT is the vector whose i^{th} coordinate is the number of $i's$ in the tableau. We then have that*

$$h_\mu = \sum_\lambda K_{\lambda\mu} s_\lambda.$$

Proof.

$$\begin{aligned} h_\mu(x) = h_{\mu_1}(x) h_{\mu_2}(x) \cdots &= \prod_{j \geq 1} \Big(\sum_{\substack{\alpha_j^1, \ldots, \alpha_j^n \geq 0 \\ \alpha_j^1 + \cdots + \alpha_j^n = \mu_j}} x_1^{\alpha_j^1} \cdots x_n^{\alpha_j^n} \Big) \\ &= \sum_{\substack{\alpha_j^1, \ldots, \alpha_j^n \geq 0 \,:\, j \geq 1 \\ \alpha_j^1 + \cdots + \alpha_j^n = \mu_j}} \prod_{i \geq 1} x_i^{\sum_j \alpha_j^i} \end{aligned}$$

By the RSK correspondence we can change variables (ie map) and instead of sumingm over all matrices $A = (a_j^i)_{1 \leq i,j \leq n}$, where i correspond to rows and j to columns, sum over $(P, Q) = \mathsf{RSK}(A)$ tableaux or (Z_P, Z_Q) Gelfand-Tsetlin patterns. We can then write (recall the notation from (3.24) and (4.4))

$$\begin{aligned} h_\mu &= \sum_{(Z_P, Z_Q)} \mathbb{1}_{type(Z_P) = \mu} \prod_{i \geq 1} x_i^{|z_Q^i| - |z_Q^{i-1}|} \\ &= \sum_\lambda \Big(\sum_{Z_P \,:\, sh(Z_P) = \lambda} \mathbb{1}_{type(Z_P) = \mu} \Big) \Big(\sum_{Z_Q \,:\, sh(Z_Q) = \lambda} \prod_{i \geq 1} x_i^{|z^i| - |z^{i-1}|} \Big) \\ &= \sum_\lambda K_{\lambda\mu} s_\lambda(x). \end{aligned}$$

□

Our final fun proposition will give a combinatorial proof of the fact (in the case of S_n) that for a group G which decomposes to irreducible representations as $G = \oplus_i \, m_i V_i$, where m_i is the multiplicity of V_i, then

$$|G| = \sum_i (dim(V_i))^2.$$

More precisely, we have

Proposition 5.5. *Let f^λ be the number of $STY(\lambda)$, i.e. the standard Young Tableaux of shape λ. Then*

$$n! = \sum_\lambda (f^\lambda)^2.$$

Proof. The proof is just the RS bijection:

$$n! = \sum_{A:\text{ permutation matrix}} 1 = \sum_\lambda \sum_{P,Q\in SYT(\lambda)} 1 = \sum_\lambda (f^\lambda)^2.$$

□

5.3. Matrix construction of RSK and Kirillov's geometric lifting. Berenstein and Kirillov's $(max, +)$ formulation. Berenstein and Kirillov [BK95] adopted a different point of view of the RSK correspondence, which was to encode the combinatorial transformations via piecewise linear transformations. This is a particularly useful approach for applications in the probabilistic models we are interested in and we will describe it now. The exposition here follows mainly the presentation in Noumi-Yamada [NY04].

Let us describe how the P tableau is constructed. This is done by successive row insertions of words $w^1, w^2, \ldots$ as described in the RS correspondence. Let us start by inserting w^1, that is the sequence of letters

$$\underbrace{1\cdots 1}_{w_1^1}\ \underbrace{2\cdots 2}_{w_2^1}\cdots \underbrace{N\cdots N}_{w_N^1}.$$

Since the letters in w^1 are ordered from smaller to larger, the insertion of w^1 will produce the one-row tableau

$$P_1 = \boxed{\overbrace{1\cdots\cdots 1}^{w_1^1}}\boxed{\overbrace{2\cdots\cdots 2}^{w_2^1}}\boxed{\cdots\cdots\cdots\cdots}\boxed{\overbrace{N\cdots\cdots N}^{w_N^1}}$$

We note that we can identify the row of a tableau with words by reading the letters from left to right. In the case of P_1, the tableau can be identified with the single word (recall the notation introduced in (3.20))

$$p^1 := 1^{p_1^1} 2^{p_2^1} \cdots N^{p_N^1} = 1^{w_1^1} 2^{w_2^1} \cdots N^{w_N^1} = w^1.$$

Next, we insert word w^2 into P_1 and this insertion will produce a new tableau $\widetilde{P}_1$. We denote this schematically as

$$P_1 \overset{w^2}{\longrightarrow} \widetilde{P}_1.$$

This insertion will change the first row p^1 of P_1 by (possibly) bumping some letters out of it and replacing them with letters from w^2. The bumped letters will form a word, which will then be inserted in the second row of the tableau, which in the case of P_1 is $\emptyset$. We denote this schematically as

$$p^1 \underset{v^2}{\overset{w^2}{\longrightarrow}} \widetilde{p}^{\,1},$$

with p^1 denoting the first row of P_1, $\widetilde{p}^{\,1}$ the first row of $\widetilde{P}_1$ and v^2 the word that will form from the bumped down letters from P_1 after the insertion of w^2.

This picture is a building block of RSK, since the row insertion of a word w in a tableau P consisting of rows $p^1, p^2, ..., p^n$, can be decomposed as

$$\begin{array}{ccc} & w =: v^1 & \\ p^1 & \xrightarrow{\quad} & \widetilde{p}^1 \\ & v^2 & \\ p^2 & \xrightarrow{\quad} & \widetilde{p}^2 \\ & \vdots & \\ & v^n & \\ p^n & \xrightarrow{\quad} & \widetilde{p}^n. \\ & v^{n+1} & \end{array} \tag{5.3}$$

This picture means that the letters that will drop down from p^1, after the insertion of $v^1 = w$, will form a word v^2 which will be inserted in p^2, forming a new row $\widetilde{p}^{\,2}$, and will bump down letters which will then form a new word v^3 to be inserted into p^3 and so on.

An important remark is that row p^i, which is constructed via the RS algorithm, will only include letters with value larger or equal to i. This is an easy consequence of the algorithm. For example, p^2 will not include 1's as '1' is the smallest possible letter and so when a '1' is inserted in the first row it will stay there, bumping out 2's, 3's,...

It will be important to have an explicit, algebraic expression of the transformation $\boldsymbol{x} \xrightarrow[\boldsymbol{b}]{\boldsymbol{a}} \widetilde{\boldsymbol{x}}$, where variables $\boldsymbol{x} := i^{x_i}(i+1)^{x_{i+1}} \cdots N^{x_N}$ and $\boldsymbol{a} := i^{a_i}(i+1)^{a_{i+1}} \cdots N^{a_N}$, for $i \geq 1$, are considered as input variables ($\boldsymbol{x}$ corresponds to the generic case of an i-th row in a tableau) and $\widetilde{\boldsymbol{x}} := i^{\widetilde{x}_i}(i+1)^{\widetilde{x}_{i+1}} \cdots N^{\widetilde{x}_N}$ and $\boldsymbol{b} := (i+1)^{b_{i+1}}(i+2)^{b_{i+2}} \cdots N^{b_N}$ are output variables, with $\widetilde{\boldsymbol{x}}$ being the new row after the insertion of $\boldsymbol{a}$ and $\boldsymbol{b}$ being the bumped down letters (as explained in the previous paragraphs, $\boldsymbol{b}$ will only have letters strictly larger than i). In particular, we want to express $\widetilde{x}_i, \widetilde{x}_{i+1}, ...$ and $b_{i+1}, b_{i+2}, ...$ as piecewise linear transformations of $x_i, x_{i+1}, ...$ and $a_i, a_{i+1}, ...$ To do so, it will be more convenient to introduce cumulative variables

$$\xi_j := x_i + \cdots + x_j, \quad \text{and} \quad \widetilde{\xi}_j := \widetilde{x}_i + \cdots + \widetilde{x}_j, \qquad \text{for } j \geq i.$$

We derive the piecewise linear transformations as follows: When inserted into $\boldsymbol{x}$, the a_i letters i will bypass the already existing letters i in $\boldsymbol{x}$ and will be appended after the last i in $\boldsymbol{x}$. Thus, the new total number of letters i will be

$$\widetilde{\xi}_i = \xi_i + a_i.$$

When this insertion is completed a number of $(i+1)$'s will be bumped off $\boldsymbol{x}$. The number of these will equal

$$b_{i+1} = \min\left(\widetilde{\xi}_i - \xi_i, \xi_{i+1} - \xi_i\right) = \min\left(\widetilde{\xi}_i, \xi_{i+1}\right) - \xi_i. \tag{5.4}$$

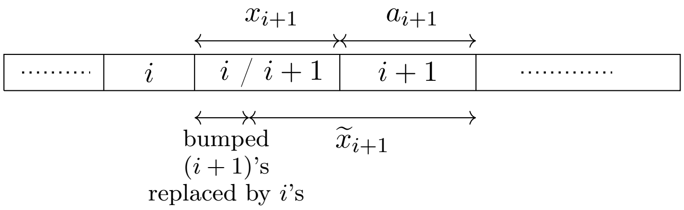
This is to be understood as follows: either a_i is smaller than the number of $(i+1)$'s in $\boldsymbol{x}$, which equals $x_{i+1} = \xi_{i+1} - \xi_i$, and so there will only be $a_i = \widetilde{\xi}_i - \xi_i$ number of $(i+1)$'s bumped down, or a_i is larger than or equal to x_{i+1}, in which case all of the $(i+1)$'s in $\boldsymbol{x}$, the number of which equals $\xi_{i+1} - \xi_i$, will be bumped down.

We also record an alternative formula for the number of bumped $(i+1)$'s, which is

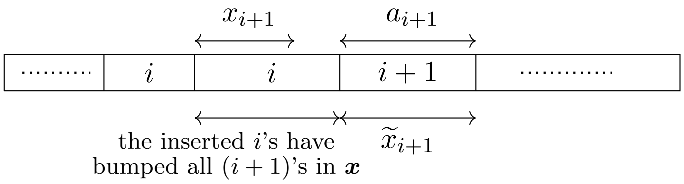
$$b_{i+1} = a_{i+1} + x_{i+1} - \widetilde{x}_{i+1} = a_{i+1} + (\xi_{i+1} - \xi_i) - (\widetilde{\xi}_{i+1} - \widetilde{\xi}_i), \tag{5.5}$$

where the first equality is to be understood as that the number of $(i+1)$'s, which will be bumped, equals the number of $(i+1)$'s that existed in $\boldsymbol{x}$ (denoted by x_{i+1}) plus the number of $(i+1)$'s that we inserted (denoted by a_{i+1}) minus the number of $(i+1)$'s that we finally see in $\widetilde{\boldsymbol{x}}$ (denoted by $\widetilde{x}_{i+1}$). This is depicted in the following figures, where blocks marked with i or $i+1$ indicate consecutive boxes occupied by i or

$i+1$. The first figure shows the case where the i's inserted in $\boldsymbol{x}$ do not bump out all $(i+1)$'s:



and the next figure depicts the situation where the i's inserted in $\boldsymbol{x}$ have bumped out all $(i+1)$'s:



in which case $\widetilde{x}_{i+1} = a_{i+1}$ and so $b_{i+1} = a_{i+1} + x_{i+1} - \widetilde{x}_{i+1} = x_{i+1}$.

We now want to get an expression for $\widetilde{\xi}_{i+1}$, which denotes the total number of numbers up to $i+1$ that exist in the output word $\widetilde{\boldsymbol{x}}$. Again, either the i's that we inserted from $\boldsymbol{a}$ did not bump all the $(i+1)$'s that existed in $\boldsymbol{x}$ (first of the two figures above), in which case the a_{i+1}-many of $(i+1)$'s, which are inserted from $\boldsymbol{a}$ will be appended at the end of the last $(i+1)$ in $\boldsymbol{x}$, giving $\widetilde{\xi}_{i+1} = \xi_{i+1} + a_{i+1}$, or the a_i-many i's in $\boldsymbol{a}$ bumped all the $(i+1)$'s in $\boldsymbol{x}$ (second of the two figures above), in which case the new $(i+1)$'s from $\boldsymbol{a}$ will be appended after the i's in $\widetilde{\boldsymbol{x}}$. In this case, $\widetilde{\xi}_{i+1} = \widetilde{\xi}_i + a_{i+1}$. Altogether, we have that

$$\widetilde{\xi}_{i+1} = \max\big(\widetilde{\xi}_i, \xi_{i+1}\big) + a_{i+1}. \tag{5.6}$$

We now can iterate this procedure through diagram (5.3). The RSK row insertion via piecewise linear transformations can be summarised as

Proposition 5.6. *Let $1 \le i \le N$. Consider two words $\boldsymbol{x} = i^{x_i}(i+1)^{x_{i+1}} \cdots N^{x_N}$ and $\boldsymbol{a} = i^{a_i}(i+1)^{a_{i+1}} \cdots N^{a_N}$. The row insertion of the word $\boldsymbol{a}$ into the word $\boldsymbol{x}$ denoted by*

$$\boldsymbol{x} \xrightarrow[\boldsymbol{b}]{\boldsymbol{a}} \widetilde{\boldsymbol{x}}\ ,$$

transforms $(\boldsymbol{x}, \boldsymbol{a})$ into a new pair $(\widetilde{\boldsymbol{x}}, \boldsymbol{b})$ with $\widetilde{\boldsymbol{x}} = i^{\widetilde{x}_i}(i+1)^{\widetilde{x}_{i+1}} \cdots N^{\widetilde{x}_N}$ and $\boldsymbol{b} = (i+1)^{b_{i+1}} \cdots N^{b_N}$, which in cumulative variables

$$\xi_j = x_i + \cdots + x_j, \quad \text{and} \quad \widetilde{\xi}_j = \widetilde{x}_i + \cdots + \widetilde{x}_j, \qquad \text{for } j \ge i, \tag{5.7}$$

is encoded via

$$\begin{cases} \widetilde{\xi}_i = \xi_i + a_i, \\ \widetilde{\xi}_k = \max\big(\widetilde{\xi}_{k-1}, \xi_k\big) + a_k, & i+1 \le k \le N \\ b_k = a_k + (\xi_k - \xi_{k-1}) - (\widetilde{\xi}_k - \widetilde{\xi}_{k-1}), & i+1 \le k \le N, \end{cases} \tag{5.8}$$

for $i < N$. If $i = N$, then $\widetilde{\xi}_N = \xi_N + a_N$ and the output $\boldsymbol{b}$ is empty and we write $\boldsymbol{b} = \emptyset$.

It is worth noticing that the recursion

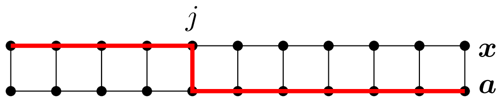
$$\widetilde{\xi}_k = \max\big(\widetilde{\xi}_{k-1}, \xi_k\big) + a_k,$$

is actually the same as the recursion of last passage percolation. To see this more clearly, we can consider the example of $\boldsymbol{x} = 1^{x_1}2^{x_2} \cdots N^{x_N}, \boldsymbol{a} = 1^{a_1}2^{a_2} \cdots N^{a_N}$ and for $\xi_j = x_1 + \cdots + x_j$ and $\widetilde{\xi}_j = \widetilde{x}_1 + \cdots \widetilde{x}_j$ we iterate as

$$
\begin{aligned}
\widetilde{\xi}_N &= \max\left(\widetilde{\xi}_{N-1} + a_N, \xi_N + a_N\right) \\
&= \max\Big(\max\left(\widetilde{\xi}_{N-2} + a_{N-1} + a_N, \xi_{N-1} + a_{N-1} + a_N\right), \xi_N + a_N \Big) \\
&\vdots \\
&= \max_{1 \le j \le N} \left(\xi_j + a_j + \cdots + a_N\right) \\
&= \max_{1 \le j \le N} \left(x_1 + \cdots + x_j + a_j + \cdots + a_N\right), \qquad (5.9)
\end{aligned}
$$

which, as shown in the figure below, is a last passage percolation on a two-row array

$\max_j \sum$ weights of nodes along red path



(5.10)

This is an indication of the relevance of RSK, and in particular this formulation, for last passage percolation and other models in the KPZ class. Moreover, this formulation is amenable to a generalisation which will be important in treating the positive temperature case relating to directed polymers. In Section 5.5 we will prove this connection in more detail.

5.4. A geometric lifting of RSK - Kirillov's "Tropical RSK".

As we have seen in Proposition 5.6, RSK can be encoded in terms of piecewise linear recursive relations, using the $(\max, +)$ algebra. Kirillov [K01] replaced the $(\max, +)$ in the set of RSK's piecewise linear relations with relations $(+, \times)$, thus establishing a *geometric lifting* of RSK, which he named *tropical* RSK. This name was given by Kirillov in honour of P. Schützenberger, see [K01], page 84, for some clues regarding the etymology of this name. However, since the term tropical has been reserved for the opposite passage from the $(+, \times)$ algebra to the $(\max, +)$, the term geometric RSK (gRSK) has now prevailed for the geometric lifting of the RSK correspondence. In this section we will present the construction of gRSK following mostly a matrix reformulation by Noumi and Yamada [NY04] motivated by discrete integrable systems. The approach is closely related to that of Proposition 5.6. Let us start with the definition of the *geometric row insertion*.

5.5. Geometric RSK via a matrix formulation.

This section will not be examinable but please enjoy it.

We start with the following definition (compare to Proposition 5.6):

Definition 5.7. *Let* $1 \le i \le N$. *Consider two words* $\boldsymbol{x} = (x_i, ..., x_N)$ *and* $\boldsymbol{a} = (a_i, ..., a_N)$. *We define the* **geometric lifting of row insertion** *or shortly* **geometric row insertion** *of the word* $\boldsymbol{a}$ *into the word* $\boldsymbol{x}$, *denoted by*

$$
\boldsymbol{x} \xrightarrow[\boldsymbol{b}]{\boldsymbol{a}} \widetilde{\boldsymbol{x}} \ ,
$$

as the transformation that takes $(\boldsymbol{x}, \boldsymbol{a})$ *into a new pair* $(\widetilde{\boldsymbol{x}}, \boldsymbol{b})$ *with* $\widetilde{\boldsymbol{x}} = (\widetilde{x}_i, \cdots \widetilde{x}_N)$ *and* $\boldsymbol{b} = (b_{i+1} \cdots b_N)$, *which in cumulative variables*

$$
\xi_j = x_i \cdots x_j, \quad \text{and} \quad \widetilde{\xi}_j = \widetilde{x}_i \cdots \widetilde{x}_j, \qquad \text{for } j \ge i, \tag{5.11}
$$

is encoded via

$$
\begin{cases}
\widetilde{\xi}_i = \xi_i \cdot a_i, \\
\widetilde{\xi}_k = a_k \big(\widetilde{\xi}_{k-1} + \xi_k\big), & i+1 \le k \le N \\
b_k = a_k \frac{\xi_k \, \widetilde{\xi}_{k-1}}{\xi_{k-1} \, \widetilde{\xi}_k}, & i+1 \le k \le N.
\end{cases} \tag{5.12}
$$

for $i < N$. *If* $i = N$, *then* $\widetilde{\xi}_N = \xi_N \cdot a_N$ *and the output* $\boldsymbol{b}$ *is empty and we write* $\boldsymbol{b} = \emptyset$.

It was observed by Noumi and Yamada that the geometric row insertion $(\boldsymbol{x},\boldsymbol{a}) \mapsto (\widetilde{\boldsymbol{x}},\boldsymbol{b})$, described in Definition 5.7 via relations (5.12), is equivalent to a system of equations, related to *discrete Toda systems* (see [NY04], Lemma 2.2 and Remark 2.3):

$$\begin{aligned} a_i x_i = \widetilde{x}_i, \quad &\text{and} \quad a_j x_j = \widetilde{x}_j b_j \quad \text{for } j \geq i+1, \\ \frac{1}{a_i} + \frac{1}{x_{i+1}} = \frac{1}{b_{i+1}}, \quad &\text{and} \quad \frac{1}{a_j} + \frac{1}{x_{j+1}} = \frac{1}{\widetilde{x}_j} + \frac{1}{b_{j+1}} \quad \text{for } j \geq i+1. \end{aligned} \tag{5.13}$$

The derivation of the system of equations in (5.13) from (5.12) is a matter of a simple algebraic manipulation. (5.13) can be put into a matrix form as:

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \bar{a}_i & 1 & & \\ & & & & \bar{a}_{i+1} & 1 & \\ & & & & & \ddots & \ddots \\ & & & & & & \bar{a}_N \end{pmatrix} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \bar{x}_i & 1 & & \\ & & & & \bar{x}_{i+1} & 1 & \\ & & & & & \ddots & \ddots \\ & & & & & & \bar{x}_N \end{pmatrix} \tag{5.14}$$

$$= \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \bar{\widetilde{x}}_i & 1 & & \\ & & & & \bar{\widetilde{x}}_{i+1} & 1 & \\ & & & & & \ddots & \ddots \\ & & & & & & \bar{\widetilde{x}}_N \end{pmatrix} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & 1 & & & \\ & & & & \bar{b}_{i+1} & 1 & \\ & & & & & \ddots & \ddots \\ & & & & & & \bar{b}_N \end{pmatrix}$$

where we have used the notation

$$\bar{x} := \frac{1}{x},$$

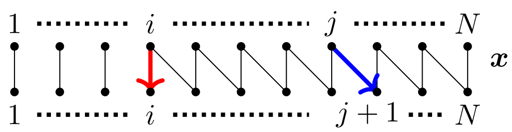
for a nonnegative real x. For a vector $\boldsymbol{x}$ with nonnegative real entries, we will denote by $\bar{\boldsymbol{x}}$ the vector whose entries are the inverses of the corresponding entries in $\boldsymbol{x}$. In the above matrices the entries that are left empty are equal to zero. Moreover, in the first three matrices the upper left corner is an identity matrix of size $i-1$, while in the fourth matrix the upper left corner is an identity matrix of size i. We put this matrix equation in a more concrete notation as follows: For a vector $\boldsymbol{x} = (x_i, ..., x_n)$ we define the matrix

$$E_i(\boldsymbol{x}) := \sum_{j=1}^{i-1} E_{jj} + \sum_{j=i}^{N} x_j E_{jj} + \sum_{j=i}^{N-1} E_{j,j+1}, \tag{5.15}$$

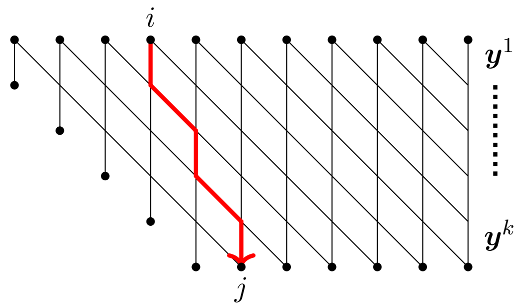
where for $1 \leq i,j \leq n$ we define the matrices $E_{ij} := (\delta_{ai}\delta_{bj})_{1\leq a,b\leq n}$ with δ_{ab} being the Kronecker delta. In other words, E_{ij} has entry (i,j) equal to 1 and all others equal to zero. We will denote $E_1(\boldsymbol{x})$ simply by $E(\boldsymbol{x})$. Then (5.14) is written more concretely as

$$E_i(\bar{\boldsymbol{a}})E_i(\bar{\boldsymbol{x}}) = E_i(\bar{\widetilde{\boldsymbol{x}}})E_{i+1}(\bar{\boldsymbol{b}}) \tag{5.16}$$

The entries of matrices $E_i(\boldsymbol{x})$, with $\boldsymbol{x} = (x_i, ..., x_N)$, can be readily read graphically from the following diagram



where on the diagonal edges and on the first $(i-1)$ vertical edges we assign the value 1 and on the rest of the vertical edges we assign the values $x_i, x_{i+1}, \dots, x_N$ in this order. Then the (k,ℓ) entry of $E_i(\boldsymbol{x})$ is given by $E_i(\boldsymbol{x})_{(k,\ell)} = \sum_{\pi:\,(1,k)\to(2,\ell)} \mathsf{wt}(\pi)$, where the sum is over all down-right paths (consisting in this case of one step), along existing edges, starting from site k in the top row to site ℓ in the bottom row and the weight of the path π is given by the product of the weights along the edges that path π traces. Furthermore, one can easily check that for products of the form $E(\boldsymbol{y}^1, ..., \boldsymbol{y}^k) := E_1(\boldsymbol{y}^1)E_2(\boldsymbol{y}^2)\cdots E_k(\boldsymbol{y}^k)$, where we understand that for $i = 1, ..., k$ the vector $\boldsymbol{y}^i = (y^i_i, ..., y^i_N)$, the entries can be read graphically from the following diagram:

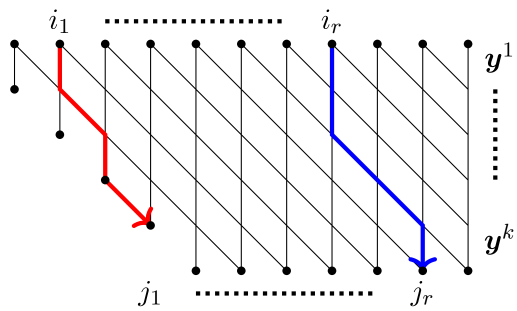


where a vertical edge connecting (a, b) to $(a+1, b)$ (in matrix coordinates) is assigned the weight y^a_b and all the diagonal edges are assigned weight one. Entry (i, j) of the matrix $E(\boldsymbol{y}^1, ..., \boldsymbol{y}^k)$ is given by $E(\boldsymbol{y}^1, ..., \boldsymbol{y}^k)_{(i,j)} = \sum_{\pi:\,(1,i)\to(k\wedge j+1,j)} \mathsf{wt}(\pi)$, where the sum is over all down-right paths, along existing edges, from site $(1, i)$ (in matrix coordinates) in the top row to site $(k \wedge j + 1, j)$, $k \wedge j := \min(k, j)$, along the lower border and where the weight of a path is $\mathsf{wt}(\pi) := \prod_{\boldsymbol{e}\in\pi} w_{\boldsymbol{e}}$, with the product over all edges $\boldsymbol{e}$ that are traced by the path π.

More remarkably, minor determinants of matrices $E(\boldsymbol{y}^1, ..., \boldsymbol{y}^k)$ have also a similar graphical representation. If we denote by $\det E(\boldsymbol{y}^1, ..., \boldsymbol{y}^k)^{i_1,...,i_r}_{j_1,...,j_r}$ the determinant of the sub-matrix of $E(\boldsymbol{y}^1, ..., \boldsymbol{y}^k)$ which includes rows $i_1 < \cdots < i_r$ and columns $j_1 < \cdots < j_r$ then

$$\det E(\boldsymbol{y}^1, ..., \boldsymbol{y}^k)^{i_1,...,i_r}_{j_1,...,j_r} = \sum_{\pi_1,...,\pi_r \in \Pi^{i_1,...,i_r}_{j_1,...,j_r}} \mathsf{wt}(\pi_1)\cdots\mathsf{wt}(\pi_r), \tag{5.17}$$

where $\Pi^{i_1,...,i_r}_{j_1,...,j_r}$ is the set of directed, non-intersecting paths, starting at locations $i_1, ..., i_r$ in the top row and ending at locations $j_1, ..., j_r$ at the bottom border of the grid :



Notice that the non-intersecting property and the orderings $i_1 < \cdots < i_r$ and $j_1 < \cdots j_r$ enforces that the path starting at i_a will end at j_a for all $a = 1, ..., r$.

Since $E(\boldsymbol{y}^1, ..., \boldsymbol{y}^k)_{(i,j)} = \sum_{\pi:\,(1,i)\to(k\wedge j+1,j)} \mathsf{wt}(\pi)$, (5.17) is a consequence of the Lindström-Gessel-Viennot theorem.

There is also a set of matrices dual to $E_i(\boldsymbol{x})$, which we now introduce, the entries and minor determinants of which are given in terms of the total weights of paths moving in the more standard up-right, directed fashion. Let us start with the $i = 1$ case, where we recall the convention that $E_1(\boldsymbol{x}) = E(\boldsymbol{x})$, and define

$$H(\boldsymbol{x}) := DE(\bar{\boldsymbol{x}})^{-1}D^{-1}, \qquad \text{with} \qquad D = \operatorname{diag}\big((-1)^{i-1}\big)_{i=1}^N. \tag{5.18}$$

An easy computation shows that $H(\boldsymbol{x}) = \sum_{1\le i\le j\le N} x_i x_{i+1}\cdots x_j E_{ij}$, that is, the (i,j) entry of $H(\boldsymbol{x})$ equals $x_i x_{i+1}\cdots x_j$ if $i\le j$ and zero otherwise.

In general, for $k\ge 1$, and $\boldsymbol{x}=(x_k,...,x_n)$, in which case we note that $H(\boldsymbol{x}) = H(x_k,...,x_N)$ is a $k\times k$ matrix, we define the $N\times N$ matrix

$$H_k(\boldsymbol{x}) := \begin{bmatrix} I_{k-1} & \mathbf{0} \\ \mathbf{0} & H(\boldsymbol{x}) \end{bmatrix}$$

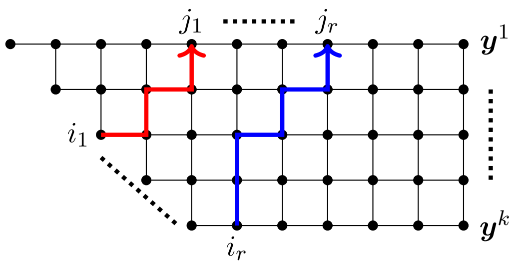
where I_{k-1} is a $(k-1)\times(k-1)$ identity matrix. Then using (5.18) equation (5.16) can be equivalently written as

$$H_i(\boldsymbol{x})H_i(\boldsymbol{a}) = H_{i+1}(\bar{\boldsymbol{b}})H_i(\widetilde{\boldsymbol{x}}) \tag{5.19}$$

Similarly to $E(\boldsymbol{y}^1,...,\boldsymbol{y}^k)$, products of the form $H(\boldsymbol{y}^1,...,\boldsymbol{y}^n) := H(\boldsymbol{y}^1)\cdots H(\boldsymbol{y}^n)$ or more generally $H(\boldsymbol{y}^1,...,\boldsymbol{y}^k) := H_k(\boldsymbol{y}^k)\cdots H_1(\boldsymbol{y}^1)$ have the property that their entries and their minor determinants are given via ensembles of non-intersecting paths. This is again a consequence of the Lindström-Gessel-Viennot theorem: denoting by $\det H(\boldsymbol{y}^1,...,\boldsymbol{y}^k)^{i_1,...,i_r}_{j_1,...,j_r}$ the minor determinant of $H(\boldsymbol{y}^1,...,\boldsymbol{y}^k)$ consisting of rows $i_1<\cdots<i_r$ and columns $j_1<\cdots<j_r$, then

$$\det H(\boldsymbol{y}^1,...,\boldsymbol{y}^k)^{i_1,...,i_r}_{j_1,...,j_r} = \sum_{\pi_1,...,\pi_r} \mathsf{wt}(\pi_1)\cdots\mathsf{wt}(\pi_r)$$

where the sum is over up-right, non crossing paths, starting at locations $i_1,...,i_r$ at the bottom border (including possibly the diagonal part) and ending at locations $j_1,...,j_r$ at the top row of the grid.



Each vertex (a,b) of the grid is assigned a weight y^a_b and in this case the total weight of a path is $\mathsf{wt}(\pi) = \prod_{(a,b)\in\pi} y^a_b$.

Noumi and Yamada [NY04] used these observations in order to give a matrix reformulation of geometric RSK allowing the output of geometric RSK to be expressed in terms of total weight (often also called partition functions) of ensembles of non-intersecting paths. The main theorem towards this is the following whose origins are in the theory of *total positivity.*

Theorem 5.8. *Given a matrix* $\mathsf{X} := (x^i_j : 1\le i\le n,\ 1\le j\le N) =: (\boldsymbol{x}^1,...,\boldsymbol{x}^n)^{\mathsf{T}}$ *the matrix equation*

$$H(\boldsymbol{x}^1)H(\boldsymbol{x}^2)\cdots H(\boldsymbol{x}^n) = H_k(\boldsymbol{y}^k)H_{k-1}(\boldsymbol{y}^{k-1})\cdots H_1(\boldsymbol{y}^1), \qquad k=\min(n,N), \tag{5.20}$$

has a unique solution $(\boldsymbol{y}^1,...,\boldsymbol{y}^k)$ *with* $\boldsymbol{y}^i := (y^i_i,...,y^i_N)$, *given by*

$$y^i_i = \frac{\tau^i_i}{\tau^{i-1}_i}, \quad \text{and} \quad y^i_j = \frac{\tau^i_j\,\tau^{i-1}_{j-1}}{\tau^{i-1}_j\,\tau^i_{j-1}} \quad \text{for} \quad i<j, \tag{5.21}$$

where

$$\tau^i_j := \sum_{\pi_1,...,\pi_i\in\Pi^{1,...,i}_{j-i+1,...,j}} \mathsf{wt}(\pi_1)\cdots\mathsf{wt}(\pi_i) \tag{5.22}$$

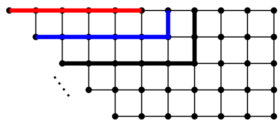
is the total weight (partition function) of an ensemble of i *non-intersecting, down-right paths* $\pi_1,...,\pi_i$, *along the entries of* X, *starting from* $(1,1),...,(1,i)$ *and ending at* $(n,j-i+1),...,(n,j)$, *respectively, with the weight of a path* π_r *given by* $\mathsf{wt}(\pi_r) := \prod_{(a,b)\in\pi_r} x^a_b$.

Proof. The fact that the left-hand side of (5.20) can be written uniquely as a product of the form of the right-hand side of (5.20) follows from a more general result of Berenstein-Fomin-Zelevinsky [?], which states that any upper triangular matrix $A = (a^i_j)_{1\le i,j\le n}$ such that

- $a^i_j = 0$, if $j < i$ or $j > i+m$ for some $m \le n$, i.e. A is a “band” upper triangular matrix . Moreover, we assume that $a^i_j = 1$ for $j = i+m$,
- the minor determinants
$$Q_{i,j} := Q_{i,j}(A) := \det A^{1,\dots,j-i+1}_{i,i+1,\dots,j},$$
are non-zero for all i, j such that $i \le j$ and $i \le m$,

can be written uniquely in the form $H_k(\boldsymbol{y}^k)H_{k-1}(\boldsymbol{y}^{k-1})\cdots H_1(\boldsymbol{y}^1)$. Let us note that the second bullet is essentially enforced by the first but we preferred to state it separately as the non-vanishing of these minor determinants is an important feature in this matrix representation as well as in terms of the path representation that follows. We refer for this general result and proofs to [NY04], Propositions 1.5, 1.6 and Theorem 2.4. The proof of this statement is a clever linear algebra using the Gauss decomposition but we prefer to omit it as we would like to go straight to the connection to path weights.

It is easy to check that the left-hand side of (5.20) satisfies the above conditions. Here we will only prove that the solution to (5.20) is given via (5.21), which shows the relation between the minor determinants of products (5.20) and partition functions of paths.

From the graphical representation of the minor determinants of $H(\boldsymbol{x}^1)H(\boldsymbol{x}^2)\cdots H(\boldsymbol{x}^n)$, we have that
$$\tau^i_j = \det\big(H(\boldsymbol{x}^1)H(\boldsymbol{x}^2)\cdots H(\boldsymbol{x}^n)\big)^{1,\dots,i}_{j-i+1,\dots,j}.$$
But by (5.20) this is equal to $\det\big(H_k(\boldsymbol{y}^k)H_{k-1}(\boldsymbol{y}^{k-1})\cdots H_1(\boldsymbol{y}^1)\big)^{1,\dots,i}_{j-i+1,\dots,j}$ and again from the graphical representation we see that this equals
$$\sum_{\gamma_1,\dots,\gamma_r} \mathsf{wt}(\gamma_1)\cdots\mathsf{wt}(\gamma_r) \equiv \sum_{\gamma_1,\dots,\gamma_r} \tag{5.23}$$


where the summation is over non-intersecting, up-right paths on the trapezoidal lattice with weights $\boldsymbol{y}$, that start from vertices $(1,1), \dots, (i,i)$ in the lower-left border of the lattice and go to vertices $(1, j-i+1), \dots, (1,j)$ at the upper border. As seen in the figure in relation (5.23), there is only one such i-tuple of paths with total weight $\prod_{1\le a\le i\,,\, a\le b\le j} y^a_b$. Writing similarly $\tau^{i-1}_j, \tau^{i-1}_{j-1}, \tau^i_{j-1}$ we see that
$$\frac{\tau^i_i}{\tau^{i-1}_i} = y^i_i, \quad \text{and for } i<j \quad \frac{\tau^i_j\,\tau^{i-1}_{j-1}}{\tau^{i-1}_j\,\tau^i_{j-1}} = y^i_j.$$
□

Let us now describe how geometric RSK can be encoded in this matrix formulation. We will make reference to the following diagram (refer to relation (5.3) for a reminder on the notation):

$$\begin{array}{ccccccccc}
 & \boldsymbol{x}^1 & & \boldsymbol{x}^2 =: \boldsymbol{x}^{2,1} & & \boldsymbol{x}^3 =: \boldsymbol{x}^{3,1} & & & \\
\emptyset & \longrightarrow & \boldsymbol{y}^{1,1} & \longrightarrow & \boldsymbol{y}^{2,1} & \longrightarrow & \boldsymbol{y}^{3,1} & \cdots \\
 & & & \boldsymbol{x}^{2,2} & & \boldsymbol{x}^{3,2} & & \\
 & & \emptyset & \longrightarrow & \boldsymbol{y}^{2,2} & \longrightarrow & \boldsymbol{y}^{3,2} & \cdots \\
 & & & & & \boldsymbol{x}^{3,3} & & \\
 & & & & \emptyset & \longrightarrow & \boldsymbol{y}^{3,3} & \cdots \\
 & & & & & & \emptyset & \cdots
\end{array}$$

where $\boldsymbol{x}^i = (x^i_1, \dots, x^i_N)$ for $i \ge 1$, is a sequence of words, which are successively row inserted via geometric RSK. We should keep in mind the useful identification of entries x^i_j with number of letters ‘j’ in a ‘*word* ’ $\boldsymbol{x}^i$ as in (3.20).

Let us now describe geometric RSK in this matrix language translating essentially from the language of RSK as described in Section **??**. For this reason we will be using the terms *tableau, row insertion, word* etc.

Initially, we have an empty tableau $\emptyset$ to which we insert the first word $\boldsymbol{x}^1$ as $\emptyset \overset{\boldsymbol{x}^1}{\underset{\emptyset}{\longrightarrow}} \boldsymbol{y}^{1,1}$. Of course, in this situation the output tableau will only have one row $\boldsymbol{y}^{1,1}$ and $\boldsymbol{y}^{1,1} = \boldsymbol{x}^1$, which we trivially encode via the matrix equation

$$H(\boldsymbol{x}^1) = H(\boldsymbol{y}^{1,1}). \tag{5.24}$$

Next, to the single-row tableau $\boldsymbol{y}^{1,1}$ we insert $\boldsymbol{x}^2$ as $\boldsymbol{y}^{1,1} \overset{\boldsymbol{x}^2}{\underset{\boldsymbol{x}^{2,2}}{\longrightarrow}} \boldsymbol{y}^{2,1}$. Here, $\boldsymbol{y}^{2,1}$ corresponds to the updated first row of the new tableau and $\boldsymbol{y}^{2,2} = \boldsymbol{x}^{2,2}$ to its second row. We notice that $\boldsymbol{x}^{2,2}$ is the word consisting of the dropdown letters after the insertion of $\boldsymbol{x}^2$ into $\boldsymbol{y}^{1,1}$, which are then inserted into the empty second row and this is the reason why the second row of the updated tableau $\boldsymbol{y}^{2,2}$ equals $\boldsymbol{x}^2$. This second set of insertions can be be encoded via a matrix equation, which is derived by multiplying (5.24) on the right by $H(\boldsymbol{x}^2)$ and using (5.20) to obtain

$$H(\boldsymbol{x}^1)H(\boldsymbol{x}^2) = H_1(\boldsymbol{y}^{1,1})H(\boldsymbol{x}^2) \overset{(5.20)}{=} H_2(\boldsymbol{y}^{2,2})H_1(\boldsymbol{y}^{2,1}). \tag{5.25}$$

The fact that this matrix multiplication and the output variables $\boldsymbol{y}^{2,2}, \boldsymbol{y}^{2,1}$ give the output of geometric row insertion with input $\boldsymbol{x}^1, \boldsymbol{x}^2$ is a consequence of Definition 5.7, its matrix reformulation (5.16), (5.20) and finally Theorem 5.8.

In a similar way, we encode the third group of row insertions :

$\boldsymbol{y}^{2,1} \overset{\boldsymbol{x}^3}{\underset{\boldsymbol{x}^{3,2}}{\longrightarrow}} \boldsymbol{y}^{3,1}$ [word $\boldsymbol{x}^3$ is row inserted to the first line $\boldsymbol{y}^{2,1}$ of the current tableau],

$\boldsymbol{y}^{2,2} \overset{\boldsymbol{x}^{3,2}}{\underset{\boldsymbol{x}^{3,3}}{\longrightarrow}} \boldsymbol{y}^{3,2}$ [word $\boldsymbol{x}^{3,2}$ formed by the dropped down letters are row inserted to the second line $\boldsymbol{y}^{2,2}$ of the current tableau],

$\emptyset \overset{\boldsymbol{x}^{3,3}}{\underset{\emptyset}{\longrightarrow}} \boldsymbol{y}^{3,3}$ [the dropdown letters from the previous insertion form the new row $\boldsymbol{y}^{3,3}$],

via multiplying on the right (5.25) by $H(\boldsymbol{x}^3)$ and using successively relation (5.20) as

$$\begin{aligned} H(\boldsymbol{x}^1)H(\boldsymbol{x}^2)H(\boldsymbol{x}^3) &= H_2(\boldsymbol{y}^{2,2})H_1(\boldsymbol{y}^{2,1})H(\boldsymbol{x}^3) \overset{(5.20)}{=} H_2(\boldsymbol{y}^{2,2})H_2(\boldsymbol{x}^{3,2})H_1(\boldsymbol{y}^{3,1}) \\ &\overset{(5.20)}{=} H_3(\boldsymbol{x}^{3,3})H_2(\boldsymbol{y}^{3,2})H_1(\boldsymbol{y}^{3,1}) = H_3(\boldsymbol{y}^{3,3})H_2(\boldsymbol{y}^{3,2})H_1(\boldsymbol{y}^{3,1}) \end{aligned}$$

We point out that the second equality above is a matrix representation of the diagram $\boldsymbol{y}^{2,1} \overset{\boldsymbol{x}^3}{\underset{\boldsymbol{x}^{3,2}}{\longrightarrow}} \boldsymbol{y}^{3,1}$, the third equality of the diagram $\boldsymbol{y}^{2,2} \overset{\boldsymbol{x}^{3,2}}{\underset{\boldsymbol{x}^{3,3}}{\longrightarrow}} \boldsymbol{y}^{3,2}$ and the fourth of the (trivial) diagram $\emptyset \overset{\boldsymbol{x}^{3,3}}{\underset{\emptyset}{\longrightarrow}} \boldsymbol{y}^{3,3}$.

This procedure continues during the first N insertions at which stage the resulting tableau will have the full depth of N rows. After that, no additional rows will be created in the subsequent tableaux and the process continues as follows

$$
\begin{array}{ccccccccc}
 & \boldsymbol{x}^{N+1} =: \boldsymbol{x}^{N+1,1} & & \boldsymbol{x}^{N+2} =: \boldsymbol{x}^{N+2,1} & & \boldsymbol{x}^{N+3} =: \boldsymbol{x}^{N+3,1} & \\
\boldsymbol{y}^{N,1} & \longrightarrow & \boldsymbol{y}^{N+1,1} & \longrightarrow & \boldsymbol{y}^{N+2,1} & \longrightarrow & \cdots \\
 & \boldsymbol{x}^{N+1,2} & & \boldsymbol{x}^{N+2,2} & & \boldsymbol{x}^{N+3,2} & \\
\boldsymbol{y}^{N,2} & \longrightarrow & \boldsymbol{y}^{N+1,2} & \longrightarrow & \boldsymbol{y}^{N+2,2} & \longrightarrow & \cdots \\
 & \boldsymbol{x}^{N+1,3} & & \boldsymbol{x}^{N+2,3} & & \boldsymbol{x}^{N+3,3} & \\
\boldsymbol{y}^{N,3} & \longrightarrow & \boldsymbol{y}^{N+1,3} & \longrightarrow & \boldsymbol{y}^{N+2,3} & \longrightarrow & \cdots \\
 & \boldsymbol{x}^{N+1,4} & & \boldsymbol{x}^{N+2,4} & & \boldsymbol{x}^{N+3,4} & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & \boldsymbol{x}^{N+1,N} & & \boldsymbol{x}^{N+2,N} & & \boldsymbol{x}^{N+3,N} & \\
\boldsymbol{y}^{N,N} & \longrightarrow & \boldsymbol{y}^{N+1,N} & \longrightarrow & \boldsymbol{y}^{N+2,N} & \longrightarrow & \cdots \\
 & \emptyset & & \emptyset & & \emptyset &
\end{array}
$$

Overall, the above, two-step procedure is encoded via the matrix equation

$$
\begin{cases}
H(\boldsymbol{x}^1)H(\boldsymbol{x}^2)\cdots H(\boldsymbol{x}^n) = H_n(\boldsymbol{y}^{n,n})H_{n-1}(\boldsymbol{y}^{n,n-1})\cdots H_1(\boldsymbol{y}^{n,1}), & \text{if } n \le N \\
H(\boldsymbol{x}^1)H(\boldsymbol{x}^2)\cdots H(\boldsymbol{x}^n) = H_N(\boldsymbol{y}^{n,N})H_{N-1}(\boldsymbol{y}^{n,N-1})\cdots H_1(\boldsymbol{y}^{n,1}), & \text{if } n \ge N
\end{cases}
\tag{5.26}
$$

We have seen how to encode (geometric) row insertion in a matrix formulation, thus producing the P tableau of geometric RSK. We can state the geometric RSK as a one-to-one correspondence between input matrices $\mathsf{X} = (x^i_j \colon 1 \le i \le n,\, 1 \le j \le N)$ and two sets of variables $P := (p^i_j \colon 1 \le i \le N \wedge n,\, i \le j \le N)$ and $Q := (q^i_j \colon 1 \le i \le N \wedge n,\, i \le j \le n)$. These will be the analogues of the P and Q tableaux in the standard RSK correspondence. For a full proof of this theorem (in particular the reconstruction of X from (P, Q)), we refer to [NY04], Section 3 and Theorem 3.8.

Theorem 5.9. *Consider a matrix* $\mathsf{X} := (x^i_j \colon 1 \le i \le n,\, 1 \le j \le N)$ *with nonnegative entries and denote by* $(\boldsymbol{x}^1, ..., \boldsymbol{x}^n)$ *its rows and by* $(\boldsymbol{x}_1, ..., \boldsymbol{x}_N)$ *its columns. Then there exists a one-to-one correspondence between* X *and a set of variables* $\boldsymbol{p}^i := (p^i_i, ..., p^i_N)$ *for* $i = 1, ..., \min(n, N)$ *and* $\boldsymbol{q}^i = (q^i_i, ..., q^i_n)$ *for* $i = 1, ..., \min(n, N)$*, which are uniquely determined via equations*

$$H(\boldsymbol{x}^1)H(\boldsymbol{x}^2)\cdots H(\boldsymbol{x}^n) = H_k(\boldsymbol{p}^k)H_{k-1}(\boldsymbol{p}^{k-1})\cdots H_1(\boldsymbol{p}^1), \qquad k = \min(n, N), \tag{5.27}$$

$$H(\boldsymbol{x}_1)H(\boldsymbol{x}_2)\cdots H(\boldsymbol{x}_N) = H_k(\boldsymbol{q}^k)H_{k-1}(\boldsymbol{q}^{k-1})\cdots H_1(\boldsymbol{q}^1), \qquad k = \min(n, N), \tag{5.28}$$

Variables (p^i_j) *and* (q^i_j) *are given in terms of the input variables* (x^i_j) *via relations* (5.20) *and* (5.22).

An immediate consequence of the formulation of gRSK as in Theorem 5.9 and in particular the matrix equations (5.27), (5.28) is that if the input matrix X is symmetric then the P and Q tableaux of gRSK are equal.

Geometric RSK on (geometric) Gelfand-Tsetlin patterns. It is useful to put geometric RSK and Theorem 5.9 under a Gelfand-Tsetlin framework. To this end, if $\boldsymbol{p}^1, ..., \boldsymbol{p}^k$ and $\boldsymbol{q}^1, ..., \boldsymbol{q}^k$ are as in (5.27), (5.28), set

$$
\begin{aligned}
z^i_j &:= p^j_j\, p^j_{j+1} \cdots p^j_{i-1}\, p^j_i \quad \text{for} \quad 1 \le j \le i \le N, \text{ and } j \le n \wedge N, \\
(z^i_j)' &:= q^j_j\, q^j_{j+1} \cdots q^j_{i-1}\, q^j_i \quad \text{for} \quad 1 \le j \le i \le n, \text{ and } j \le n \wedge N.
\end{aligned}
$$

Then Theorem 5.9 establishes a bijection between matrices $\mathsf{X} := (x^i_j \colon 1 \le i \le n,\, 1 \le j \le N)$ with nonnegative entries and a pair $(\mathsf{Z}, \mathsf{Z}') = \mathsf{gRSK}(\mathsf{X})$. We will call the arrays $\mathsf{Z} = (z^i_j \colon 1 \le j \le i \le N, j \le n \wedge N)$ and $\mathsf{Z}' = ((z^i_j)' \colon 1 \le j \le i \le n, j \le n \wedge N)$, **geometric Gelfand-Tsetlin patterns**, even though in general they do not satisfy the interlacing constraints $z^{i+1}_{j+1} \le z^i_j \le z^{i+1}_j$ (however, they do degenerate to genuine Gelfand-Tsetlin patterns in the combinatorial limit described in the next paragraph). For short, we will often denote geometric Gelfand-Tsetlin patterns by gGT, while at other times, when it is clear from the context, we may omit the adjective "goemetric".

Bearing in mind property (5.21) we obtain that variables z^i_j are given in terms of ratios of partition functions

$$z^i_j = \frac{\tau^j_i}{\tau^j_{i-1}} = \frac{\sum_{\pi_1,\dots,\pi_j} \mathsf{wt}(\pi_1)\cdots\mathsf{wt}(\pi_j)}{\sum_{\pi_1,\dots,\pi_{j-1}} \mathsf{wt}(\pi_1)\cdots\mathsf{wt}(\pi_{j-1})}, \tag{5.29}$$

where the sum in the numerator is over directed, non-intersecting paths along entries of X starting from $(1,1),\dots,(1,j)$ and ending at $(n,i-j+1),\dots,(n,i)$, respectively, and the denominator sum is over directed, non-intersecting paths starting from $(1,1),\dots,(1,j-1)$ and ending at $(n,i-j+2),\dots,(n,i)$, respectively. In particular,

$$z^N_1 = \sum_{\pi\colon (1,1)\to(n,N)} \mathsf{wt}(\pi) = \sum_{\pi\colon (1,1)\to(n,N)} \prod_{(a,b)\in\pi} x^a_b, \tag{5.30}$$

which defines the **partition function** of the **directed polymer model**.

Passage to standard (combinatorial) RSK setting. We will now see how the geometric RSK framework degenerates to the standard RSK framework and how in this way we can obtain via Theorem 5.9 both Schensted's and Greene's theorems as well as the links between RSK and last passage percolation alluded to in (5.9) and (5.10).

Replacing in (5.12) variables $\xi_k, \widetilde{\xi}_k, a_k, b_k$ by $e^{\xi_k/\varepsilon}, e^{\widetilde{\xi}_k/\varepsilon}, e^{a_k/\varepsilon}, e^{b_k/\varepsilon}$, taking the logarithm on both sides of each relation therein and multiplying by ε, the set of equations (5.12) may be written as

$$\begin{cases} \widetilde{\xi}_i = \xi_i + a_i, \\ \widetilde{\xi}_k = a_k + \varepsilon \log\left(e^{\widetilde{\xi}_{k-1}/\varepsilon} + e^{\xi_k/\varepsilon}\right), & i+1 \le k \le N \\ b_k = a_k + (\xi_k - \xi_{k-1}) - (\widetilde{\xi}_k - \widetilde{\xi}_{k-1}), & i+1 \le k \le N. \end{cases} \tag{5.31}$$

Taking now the limit $\varepsilon \to 0$ these reduce to the piecewise linear transformations (5.8) defining the standard RSK correspondence. Replacing also the variables x^i_j, p^i_j, q^i_j in Theorem 5.9 by $e^{x^i_j/\varepsilon}, e^{p^i_j/\varepsilon}, e^{q^i_j/\varepsilon}$ we obtain in the limit $\varepsilon \to 0$ the RSK correspondence, in the sense that variables (p^i_j) and (q^i_j) encode the P and Q tableaux of the standard RSK.

In particular, the solution to the degeneration, as $\varepsilon \to 0$, of problem (5.27) is given via the degeneration of relations (5.21), (5.22) as:

$$p^i_i = \sigma^i_i - \sigma^{i-1}_i \quad \text{and} \quad p^i_j = \sigma^i_j + \sigma^{i-1}_{j-1} - \sigma^{i-1}_j - \sigma^i_{j-1} \quad \text{for} \quad i<j,$$

$$\text{with} \qquad \sigma^i_j := \max_{\pi_1,\dots,\pi_i \in \Pi^{1,\dots,i}_{j-i+1,\dots,j}} \sum_{k=1}^i \mathsf{wt}(\pi_k)$$

being last passage percolation functionals corresponding to ensembles of i non-intersecting, down-right paths $\pi_1,\dots,\pi_i$, starting from $(1,1),\dots,(1,i)$ and ending at $(n,j-i+1),\dots,(n,j)$, respectively. The weight of a path π_r in this case is $\mathsf{wt}(\pi_r) := \sum_{(a,b)\in\pi_r} x^a_b$.

Passing to the Gelfand-Tsetlin variables, we set

$$z^i_j := p^j_j + p^j_{j+1} + \cdots + p^j_{i-1} + p^j_i = \sigma^j_i - \sigma^j_{i-1}$$

for $1 \le j \le i \le N$, and $j \le n \wedge N$. From this we get that

$$z^N_1 + \cdots z^N_j := \sigma^j_N,$$

which in the case $j=1$ is Schensted's theorem.

6. Pieri rules and integrable Markov processes.

6.1. Pieri Rule. Pieri rule is the rule of how to expand $s_\nu h_n$ and $s_\nu e_n$, for ν a partition, in terms of Schur functions. Notice that the function $s_\nu h_n$ and $s_\nu e_n$ are symmetric and since Schur functions are a basis for the ring of symmetric functions, such an expansion exists and the question is to determine the coefficients. We will see that the Pieri rule has probabilistic interpretations, in particular showing that the evolution of the shape of Young tableaux through RSK is Markovian.

The Pieri rule is a particular case of the expansion

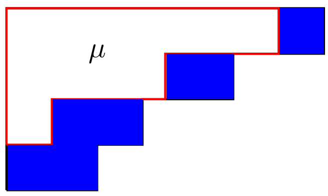
$$s_\mu s_\nu = \sum_\lambda c^\lambda_{\mu\nu} s_\lambda,$$

with the coefficients $c^\lambda_{\mu\nu}$ called the *Littlewood-Richardson* coefficients. These coefficients have a beautiful and deep combinatorial interpretation that goes through *Schtzenberger* involution or jeu-de-taquin (the *" teasing game"*). Littlewood-Richardson coefficient also have a wide range of application in algebraic geometry (Schubert calculus, Grassmanians etc.) and representations theory. Unfortunately, I don't think we will have time to go through all these...but you can take a look at the appendix of Stanley's book.

Let us just expose Pieri's rules and the links to probability. We need to start with some definitions.

Definition 6.1. *Let λ, μ be partitions such that $\mu \subset \lambda$, i.e. the Young diagrams with shape μ lies inside the skew diagram with shape λ. We define the* **skew partition** *λ/μ to be the skew-shaped diagram that remains after removing the boxes of μ from the Young diagram corresponding to λ. That is, $\lambda/mu = (\lambda_1 - \mu_1, \lambda_2 - \mu_2, ...)$.*

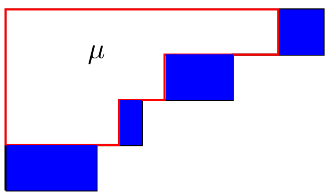
An example of a skew partition and its skew Young diagram depiction is this:

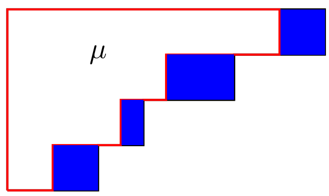
$\lambda =$ 

In this figure the blue shaded area is the skew partition λ/μ.

Definition 6.2. *A skew partition λ/μ is called a* **horizontal strip** *if it contains no two boxes on top of each other.*

Two (typical) examples of horizontal skew-partitions are :

$\lambda =$ 

or $\lambda =$ 

(6.1)

where again the blue are depicts λ/μ.

Definition 6.3. *A skew partition λ/μ is called a* **vertical strip** *if it contains no two boxes at the same horizontal level.*

We can now state the first Pieri rule and give a bijective proof of it. For an inner product proof look at Stanley's book.

Theorem 6.4 (Pieri rule). *Let ν be a partition and let s_ν be the corresponding Schur functions. Let also s_n be the Schur function that corresponds to the partition (n), which consists of one part with n boxes. We then have the identity*

$$s_\nu s_n = \sum_{\lambda:\ \lambda/\nu \text{ is a horizontal strip of size } n} s_\lambda.$$

Since $s_n = h_n$ (explain why) the same identity holds for the product $s_\nu h_n$.

Proof. The bijective proof is another nice application of the RSK algorithm.

We start with the combinatorial definition of the Schur functions and write

$$\begin{aligned} s_\nu s_n &= \sum_{T \in SSYT(\nu)} x_1^{\alpha_1(T)} x_2^{\alpha_2(T)} \cdots \sum_{T' \in SSYT((n))} x_1^{\alpha_1(T')} x_2^{\alpha_2(T')} \cdots \\ &= \sum_{T \in SSYT(\nu), T' \in SSYT((n))} x_1^{\alpha_1(T)+\alpha_1(T')} x_2^{\alpha_2(T)+\alpha_2(T')} \cdots \end{aligned} \tag{6.2}$$

Setting $\beta_i := \alpha_i(T) + \alpha_i(T')$ for $i \geq 1$, Pieri's rule amounts to finding a bijection between, on the one hand, the set of pairs (T, T') of $T \in SSYT(\nu)$ and $T' \in SSYT((n))$ with $type(T) = (\alpha_1(T), \alpha_2(T), ...)$ and $type(T') = (\alpha_1(T'), \alpha_2(T'), ...)$ and, on the other hand, the union $\bigcup_{\lambda:\ \lambda/\nu \text{ is a horizontal strip of size n}} \mathcal{T}_\lambda^\beta$ of the set of $SSYT(\lambda)$ with λ/ν is a horizontal strip of size n and type $\beta = \alpha(T) + \alpha(T')$.

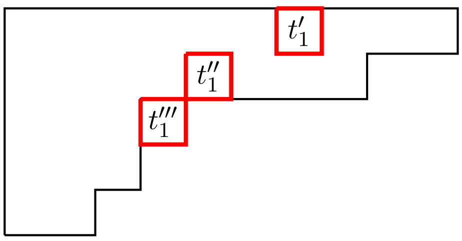
To do so, let encode a $T' \in SSYT((n))$ via the sequence of numbers $t'_1 t'_2 \cdots t'_n$ as they appear in its boxes. Let us now start inserting letters $t'_1, t'_2, \ldots, t'_n$ in this sequence into tableau T via RSK (row insertion), i.e.

$$\big(((T \leftarrow t'_1) \leftarrow t'_2) \leftarrow \cdots\big) \leftarrow t'_n. \tag{6.3}$$

The fact that the resulting tableau in (6.3) will have an additional horizontal strip added to ν will be a consequence of the fact that $t'_1 \leq t'_2 \leq \cdots \leq t'_n$ and the following consideration: Let us look at the insertion $T \leftarrow t'_1$:

(i) either t'_1 will be inserted at the end of the first row and the insertion will finish or

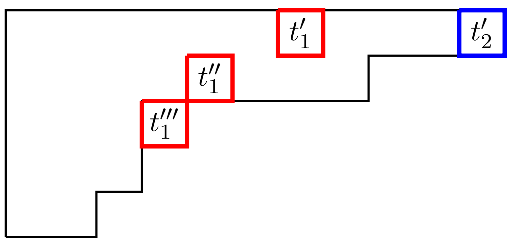
(ii) it will bump an entry of the first row, denote it by t''_2, which will then be inserted in the second row bumping another entry t'''_1 etc. initiating a cascade until eventually a new box is created at the end of a row. For example:

$T \leftarrow t'_1 =$ 

Notice that the red path will move towards the left, *weakly*, as we move downwards. Let us see this by seeing why the placement of t''_1 in the second row should be (weakly) in the left of the placement of t'_1 in the first row: t'_1 took the place of t''_1, which means that the box just below that position would contain a number a_1 strictly larger than t''_1. So when t''_1 gets to be inserted in the second row, it cannot go beyond the box of a_1.

Let us now insert t'_2. We know that $t'_1 \leq t'_2$, so the box where t'_2 is inserted in the first row will be in the right of that of t'_1. There are two possibilities:

(i) It may be inserted at the end of the row in which case the insertion ends and the resulting tableau after both t'_1 and t'_2 is clearly obtained from T by adding a horizontal strip of two boxes (occupied by t'_2 and t'''_1) like in the picture:

$(T \leftarrow t'_1) \leftarrow t'_2 =$ 

(ii) Or it might bump an entry t''_2, which will then be inserted into the second row. Let's see what happens in this case:

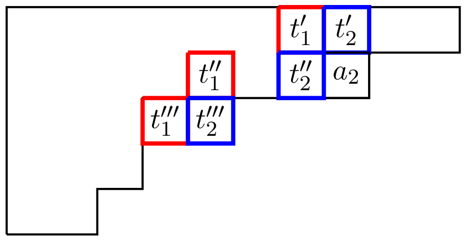
Let a_2 be the entry of the box below that of the box used to be occupied by t_2'' (and which now is occupied by t_2'. We have that

$$t_1' \leq t_2' < t_2'' < a_2$$

with the second inequality as a consequence of the fact that t_2' bumped t_2''. Moreover, we have that $t_1'' \leq t_2''$ since t_1'' used to occupy in the T tableau the box of t_1', which is in the left of that that t_2'' used to occupy before it was bumped. This inequality implies that, when inserted in the second row, t_2'' will be strictly to the right t_1'' but also weakly to the left of a_2. This process will continue, it t_1''' was not appended at the end of the row. Let's see what happens if t_1''' was appended at the end of the next row.

Before being bumped, t_1''' was occupying the box which is currently occupied by t_1'', hence $t_1''' \leq t_2'''$ since the box of t_2''' (which will occupied by t_2'' was on the right of that occupied by t_1''' (before being bumped). But since t_2''' is larger than t_1''', when inserted in the row that t_1''' currently lies, it will bypass it. Since we have assumed that t_1''' is currently at the end of the row, t_2''' will go at the end of the same row, next to t_1'''. See the picture:

$$\big(T \leftarrow t_1'\big) \leftarrow t_2' =$$



It is now clear that $\big(T \leftarrow t_1'\big) \leftarrow t_2'$ is obtained by T by adding a horizontal line, consisting of the boxes occupied by t_1''' and t_2'''.

The process continues in the same way, therefore $\big(((T \leftarrow t_1') \leftarrow t_2') \leftarrow \cdots\big) \leftarrow t_n'$ is obtained from T by adding a horizontal strip. $\square$

Exercise 20. *Prove the dual Pieri rule that*

$$s_\nu s_{1^n} = s_\nu e_n = \sum_{\lambda:\ \lambda/\nu \text{ is a vertical strip of size } n} s_\lambda.$$

Use the homomorphism ω and the fact[†] *that $\omega(s_\lambda) = s_{\lambda'}$, where for a partition λ, λ' is the partition whose rows (in the Young diagram representation) are the columns of λ.*

6.2. Solvable particle systems. We will recast RSK into a two-dimensional (triangular array) interacting particle system, which couples two one-dimensional interacting particle systems. The one $1d$ particle system will be the right diagonal of the ϑ pattern and the other its bottom row (which is also the shape of the Young tableau). Both systems will be Markovian but while the Markovianity of the diagonal will be obvious the Morkovianity of the bottom row (or, equivalently, the evolution of the shape of the Young tableau of RSK) is very non obvious. The Markovian evolution of the shape will be related to the Pieri rule.

To start, we will need to recall the Gelfand-Tsetlin representation of Young tableaux from Section 3.7. This is a triangular array of numbers $(z^i_j)_{1\leq j\leq i\leq N}$ which interlace, meaning that

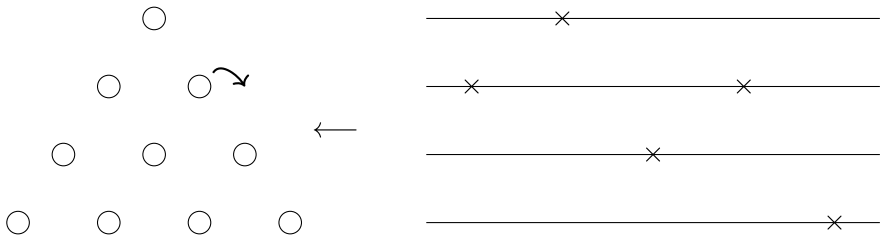
$$z^{i+1}_{j+1} \leq z^i_j \leq z^{i+1}_j. \tag{6.4}$$

We also recall the encoding of Young tableaux that the Gelfand-Tsetlin patterns offer, which is that

$$z^i_j \text{ is the total number of boxes in row } j\text{, which contain numbers up to } i$$

[†]if you want understand why this fact holds, look at Stanley's book, Section 7.14. This fact used the *dual* RSK, which we didn't cover

A matrix, which serves as the input of RSK will be considered as encoding a sequence of signals which trigger particle motion. Pictorially we can image



(6.5)

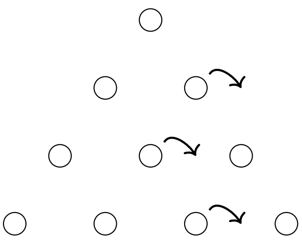
where the lines in the left can be thought of as the input array

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

The first pulse being on line 2, will trigger, via row insertion, a move of particle z_2^1 and this move will trickle downwards in teh gelfand-Tsetlin pattern in a particular way, which is consisent with the row insertion. The rule is the following:

- a particle z_j^i can only move to the right by 1.
  - if $z_j^i < z_j^{i+1}$, then the motion will propagate to particle z_{j+1}^{i+1}, which will move to the right by 1 step,
  - if $z_j^i = z_j^{i+1}$, then the motion will propagate to particle z_j^{i+1}, which will move to the right by 1 step,
- the process continues in the above fashion until the motion reached the bottom room of the Gelfand-Tsetlin pattern.

For example, a particle motion of the form



indicates that :

1. there was a "2" that was row-inserted, which causes z_1^2 to jump to the right by 1,
2. before its jump, z_1^2 was strictly less than z_1^3 and so z_1^2 *"pulls"* particle z_2^3 to the right by 1
3. before its jump, z_2^3 was equal to z_3^4 and so it pushes particle z_3^4 to the right by 1.

To see why this motion is consistent with row insertion, we invite the reader to work out the particle representation of the row insertion of the word $(2, 3)$ into the Young tableau / Gelfand-tsetlin pattern:

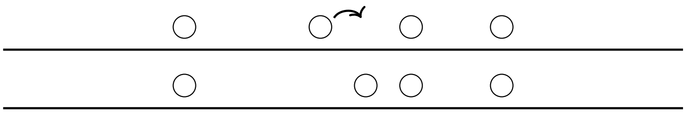
$$\begin{array}{|c|c|c|c|c|c|} \hline 1 & 1 & 2 & 2 & 3 & 4 \\ \hline 2 & 2 & 3 & 4 \\ \cline{1-4} 3 & 4 \\ \cline{1-2} 2 \\ \cline{1-1} \end{array} \quad \equiv \quad \begin{array}{ccccccc} & & & 2 & & & \\ & & 2 & & 4 & & \\ & 1 & & 3 & & 5 & \\ 1 & & 2 & & 4 & & 6 \end{array} \quad \longleftarrow \ (2,3)$$

We can view the right diagonal $(z_1^i : i = 1, ..., N)$ of the Gelfand-Tsetlin pattern as a one-dimensional particle system, called *push-TASEP* (TASEP stands for Totally Asymmetric Exclusion Process). Indeed, setting

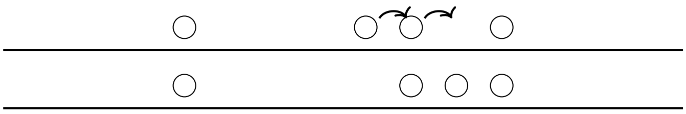
$$x_i := z_1^i - N + i, \qquad \text{for } i = 1, ..., N$$

and given that $z_1^i \leq z_1^{i+1}$, we have that $x_i < x_{i+1}$ and so we can map the diagonal to a linear system of particles where no two particles occupy the same location. For example, the diagonal $(2, 4, 5, 6)$ of the above gelfand-Tsetlin pattern maps to particles

If a particle that attempts to jump is blocked by the particle in front of it, then it pushes that particle one step ahead. Recall that the location of the first particle from the left is $x_1 = z_1^1 - N + 1$, the location of the second particle is $x_2 = z_2^1 - N + 2$ and so on... If, now, a 2 is inserted in the tableau, the second particle (from the left) will want to jump to the right as



If, now, another 2 is inserted, again the second particle (from the left) will want to jump to the right but in this case it is blocked by the third particle and so it will push it also to the right as



It is clear that push-TASEP is Markovian because knowing the randomness (ie which particle is prompted to jump) and the current configuration, we can determine the next configuration. Moreover, it is easy to see that the whole Gelfand-Tsetlin configuration $(z_j^i : 1 \leq j \leq i \leq N)$ is Markovian. However, it is not obvious that the evolution of the bottom row $z^N(n) := (z_i^N(n) : i = 1, ..., N)$ is Markovian on its own, that is, that knowing the random input $w^n = 1^{w_1^n} 2^{w_2^n} \cdots N^{w_N^n}$ and the configuration $z^N(n)$, we can determine (the statistics of) the next configuration without any further information or the history of $(z^N(t) : t < n)$. We will determine that this is the case when the distributions of (W_j^i) are independent geometric with parameters $p_i q_j$. This consideration is important because it will turn out the Markovian dynamics of the bottom row are recognised as (analogues) of Markovian dynamics coming from Random Matrix Theory and so the statistics of the largest particle of push-TASEP are identified as the (analogues of the) statistics of largest eignevalue a random matrix ensemble.

The bottom row of a Gelfand-Tsetlin pattern can be viewed as a function of the Gelfand-Tsetlin pattern, in particular $(z_i^N : i = 1, ..., N)$ is the projection of $Z := (z_j^i : 1 \leq j \leq i \leq N)$ onto its bottom row. Given that Z evolves as a Markov process, the question is a particular case of the following question :

when is a function of a Markov process a Markov process itself ?

This question naturally brings us to the point that we should do a quick recap of Markov process and prove the main theorem on Markov functions, which is known as the Pitman-Rogers theorem.

6.3. A quick reminder / introduction of Markov processes and the Pitman-Rogers theorem. A Markov process is a stochastic process $(X_n)_{n \geq 0}$ such that the future only depends on the present and not the past. In particular

$$\mathbb{P}(X_{n+1} = x_{n+1} \,|\, X_n, ..., X_0) = \mathbb{P}(X_{n+1} = x_{n+1} \,|\, X_n).$$

Recall that for two variables X, Y

$$\mathbb{P}(Y = y \,|\, X = x) = \frac{\mathbb{P}(Y = y, X = x)}{\mathbb{P}(X = x)},$$

by Bayes rule.

An important notion in Markov processes is the **transition probability matrix**. This is defined as

$$Q_n(y, x) := \mathbb{P}(X_{n+1} = y \,| X_n = x), \qquad \text{for } x, y \text{ in the state space } \mathcal{S} \text{ of } (X_n)$$

A Markov process is called *(time) homogeneous*, if the transition matrix Q_n is independent of the time n, i.e. $Q_n(x,y) = Q(x,y)$ for all $x, y \in \mathcal{S}$.

Another thing to note is that $Q_n(x,y)$ are what is called *stochastic matrices*, i.e. $\sum_{y\in\mathcal{S}} Q_n(x,y) = 1$ for all $x \in \mathcal{S}$.

Markov processes are essentially the only processes for which a complete theory exists. If a process is not Markov, ie it has memory, then its study can be formidable. Some such examples are the self-avoiding walk, reinforced random walk etc.

Another interesting point is the relation between Markov processes and differential operators. To understand the link, we can look at the Simple Random Walk (SRW) on $\mathbb{Z}^d$, that is the Markov process such that

$$Q(x,y) = \frac{1}{2d}\mathbb{1}_{|x-y|_1=1}$$

. Viewed as an operator Q acts on functions as

$$(Qf)(x) = \sum_{y\in\mathcal{S}} Q(x,y)f(y),$$

and in the special case of SRW

$$(Q^{\mathrm{SRW}} f)(x) = \frac{1}{2d} \sum_{y\in\mathbb{Z}^d :\, |y-x|_1=1} f(y).$$

The operator $Q - I$, where I is the identity operator is also distinguished and often it is call the *generator* of the Markov process. It is important as it converts Markov process to *martingales* - another distinguished class of stochastic processes. In the case of SRW we have that

$$\big((Q^{\mathrm{SRW}} - I)f\big)(x) = \frac{1}{2d} \sum_{y\in\mathbb{Z}^d :\, |y-x|_1=1} \big(f(y) - f(x)\big)$$

and this is the *discrete Laplacian* on $\mathbb{Z}^d$, i.e. in a suitable spatial scaling limit it converges to $\frac{1}{2}\Delta$.

All in all, having Markov processes is important as we can use the theory. In general, a function of a Markov process is not Markov (can you think of such a situation ?), so knowing when the Markovianity is maintained is important. A criterion is provided by the following theorem:

Theorem 6.5 (Pitman-Rogers [RP81]**).** *Consider a discrete time Markov process $Z(\cdot)$ on a measurable space $(\mathcal{Z},\mu)$ with transition probability kernel Π and a measurable function $\Phi : \mathcal{Z} \to \mathcal{X}$, with $\mathcal{X}$ a measurable space. Assume that there exists a kernel $\mathsf{P}(\cdot,\cdot)\colon \mathcal{X}\times\mathcal{X} \to \mathbb{R}$ such that for almost every $x \in \mathcal{X}$, $\mathsf{P}(x,\cdot)$ is a probability measure and a kernel $\mathsf{K}(\cdot,\cdot)\colon \mathcal{X}\times\mathcal{Z} \to \mathbb{R}$ satisfying:*

(i) for all $x \in \mathcal{X}$, $\mathsf{K}(x, \Phi^{-1}(x)) = 1$,

(ii) the inter-twinning relation $\mathsf{K}\Pi = \mathsf{P}\mathsf{K}$ holds.

If, for arbitrary $x \in \mathcal{X}$, the initial distribution of the Markov process $Z(\cdot)$ is $\mathsf{K}(x,\cdot)/\int_{\mathcal{Z}} \mathsf{K}(x,z)\mu(\mathrm{d}z)$, then it holds that

(i) The process $X_n = \Phi(Z_t)$ is Markov with respect to its own filtration $\mathcal{X}_n := \sigma\{X_s \colon s \le n\}$ with transition probability kernel P and initial condition $X_0 = x$,

(ii) For all $x \in \mathcal{X}$ and all bounded Borel functions f on $\mathcal{Z}$,

$$\mathbb{E}\big[f(Z_t)\,\big|\, X_s,\, s<t,\, \mathsf{X}_t = x\big] = (\mathsf{K}f)(x).$$

Proof. For simplicity, let us assume that μ is the counting measure and that $\sum_z \mathsf{K}(x,z) = 1$. Let us compute the probability

$$\mathbb{P}(X_n = x_n \,|\, X_{n-1} = x_{n-1}, ..., X_1 = x_1, X_0 = x_0) = \frac{\mathbb{P}(X_n = x_n, X_{n-1} = x_{n-1}, ..., X_1 = x_1, X_0 = x_0)}{\mathbb{P}(X_{n-1} = x_{n-1}, ..., X_1 = x_1, X_0 = x_0)},$$

and let us concentrate on the numerator, which we can write as

$$\mathbb{P}(X_n = x_n, ..., X_1 = x_1, X_0 = x_0) = \sum_{z_n \in \Phi^{-1}(x_n), ..., z_0 \in \Phi^{-1}(x_0)} \mathbb{P}(Z_n = z_n, ..., Z_0 = z_0)$$

and since (Z_n) is Markov, we can write this as

$$\begin{aligned} &\sum_{z_n \in \Phi^{-1}(x_n), ..., z_0 \in \Phi^{-1}(x_0)} \mathbb{P}(Z_n = z_n \,|\, Z_{n-1} = z_{n-1}) \cdots \mathbb{P}(Z_1 \,|\, Z_0 = z_0) \mathbb{P}(Z_0 = z_0) \\ &= \sum_{z_n \in \Phi^{-1}(x_n), ..., z_0 \in \Phi^{-1}(x_0)} \Pi(z_{n-1}, z_n) \cdots \Pi(z_0, z_1) \mathsf{K}(x_0, z_0), \end{aligned}$$

where we also used the fact that the initial distribution, ie the distribution of Z_0 is $\mathsf{K}(x_0, \cdot)$. Let us now compute using, first, property (i) and then the intertwining property (ii)

$$\begin{aligned} \sum_{z_0 \in \Phi^{-1}(x_0)} \Pi(z_0, z_1) \mathsf{K}(x_0, z_0) &= \sum_{z_0} \Pi(z_0, z_1) \mathsf{K}(x_0, z_0) \\ &= \sum_{z_0} \mathsf{K}(x_0, z_0) \Pi(z_0, z_1) \\ &= \sum_{\tilde{x}_0} \mathsf{P}(x_0, \tilde{x}_0) \mathsf{K}(\tilde{x}_0, z_1) \end{aligned}$$

but, again, because of assumption (i) this is equal to

$$\mathsf{P}(x_0, x_1) \mathsf{K}(x_1, z_1)$$

feeding this into the next summation, we have that

$$\begin{aligned} \sum_{z_1 \in \Phi^{-1}(x_1)} \Pi(z_1, z_2) \sum_{z_0 \in \Phi^{-1}(x_0)} \Pi(z_0, z_1) \mathsf{K}(x_0, z_1) &= \sum_{z_1 \in \Phi^{-1}(x_1)} \Pi(z_1, z_2) \mathsf{P}(x_0, x_1) \mathsf{K}(x_1, z_1) \\ &= \sum_{z_1 \in \Phi^{-1}(x_1)} \mathsf{K}(x_1, z_1) \Pi(z_1, z_2) \mathsf{P}(x_0, x_1) \\ &= \mathsf{K}(x_2, z_2) \mathsf{P}(x_1, x_2) \mathsf{P}(x_0, x_1) \end{aligned}$$

where in the last step we used the intertwining property in the same way we used it above. Repeating this procedure we obtain that

$$\begin{aligned} \mathbb{P}(X_n = x_n, ..., X_1 = x_1, X_0 = x_0) &= \sum_{z_n \in \Phi^{-1}(x_n)} \mathsf{K}(x_n, z_n)\, \mathsf{P}(x_{n-1}, x_n) \cdots \mathsf{P}(x_1, x_2) \mathsf{P}(x_0, x_1) \\ &= \mathsf{P}(x_{n-1}, x_n) \cdots \mathsf{P}(x_1, x_2) \mathsf{P}(x_0, x_1), \end{aligned}$$

where in the second equality we used assumption (i). This clearly leads to

$$\mathbb{P}(X_n = x_n \,|\, X_{n-1} = x_{n-1}, ..., X_1 = x_1, X_0 = x_0) = \mathsf{P}(x_{n-1}, x_n),$$

which is the desired Markov property.

Claim *(ii)* is left as an exercise. □

6.4. Application to RSK dynamics. (this subsection will not be examinable but please enjoy it)

We now want to apply the Pitman-Rogers theorem to show that the bottom row of the Gelfand-Tsetlin pattern (or the shape of the Young tableau) induced by the RSK dynamics is Markovian. This is not the case for any distribution on the entries of the input matrix $W = (w^i_j)$. It happens only when the entries have either a geometric distribution or an exponential distribution (the latter can be recovered as a limit of geometric random variables).

For the sake of exposition, let us assume exponential distributions. In this case and referring to (6.5), the crosses on the input lines constitute a Poisson Point Process (is the inter-arrival times has exponential distributions). As we said, and we will refresh below, the evolution of the whole pattern is Markovian and denote its transition matrix by $\Pi(Z, Z')$. To show that the bottom row of the Gelfand-Tsetlin pattern evolves markovianly, we need to find two kernels, $P(\lambda, \nu)$, where $\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N)$ and

$\nu = (\nu_1 \geq \nu_2 \geq \cdots \geq \nu_N)$ (λ, ν being partitions representing the state space of shapes) and another kernel $K(\lambda, Z)$ that intertwines P with the transition kernel Π of the whole pattern. The question is how to guess these !

The answer comes from the structure of RSK, its relations to Schur functions and the structure of Schur functions via its combinatorial definition and the Pieri rule.

- First, let us recall the RSK dynamics: on a GT pattern $Z = (z^i_j \colon 1 \leq j \leq i \leq n)$ only particles $z^i_1, i = 1, 2, ..., n$ jump of their own volition and they do so at exponential times at rate x_i, respectively. The jumps consist of one step to the right and trickle down the pattern as follows: if particle z^i_j jumps and before the jump took place we had $z^i_j = z^{i+1}_j$, then particle z^{i+1}_j is pushed one step to the right along with z^i_j. If $z^i_j < z^{i+1}_j$, then particle z^{i+1}_{j+1} is pulled one step to the right along with z^i_j. The jumps trickle down until they reach the bottom of the pattern. The Markov generator of this process can be easily written. Concretely, in the case $N = 2$, it may be written as

$$\Pi(Z, \widetilde{Z}) = x_1 \mathbb{1}_{\{(\widetilde{z}^1, \widetilde{z}^2) = (z^1+e_1, z^2+e_1)\}} \mathbb{1}_{\{z^1_1 = z^2_1\}} + x_1 \mathbb{1}_{\{(\widetilde{z}^1, \widetilde{z}^2) = (z^1+e_1, z^2+e_2)\}} \mathbb{1}_{\{z^1_1 < z^2_1\}} + x_2 \mathbb{1}_{\{(\widetilde{z}^1, \widetilde{z}^2) = (z^1, z^2+e_1)\}},$$

  where we have denoted the base vectors on $\mathbb{R}^2$ by $e_1 := (1, 0)$ and $e_2 := (0, 1)$. Observe that

$$\sum_{\widetilde{\mathsf{Z}}} \Pi(\mathsf{Z}, \widetilde{\mathsf{Z}}) = x_1 + x_2 = h_1(x_1, x_2), \tag{6.6}$$

  the complete, symmetric function of degree one, in variables x_1, x_2.
- Next let us recall the Pieri and in fact we will just need the particular case of multiplying the Schur function with $h_1(x_1, ..., x_n) = \sum_{i=1}^n x_i$. In this case, the Pieri rule may be written as

$$h_1 s_\lambda = \sum_{\nu \succ_1 \lambda} s_\nu = \sum_{i=1}^n s_{\lambda + \boldsymbol{e}_i}. \tag{6.7}$$

  where the symbol $\nu \succ_1 \lambda$ means adding one box at the end of one of the rows of the Young diagram with shape λ or in terms of Gelfand-Tsetlin parametrisation to one of the particles of its bottom row jumping one step to the right.
- Third, let us recall the combinatorial form of the Schur functions, written in Gelfand-Tsetlin variables as

$$s_\lambda(x) = \sum_{\substack{Z \colon \text{Gelfand-Tsetlin pattern} \\ \text{with shape } \lambda, \text{ i.e } z^N = \lambda}} \prod_{i=1}^N x_i^{|z^i| - |z^{i-1}|} =: \sum_{\substack{Z \colon \text{Gelfand-Tsetlin pattern} \\ \text{with shape } \lambda, \text{ i.e } z^N = \lambda}} K(\lambda, Z) \tag{6.8}$$

  We should mention that the above formula is also called **branching rule** as it also has a representation theoretic significance (it describes the *characters of the induced representations* when we restrict GL_n to GL_{n-1}). For the moment let us now expand on this but restrict to only use the terminology *branching rule.*

Let us now see how these three elements come together to produce a guess for the desired intertwining. Let us start with (6.6), then the Pieri rule and then move with the branching rule to produce the sequence of identities:

$$\sum_{Z, \widetilde{Z}} K(\lambda, Z) \Pi(Z, \widetilde{Z}) \overset{(6.6)}{=} h_1 \sum_Z K(\lambda, Z) \overset{(6.7)}{=} \sum_{i=1}^n s_{\lambda + e_i}(x) \overset{(6.8)}{=} \sum_{i=1}^n \sum_{\widetilde{Z}} K(\lambda + e_i, \widetilde{Z})$$
$$= \sum_{\nu \colon |\nu - \lambda|_1 = 1} \sum_{\widetilde{Z}} P(\lambda, \nu) \, K(\nu, \widetilde{Z}),$$

where $P(\lambda, \nu) = \mathbb{1}_{|\nu-\lambda|_1=1}$ if λ, ν are partitions. Now one can be imaginative and wonder whether the above equality still holds if one dropped the summations of $\widetilde{Z}$, ie whether

$$\sum_Z K(\lambda, Z)\Pi(Z, \widetilde{Z}) \stackrel{???}{=} \sum_{\nu:\, |\nu-\lambda|_1=1} P(\lambda, \nu)\, K(\nu, \widetilde{Z}) \tag{6.9}$$
$$\iff K\Pi(\lambda, \widetilde{Z}) \stackrel{???}{=} PK(\lambda, \widetilde{Z}),$$

which is the intertwining we seek. Once the guess is made, it only remains to check whether this is correct and this can only be done via a direct check. This turns out to be true but we will skip the computation, which is straightforward but a bit tedious (remember, mathematics have two parts: one is the inspiration and the other is the hard work !).

One thing to note, though, is that $P(\lambda, \nu)$ is *not* a probability kernel as $\sum_\nu P(\lambda, \nu)$ equals N rather than 1. The question is how to turn this into a probability kernel. The guess of dividing by N is not quite right...The correct is to consider the kernel

$$\hat{P}(\lambda, \nu) = \frac{s_\nu(x)}{s_\lambda(x)} P(\lambda, \nu). \tag{6.10}$$

The fact that this is a probability kernel is a consequence of the Pieri rule. Actually, in another language, Pieri rule can be interpreted as saying that the Schur functions are **harmonic functions** for operator P and the tranformation (6.10) is actually a particular case what is called **Doob's transform** in stochastic processes. Intertwining (6.9) can be rewritten as

$$\sum_Z \frac{1}{s_\lambda(x)} K(\lambda, Z)\Pi(Z, \widetilde{Z}) = \sum_{\nu:\, |\nu-\lambda|_1=1} \frac{s_\nu(x)}{s_\lambda(x)} P(\lambda, \nu)\, \frac{1}{s_\nu(x)} K(\nu, \widetilde{Z})$$
$$\iff \hat{K}\hat{\Pi}(\lambda, \widetilde{Z}) = \hat{P}\hat{K}(\lambda, \widetilde{Z}),$$

where

$$\hat{K}(\lambda, Z) := \frac{1}{s_\lambda(x)} K(\lambda, Z).$$

So, finally, applying the Pitman-Rogers theorem, the shape of the Young tableau induced by RSK is a Markov process with transition probability kernel

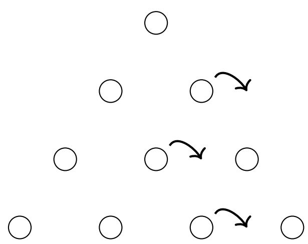
$$\hat{P}(\lambda, \nu) = \frac{s_\nu(x)}{s_\lambda(x)} P(\lambda, \nu).$$

Remark 6.6. The fact that kernel $\hat{P}$ is a probability (and also a Markovian) kernel is a consequence of the Pieri rule.

6.5. q-deformed dynamics and Macdonald functions.

This subsection is also not examinable but please enjoy it.

Let us start with a definition of dynamics on Gelfand-Tsetlin patterns



where the jump rates / probabilities will depend on the relative distance of the particles via a parameter q, which can be tuned to several limits.

Definition 6.7 (q-Whittaker $2d$ growth model). *Let $x_1, ..., x_n$ be positive numbers. Each of the particles z_j^k in a Gelfand-Tsetlin pattern $(z_j^k\colon 1 \le j \le k \le n)$ jumps, independently of others, to the right* *by one step at rate*

$$x_k \, \frac{(1-q^{z_{j-1}^{k-1}-z_j^k})(1-q^{z_j^k-z_{j+1}^k+1})}{1-q^{z_j^k-z_j^{k-1}+1}}, \tag{6.11}$$

and when it jumps it pushes along the string of particles $z_j^{k+1}, z_j^{k+2}, ...$ *with the property that* $z_j^k = z_j^{k+1} = z_j^{k+2} = ...$ *Notice that if* $z_j^k = z_{j-1}^{k-1}$ *then the jump of* z_j^k *is suppressed (the rate in this case is equal to zero), which is consistent with presrving the interlacing property. We implicitly use the convention that terms which contain particles that are not included in the Gelfand-Tsetlin pattern are omitted from expression* (6.11)*.*

Let us remark that the q-Whittaker dynamics are different than the dynamics induced by RSK. This is because in the latter the independent jumps only take place on the diagonal z_1^k with $k = 1, ..., n$ and the jumps propagate to the rest of the Gelfand-Tsetlin pattern, while in the q-Whittaker dynamics each particle has its own independent exponential clock that initiates jumps. In fact, the above dynamics are a deformation of what is known as Warren dynamics (from Jon Warren of Warwick !) [[W07]. The above deformation was introduce in [BC14]. The name q-Whittaker comes from the special functions that are involved in these dynamics and q-deform Schur functions (but also the Whittaker functions).

Notice that the rates of particles $(z_k^k \colon k = 1, ..., n)$ are just given by $x_k(1 - q^{z_{k-1}^{k-1}-z_k^k})$, which means that the evolution $(z_k^k \colon k = 1, ..., n)$ is also Markovian: it is a q deformation of TASEP, called q-TASEP. The probability of a particle to jump depends on its distance to the following particle. Thus, we see again that particle z_n^n has a double nature: on the one hand that of the smallest particle in a string of q-TASEP and on the other that of the smallest particle in a Dyson-like process.

But how does one come up with such crazy dynamics ? Macdonald polynomials were defined by Macdonald in [M88], see also [M98], as a family of symmetric polynomials, depending on two parameters q, t in a way that they degenerate, in certain limits of q, t to several other families of symmetric polynomials that includes Schur, Hall-Littlewood, Jack, zonal etc. The way that Macdonald defined these functions was via the inner product approach and bi-orthgonalisation, similar to the third definition of Schur functions, see Section 3.5. In particular, Macdonald proved the following theorem

Theorem 6.8 ([M88])**.** *Consider the inner product* $\langle \cdot, \cdot \rangle_{(q,t)}$ *defined via its values on power symmetric polynomials as*

$$\langle p_\lambda, p_\mu \rangle_{(q,t)} = z_\lambda(q,t)\delta_{\lambda,\mu}, \qquad \text{with} \qquad z_\lambda(q,t) := z_\lambda \prod_{i=1}^{\ell(\lambda)} \frac{1-q^{\lambda_i}}{1-t^{\lambda_i}}, \tag{6.12}$$

and z_λ *as in* (2.1)*. Then, for each partition* λ*, there exists a unique symmetric function* $P_\lambda(x_1, x_2, ...) = P_\lambda(x_1, x_2, ...; q, t)$ *such that*

$$\text{(A)} \qquad P_\lambda = m_\lambda + \sum_{\mu<\lambda} u_{\lambda\,\mu}(q,t)\, m_\mu \qquad \text{and}$$

$$\text{(B)} \qquad \langle P_\lambda, P_\mu \rangle_{(q,t)} = 0, \qquad \text{if} \qquad \lambda \neq \mu.$$

In (A) *the coefficients* $u_{\lambda\,\mu}(q,t)$ *are rational functions in* q, t*.*

We notice that when $q = t$, then $z_\lambda(q,q) = z_\lambda$ and, thus, the inner product in (6.12) becomes the same as in the Schur case. So in this case, by the uniqueness, the Macdoland polynomials are identical to the Schur. When $t = 0$, we talk about the q-Whittaker function (the name because they provide a q-deformation of the Whittaker functions. Macodnald functions are important because they interpolate in-between many other distinguished special functions that appear in many fields of mathematics (from mathematical physics, representation theory, number theory and even statistics). We refer to the diagram in Figure 1 for a glimpse. Macdonald are eigenfunctions of the different operator:

Macdonald-diag.png

FIGURE 1. The above diagram shows possible degenerations of Macdonald functions and related integrable stochastic processes. The image is taken from slides of a talk by Borodin, `https://math.temple.edu/events/seminars/grosswald/past/oct2015/IPLecture2and3.pdf`

$$D = \sum_{i=1}^{n} \Big(\prod_{i \neq j} \frac{t x_i - x_j}{x_i - x_j} \Big) T_{q,x_i}$$

with the operator T_{q,x_i}, for $i = 1, ..., n$, defined via its action on a function $f(x_1, ..., x_n)$ as

$$(T_{q,x_i} f)(x_1, ..., x_n) = f(x_1, ..., q x_i, ..., x_n).$$

The eigenvalue $c_{\lambda\lambda}$ is also explicit and given by $q^{\lambda_1} t^{n-1} + q^{\lambda_2} t^{n-2} + \cdots + q^{\lambda_n}$, see [M98], VI (4.15).

Macdonald polynomials satisfy the Cauchy identity:

$$\sum_{\lambda} P_\lambda(x; q, t) Q_\lambda(y; q, t) = H(x, y) \qquad \text{with}$$

$$H(x, y) := H(x, y; q, t) := \prod_{i,j} \frac{(t x_i y_j; q)_\infty}{(x_i y_j; q)_\infty},$$

where in the last expression $(a; q)_\infty := \prod_{i=0}^{\infty} (1 - a q^i)$ is the q-**Polchammer symbol** and $Q_\lambda(\cdot; q, t)$ are the dual Macdoland polynomials:

$$Q_\lambda(\cdot; q, t) := \frac{P_\lambda(\cdot; q, t)}{\langle P_\lambda(\cdot; q, t), P_\lambda(\cdot; q, t) \rangle_{(q,t)}}.$$

Moreover, Macdonald polynomials satisfy a Pieri identity:

$$P_\mu e_r = \sum_{\lambda/\mu \text{ is a vertical } r \text{ strip}} \psi'_{\lambda/\mu} P_\lambda, \tag{6.13}$$

with

$$\psi'_{\lambda/\mu} = \prod_{i<j:\, \lambda_i = \mu_i,\, \lambda_j = \mu_j + 1} \frac{(1 - q^{\mu_i - \mu_j} t^{j-i-1})(1 - q^{\lambda_i - \lambda_j} t^{j-i+1})}{(1 - q^{\mu_i - \mu_j} t^{j-i})(1 - q^{\lambda_i - \lambda_j} t^{j-i})},$$

for λ/μ a vertical r-strip. The Pieri rule (for $t = 0$) is what motivates the dynamics in the q-Whittaker model.

7. The Murnaghan-Nakayama rule: Schur functions and characters of the symmetric group.

We want in this section to prove formula (2.5) which represents the characters of the symmetric group in terms of (border-strip) Young tableaux. We will do this exploiting certain relations with Schur functions. Via these relations, we will also provide a different derivation of the orthogonality of the characters of the symmetric group; a fact that we have seen it is a general property of characters. The text reference here is [S23], Chapter 7.17.

To start, recall the notion of a *border strip tableau* from Section 2.4.1. The starting theorem is the following, which should be viewed as an analogue of the Pieri rule:

Theorem 7.1. *Let μ be a partition, $r \in \mathbb{N}$ and p_r the power symmetric polynomials of order r. Then*

$$s_\mu p_r = \sum_{\lambda:\ \lambda/\mu \text{ border-strip tableau of size } r} (-1)^{\mathrm{ht}(\lambda/\mu)} s_\lambda$$

Proof. We will use the determinantal formula for Schur functions. Let us denote by δ the partition $\delta := (n-1, n-2, ..., 1)$ and for a partition λ the determinant

$$a_\lambda(x) := \det\left(x_i^{\lambda_j}\right) \tag{7.1}$$

In particular, in this notation, we have that

$$s_\lambda(x) = \frac{\det\left(x_i^{\lambda_j+n-j}\right)_{1\le i,j\le n}}{\det\left(x_i^{n-j}\right)_{1\le i,j\le n}} = \frac{a_{\lambda+\delta}(x)}{a_\delta(x)},$$

and so it suffices to show that

$$a_{\mu+\delta}(x)\, p_r = \sum_{\lambda:\ \lambda/\mu \text{ border-strip tableau of size } r} (-1)^{\mathrm{ht}(\lambda/\mu)} a_\lambda(x)$$

Starting from the left-hand side, we have that

$$\begin{aligned} a_{\mu+\delta}(x) p_r &= \sum_{\sigma\in S_n} sgn(\sigma) \prod_{i=1}^n x_i^{\mu_{\sigma(i)}+n-\sigma(i)} \cdot \sum_{k=1}^n x_k^r \\ &= \sum_{k=1}^n \sum_{\sigma\in S_n} sgn(\sigma) \prod_{i=1}^n x_i^{\mu_{\sigma(i)}+n-\sigma(i)} x_k^r \\ &= \sum_{k=1}^n \sum_{\sigma\in S_n} sgn(\sigma) \prod_{i=1}^n x_i^{\mu_{\sigma(i)}+n-\sigma(i)} \prod_{i=1}^n x_i^{r\delta_{k,i}}, \end{aligned}$$

where in the last $\delta_{k,i}$ is the Kronecker delta. We can continue as

$$\begin{aligned} a_{\mu+\delta}(x) p_r &= \sum_{k=1}^n \sum_{\sigma\in S_n} sgn(\sigma) \prod_{i=1}^n x_i^{\mu_{\sigma(i)}+n-\sigma(i)+r\delta_{k,i}} \\ &= \sum_{k=1}^n \det\left(x_i^{\mu_j+n-j+r\delta_{k,i}}\right)_{1\le i,j\le n}. \end{aligned}$$

Now, it actually turns out that the latter can also be written as

$$\sum_{k=1}^n \det\left(x_i^{\mu_j+n-j+r\delta_{k,j}}\right)_{1\le i,j\le n}, \tag{7.2}$$

(note the change from $\delta_{k,i}$) to $\delta_{k,j}$! (Justify this...)

We now want to rearrange the order of the columns in the determinant, in order to write in the form (7.1), thus, resembling the numerator in the determinantal formula for Schur functions. For this, we need to reorder the sequence

$$(\mu_1 + n - 1\,,\ ...,\underbrace{\mu_k + n - k + r}_{k-\text{position}}, ...,\ \mu_n).$$

in descending order, since in the definition (7.1), λ needs to be a partition. If the k^{th} position turns out to be equal to another entry of the vector, then the determinant will be zero. If not, then this entry will move forward a number of places. In particular, if ℓ is largest number such that

$$\mu_{\ell-1} + n - (\ell - 1) > \mu_k + n - k + r,$$

then we will have the vector

$$(\mu_1 + n - 1\,, ..., \mu_{\ell-1} + n - (\ell - 1)\,, \underbrace{\mu_k + n - k + r}_{\ell-\text{position}}\,, \mu_\ell + n - \ell, ..., \mu_{k-1} + n - (k - 1)\,, \mu_k + n - k, ..., \mu_n).$$

whose coordinates are ordered in decreasing fashion. We want to represent this vector in the form $\lambda + \delta = (\lambda_1 + n - 1\,, \lambda_2 + n - 2\,, ..., \lambda_n)$, for a partition λ. This will be the case if we set

$$(\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n) = (\mu_1 \geq ... \geq \mu_{\ell-1} \geq \mu_k + \ell - k + r \geq \mu_\ell + 1 \geq ... \geq \mu_{k-1} + 1 \geq \mu_{k+1} \geq ... \geq \mu_n).$$

In this way

$$\begin{aligned} \lambda/\mu &= (\lambda_1 - \mu_1, \lambda_2 - \mu_2, ..., \lambda_n - \mu_n) && (7.3)\\ &= (0, ..., 0, \underbrace{\mu_k - \mu_\ell + \ell - k + r}_{\ell-\text{position})}, \mu_\ell - \mu_{\ell+1} + 1, ..., \mu_{k-1} - \mu_k + 1, 0, ..., 0). && (7.4) \end{aligned}$$

This is a border strip (check !) that starts from row ℓ and goes to row k, hence having height $\ell - k$, and the number of boxes it has is

$$(\mu_k - \mu_\ell + \ell - k + r) + (\mu_\ell - \mu_{\ell+1} + 1) + \cdots + (\mu_{k-1} - \mu_k + 1) = r.$$

Coming back to (7.2), we have that

$$\begin{aligned} a_{\mu+\delta} p_r &= \sum_{k=1}^{n} \det\left(x_i^{\mu_j+n-j+r\delta_{k,j}}\right)_{1\leq i,j\leq n} \\ &= \sum_{k=1}^{n} (-1)^{\ell-k} a_{\lambda+\delta} \\ &= \sum_{k=1}^{n} (-1)^{\mathrm{ht}(\lambda/\mu)} a_{\lambda+\delta}, \end{aligned}$$

with λ being (7.3), where the factor $(-1)^{\ell-k}$ is the result of the commutation of the columns in the first determinant, above, that we performed. Dividing by the Vandermonde determinant a_δ, we get

$$s_\mu p_r = \sum_{k=1}^{n} (-1)^{\mathrm{ht}(\lambda/\mu)} s_\lambda,$$

and this leads to the desired result as the summation over k indicates the possible lower end of the border strip. □

We want to extend the above theorem to the case of $s_\mu p_\alpha$, where $\alpha = (\alpha_1, \alpha_2, ...)$ with α_i being non-negative integers, in which case $p_\alpha := p_{\alpha_1} p_{\alpha_2} \cdots$. For this, we will need to extend the notion of border-strip to that of *border-strip tableau* of type $\alpha = (\alpha_1, \alpha_2, ...)$. This will be a Young diagram where

- every column and row have weakly increasing entries,
- integer i appears α_i times,
- the set of squares occupied by i forms a border-strip.

An example is

1	1	1	1	6	6	6
1	2	2	5	6		
3	3	5	5	6		
3	5	5	6	6		

We can, now, state the generalisation of Theorem 7.1:

Theorem 7.2. *If μ is a partition and $\alpha = (\alpha_1, \alpha_2, ...)$, then*

$$s_\mu p_\alpha = \sum_\lambda \chi^{\lambda/\mu}(\alpha)\, s_\lambda, \qquad \text{where} \tag{7.5}$$

$$\chi^{\lambda/\mu}(\alpha) = \sum_{T:\ \text{border-strip tableau of shape } \lambda/\mu \text{ and type } \alpha} (-1)^{\mathrm{ht}(T)}. \tag{7.6}$$

The notation $\chi^{\lambda/\mu}$ is suggestive that these will be character. In the particular case that $\mu = \emptyset$ we have that

$$p_\alpha = \sum_\lambda \chi^\lambda(\alpha)\, s_\lambda. \tag{7.7}$$

Proof. We write $p_\alpha = p_{\alpha_1} p_{\alpha_2} \cdots$, and then apply successively Theorem 7.1 to $\Big((s_\mu p_{\alpha_1}) p_{\alpha_2}\Big) \cdots$. □

We can now state the formula for the characters of the symmetric group

Theorem 7.3. *The characters $\chi^\lambda(\alpha)$ of the symmetric group corresponding to irreducible representations indexed by λ is given by the formula*

$$\chi^\lambda(\alpha) = \sum_{T:\ \text{border-strip tableaux of shape } \lambda \text{ and type } \alpha} (-1)^{\mathrm{ht}(T)}.$$

(Sketch of the) Proof - (obviously not examinable). This is a consequence of the Schur-Weyl duality in representation theory, a deep result which says that the tensor product $V^{\otimes n}$ of a vector space V can be decomposed as a representation of $S_n \times \mathrm{GL}(V)$ as

$$V^{\otimes n} = \oplus_{|\lambda|=n} V_\lambda \otimes \mathbb{S}_\lambda V,$$

where V_λ are the irreducible representations of S_n and $\mathbb{S}_\lambda V$ are the irreducible representations of $\mathrm{GL}(V)$. Taking the trace in the above we obtain

$$p_\alpha = \sum_\lambda \chi^\lambda(\alpha)\, s_\lambda. \tag{7.8}$$

where $\chi^\lambda(\alpha)$ are the characters of irreducibles of S_n parametrised by λ. But in (7.7) we proved that this relation holds for $\chi^\lambda(\alpha)$ being as in (7.6), from which the assertion follows as Schur functions form a basis of the symmetric functions. □

We can now re-derive the orthogonality of characters of the symmetric group:

Theorem 7.4. *We have that*

$$\sum_\lambda \chi^\lambda(\alpha)\, \chi^\lambda(\beta) = z_\alpha \delta_{\alpha,\beta}. \tag{7.9}$$

for $z_\alpha = \prod_{i \ge 1} i^{\alpha_i} \alpha_i!$

Proof. Use identity (7.7) and compute

$$\begin{aligned} z_\alpha\,\delta_{\alpha,\beta} &= \langle p_\alpha, p_\beta\rangle \\ &= \sum_{\lambda,\mu} \chi^\lambda(\alpha)\chi^\mu(\beta)\langle s_\lambda(\alpha), s_\mu(\beta)\rangle \\ &= \sum_{\lambda,\mu} \chi^\lambda(\alpha)\chi^\mu(\beta)\delta_{\lambda,\mu} \\ &= \sum_{\lambda} \chi^\lambda(\alpha)\chi^\lambda(\beta). \end{aligned}$$

□

We also have another orthogonality relation, which reads as follows:

Theorem 7.5. *The following relation holds*

$$\sum_\alpha \frac{1}{z_\alpha}\chi^\lambda(\alpha)\chi^\mu(\alpha) = \delta_{\lambda,\mu}. \tag{7.10}$$

Proof. We can also get the above identity via an identity dual to (7.7), which expresses the Schur functions in a power symmetric function expansion as

$$s_\lambda = \sum_\alpha \frac{1}{z_\alpha}\chi^\lambda(\alpha)p_\alpha. \tag{7.11}$$

Relation (7.11) follows from (7.7) by taking inner products, which implies that

$$s_\lambda(\alpha) = \langle p_\alpha, \chi^\lambda(\alpha)\rangle,$$

and using the orthogonality relation $\langle p_\alpha, p_\beta\rangle = z_\alpha\delta_{\alpha,\beta}$, we obtain

$$s_\lambda(\alpha) = \sum_\alpha \frac{1}{z_\alpha}\langle p_\alpha, \chi^\lambda(\alpha)\rangle p_\alpha = \sum_\alpha \frac{1}{z_\alpha}\chi^\lambda(\alpha)p_\alpha.$$

We can now obtain (7.10) as

$$\begin{aligned} \delta_{\lambda,\mu} &= \langle s_\lambda, s_\mu\rangle \\ &= \langle \sum_\alpha \frac{1}{z_\alpha}\chi^\lambda(\alpha)p_\alpha, \sum_\beta \frac{1}{z_\beta}\chi^\mu(\beta)p_\beta\rangle \\ &= \sum_{\alpha,\beta} \frac{1}{z_\alpha\, z_\beta}\chi^\lambda(\alpha)\chi^\mu(\beta)\langle p_\alpha, p_\beta\rangle \\ &= \sum_{\alpha,\beta} \frac{1}{z_\alpha\, z_\beta}\chi^\lambda(\alpha)\chi^\mu(\beta)\, z_\alpha\delta_{\alpha,\beta} \\ &= \sum_\alpha \frac{1}{z_\alpha}\chi^\lambda(\alpha)\chi^\mu(\alpha). \end{aligned}$$

□

Remark 7.6. *The orthogonality relations* (7.9) *and* (7.10) *that we obtained through the Murnanghan-Nakayama rule are the interpretation of the two fundamental character identities (orthogonality) that is valid for general groups. In Theorem 2.21 we showed that for the inner product*

$$\langle \chi, \psi\rangle := \frac{1}{|G|}\sum_{g\in G}\chi(g)\psi(g^{-1}),$$

we have the orthogonality relation $\langle \chi, \psi\rangle = \delta_{\chi,\psi}$ for irreducible characters χ, ψ. Relation (7.10) *can be seen as a version of this orthogonality relation (do the matching !)*

Relation (7.9) can be seen as a version of the second character relation, which we didn't prove and which reads as

$$\sum_{\chi:\ \text{irreducible characters of } G} \chi(K)\overline{\chi(L)} = \frac{|G|}{|K|}\delta_{K,L},$$

for K, L conjugacy classes of G. In the particular case of S_n, the conjugacy classes of permutations are determined by the type α of the permutation, the irreducible characters χ^λ are indexed by partitions λ. Moreover, we have proved in (2.2) that for conjugacy class K_α corresponding to a permutation of type α, we have that $\frac{|S_n|}{|K_\alpha|} = z_\alpha$. From these considerations, relation

$$\sum_\lambda \chi^\lambda(\alpha)\,\chi^\lambda(\beta) = z_\alpha \delta_{\alpha,\beta}.$$

that we proved in Theorem 7.4, follows.

8. Supplementary exercises

Exercise 21. *Let $x = (x_1, x_2, ...)$ be a sequence of indeterminates and for i, j integers, denote by $x^{(i,j]} := (0, ..., 0, x_{i+1}, x_{i+2}, ..., x_j, 0, 0, ...)$. Let $e_n(x), h_n(x)$ be the elementary and complete homogeneous symmetric functions, respectively. Show that for any i, j, N such that $i \le N, j \le N$, it holds that*

$$\sum_{r=0}^{n} (-1)^n e_n\big(x^{(i,N]}\big) h_{n-r}\big(x^{(j,N]}\big) = \begin{cases} h_n\big(x^{(j,i]}\big)(x), & \text{if } j \le i, \\ \\ (-1)^n e_n\big(x^{(i,j]}\big), & \text{if } i \le j. \end{cases}$$

Remark, in particular, what happens in the case $i = j$.

Exercise 22. *Show the Littlewood identity:*

$$\sum_\lambda s_\lambda(x) = \prod_i \frac{1}{1-x_i} \prod_{i<j} \frac{1}{1-x_i x_j}.$$

Exercise 23. *Let a permutation $\sigma \in S_n$ be an involution, i.e. $\sigma^2 = Id$. Show that the number of involutions is equal to the number of standard Young tableaux with n boxes.*

Exercise 24. *Let $p(n, k)$ be the number of partitions of n, which have k parts. Show that*

$$\sum_{n,k\ge 0} p(n,k) t^k x^n = \prod_{i\ge 1} \frac{1}{1-x^i t},$$

and that also

$$\sum_{n,k\ge 0} p(n,k) t^k x^n = \sum_{n\ge 1} \frac{x^n t^n}{(1-x)(1-x^2)\cdots(1-x^n)}.$$

Exercise 25. **A**. *Let $e_k(n) := e_k(x_1, ..., x_n)$ for indeterminates $x_1, x_2, ...$ and $h_k(n) := h_k(x_1, ..., x_n)$. Show that they satisfy the following recursions:*

$$e_k(n) = e_k(n-1) + x_n e_{k-1}(n-1), \qquad e_k(0) = \delta_{k,0},$$

and

$$h_k(n) = h_k(n-1) + x_n h_{k-1}(n), \qquad h_k(0) = \delta_{k,0}.$$

B. *Define the Stirling number of the first kind as*

$$c(n,k) := \text{the number of } \pi \in S_n \text{ with } k \text{ disjoint cycles},$$

and the Stirling numbers of the second kind as

$$S(n,k) = \text{the number of partitions of the set } \{1, ..., n\} \text{ into } k \text{ subsets}.$$

Show that the following recursions hold

$$c(n,k) = c(n-1,k-1) + (n-1)c(n-1,k), \qquad c(0,k) = \delta_{k,0}$$

and

$$S(n,k) = S(n-1,k-1) + kS(n-1,k), \qquad c(0,k) = \delta_{k,0}.$$

C. *Show the following identities*

$$\text{(i)} \qquad \binom{n}{k} = e_k(\underbrace{1,...,1}_{n\ times}) = h_k(\underbrace{1,...,1}_{n-k+1\ times}),$$

$$\text{(ii)} \qquad c(n,k) = e_{n-k}(1,2,...,n-1),$$

$$\text{(iii)} \qquad S(n,k) = h_{n-k}(1,2,...,k).$$

References

[BK95] A. Berenstein, A.N. Kirillov, Groups generated by involutions, Gelfand-Tsetlin patterns, and combinatorics of Young tableaux, *Algebra i Analiz* 7 (1995) 92–152.

[B11] A. Borodin, Determinantal point processes, In: *Oxford handbook of random matrix theory.* Oxford University Press , New York, pp. 231-249, (2011)

[BC14] A. Borodin, I. Corwin, Macdonald processes, *Probability Theory and Related Fields*, 158(1-2), 225-400, (2014).

[BG16] A. Borodin, V. Gorin. Lectures on integrable probability. *Proc. of Symposia in Prue Math.* Vol. 91, edited by V. Sidoravicius and S. Smirnov (2016)

[BP14] A. Borodin, L. Petrov, Integrable probability: From representation theory to Macdonald processes, *Prob. Surv.* 11 (2014): 1-58.

[BP16a] A. Borodin, L. Petrov, Lectures on Integrable probability: Stochastic vertex models and symmetric functions,

[J05] K. Johansson, Random matrices and determinantal processes, *Lectures notes, Les Houches*, (2005), arXiv:math-ph/0510038

[K01] A. N. Kirillov, Introduction to tropical combinatorics, *Physics and Combinatorics. Proc. Nagoya 2000 2nd Internat. Workshop* (A. N. Kirillov and N. Liskova, eds.), World Scientific, Singapore, 2001, pp. 82–150.

[M88] I. G. Macdonald, A new class of symmetric functions. *Publ. IRMA Strasbourg*, 372, 131-171, (1988).

[M98] I. G. Macdonald, Symmetric functions and Hall polynomials, *Oxford university press*, (1998).

[NY04] M. Noumi, Y. Yamada, Tropical Robinson-Schensted-Knuth correspondence and birational Weyl group actions, *Representation theory of algebraic groups and quantum groups*, 40, 371-442, (2004).

[O01] A. Okounkov, Infinite wedge and random partitions, *Selecta Math.*,**7**:57–81 (2001)

[RP81] L. C. G. Rogers, J. W. Pitman, Markov functions. *Ann. Prob.*, 573-582, (1981).

[S13] B.E. Sagan, The symmetric group: representations, combinatorial algorithms, and symmetric functions, *Springer Science & Business Media*, Vol. 203, (2013).

[S79] B. Simon, Trace ideals and their applications. Vol. 35. *C*ambridge: Cambridge University Press, 1979.

[S23] R. Stanley, Enumerative Combinatorics: Volume 2, *Cambridge University Press*, (2023).

[[W07] J. Warren, Dyson's Brownian motions, intertwining and interlacing, E. J. Prob. 12 (2007), 573–590.

[Z22] N. Zygouras, Some algebraic structures in KPZ universality *Probability Surveys* 19, 590-700, (2022)

Department of Statistics, University of Warwick, Coventry CV4 7AL, UK
Email address: `N.Zygouras@warwick.ac.uk`